

**PONTIFICIA UNIVERSIDAD  
CATÓLICA DEL PERÚ**

**FACULTAD DE CIENCIAS SOCIALES**



El impacto del tiempo de desplazamiento a las instituciones educativas en el  
rendimiento académico de los estudiantes del Perú: Un estudio comparativo del  
nivel educativo primario y secundario

Tesis para obtener el título profesional de Licenciado en Economía presentado

por:

Isla Gavonel, Ayar Antonio

Asesor:

León Jara Almonte, Juan Jesús Martín

Lima, 2025

## Informe de Similitud

Yo, Leon Jara Almonte, Juan Jesus Martin, docente de la Facultad de Ciencias Sociales de la Pontificia Universidad Católica del Perú, asesor(a) de la tesis/trabajo de investigación titulado El impacto del tiempo de desplazamiento a las instituciones educativas en el rendimiento académico de los estudiantes del Perú: Un estudio comparativo del nivel educativo primario y secundario, del/de la autor(a) Isla Gavonel, Ayar Antonio, dejo constancia de lo siguiente:

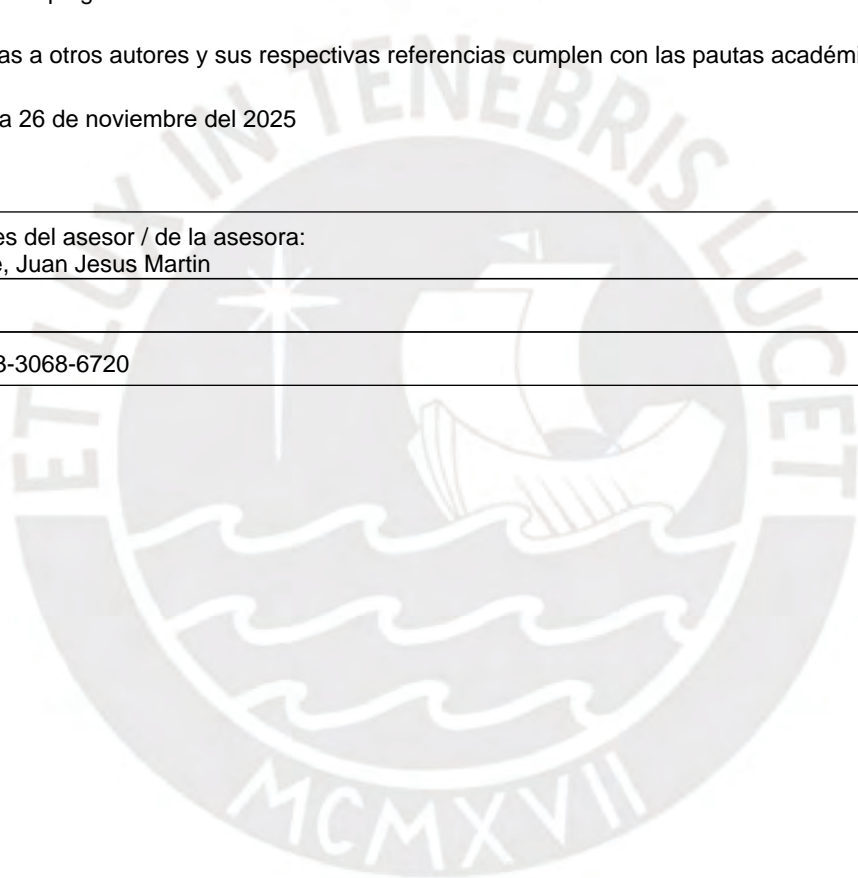
- El mencionado documento tiene un índice de puntuación de similitud de 10%. Así lo consigna el reporte de similitud emitido por el software Turnitin el 26/11/2025.
- He revisado con detalle dicho reporte y la Tesis o Trabajo de Suficiencia Profesional, y no se advierte indicios de plagio.
- Las citas a otros autores y sus respectivas referencias cumplen con las pautas académicas.

Lugar y fecha: Lima 26 de noviembre del 2025

Apellidos y nombres del asesor / de la asesora: Leon Jara Almonte, Juan Jesus Martin
---

DNI: 10548854
---------------

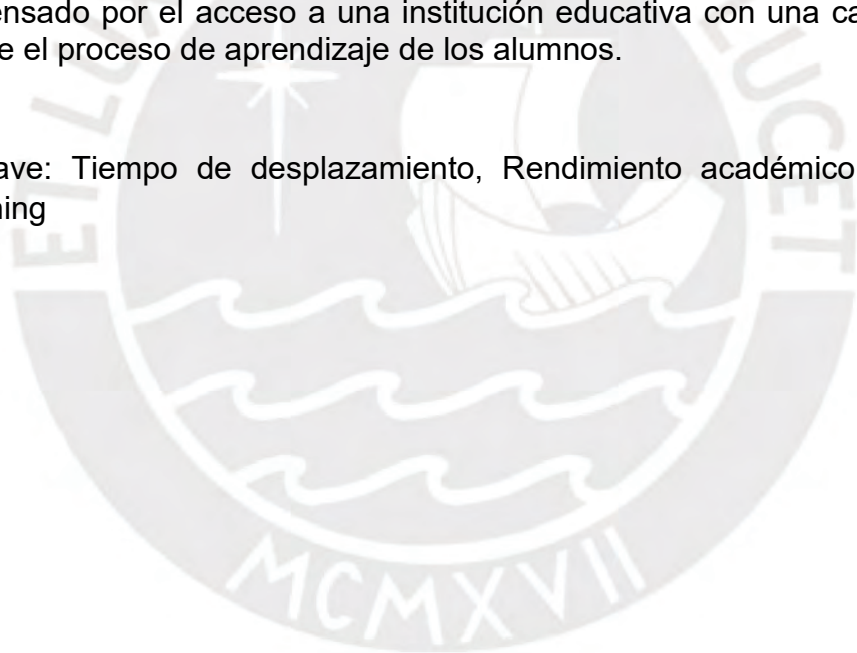
ORCID: 0000-0003-3068-6720
----------------------------



## Resumen

Con la meta de la cobertura universal de la Educación Básica Regular en Perú, la oferta educativa nacional ha estado progresando para satisfacer las necesidades escolares de los niños y sus familias. No obstante, la desigualdad de la mayor oferta educativa de escuelas de nivel primario frente a las de nivel secundario incita a analizar si esta diferencia en el mercado educativo puede afectar negativamente el rendimiento académico de los alumnos, particularmente mediante el aumento del tiempo de desplazamiento debido a la eventual transición entre los niveles educativos mencionados durante la vida escolar. De esta manera, el objetivo principal del presente trabajo es identificar cómo afecta el tiempo de desplazamiento de ida desde la residencia hacia la institución educativa al rendimiento académico de los alumnos en el nivel primario y secundario en Perú. La correspondiente hipótesis planteada indica que el impacto es negativo para ambos niveles, aunque siendo más profundo para los de secundaria. Mediante el uso de la base de datos Niños del Milenio y la técnica de Propensity Score Matching con covariables en diferentes periodos, los resultados muestran que el tiempo de desplazamiento no afecta significativamente al rendimiento de los alumnos de secundaria, mientras que para primaria este impacto es positivo y significativo, particularmente para los que residen en zonas rurales. Estas estimaciones indicarían que para ciertas familias un mayor tiempo de desplazamiento sería compensado por el acceso a una institución educativa con una calidad mayor que beneficie el proceso de aprendizaje de los alumnos.

Palabras clave: Tiempo de desplazamiento, Rendimiento académico, Propensity Score Matching



## Índice de Contenido

Introducción .....	1
Capítulo 1. Marco teórico y Revisión de Literatura .....	4
1.1 La Oferta educativa .....	4
1.2 Tiempo de desplazamiento.....	7
1.3 Rendimiento académico .....	7
1.4 Modelo Económico .....	9
1.5 El efecto del tiempo de desplazamiento en el rendimiento académico .....	11
Capítulo 2. Objetivos e hipótesis de la investigación .....	16
Capítulo 3. Metodología .....	17
3.1 Descripción de la Base de Datos.....	17
3.2 Descripción del modelo de estimación .....	18
3.3 Descripción de las variables del modelo de estimación .....	21
Capítulo 4. Hechos estilizados .....	28
4.1 Análisis de covariables utilizadas en la ronda 3 .....	29
4.1.1 Covariables relacionadas a las características del alumno .....	29
4.1.2 Covariables relacionadas a características de la familia del alumno .....	32
4.1.3 Covariables de la oferta educativa y habilidades innatas del alumno.....	37
4.2 Análisis de covariables utilizadas en la ronda 5 .....	39
4.2.1 Covariables relacionadas a las características del alumno .....	40
4.2.2 Covariables relacionadas a características de la familia del alumno .....	43
4.2.3 Covariables de la oferta educativa y habilidades innatas del alumno.....	48
4.3 Análisis previo del tiempo de desplazamiento y rendimiento académico .....	51
Capítulo 5. Resultados de las estimaciones principales .....	54
5.1 Prueba de matemáticas de los alumnos de primaria (ronda 3) .....	54
5.1.1 Resultados con el punto de corte de 20 minutos .....	54
5.1.2 Resultados con los puntos de corte de 25 y 30 minutos .....	57
5.2 Prueba de matemáticas de los alumnos de secundaria (ronda 5).....	59
5.2.1 Resultados con el punto de corte de 20 minutos .....	59
5.2.2 Resultados con los puntos de corte de 25 y 30 minutos .....	62
Capítulo 6. Interpretación de los resultados principales .....	64
Capítulo 7. Análisis de sensibilidad .....	66
7.1 Uso de diferentes métodos de emparejamiento .....	66
7.2 Estimación de resultados por zona de residencia y análisis de linealidad.....	69
Conclusiones .....	84

Limitaciones.....	86
Referencias Bibliográficas: .....	87
Anexos .....	92
Anexo A: Regresiones logit para el cálculo del Propensity score con los puntos de corte de 20, 25 y 30 minutos en las rondas 3 y 5 (5 vecinos más cercanos y caliper 0.005) .....	92
Anexo B: Tablas y gráficos de test de medias, reducción de sesgo y soporte común para los alumnos de primaria en la ronda 3 con los cortes de 25 y 30 minutos (5 vecinos más cercanos y caliper 0.005).....	96
Anexo C: Tablas y gráficos de test de medias, reducción de sesgo y soporte común para los alumnos de secundaria en la ronda 5 con los cortes de 25 y 30 minutos (5 vecinos más cercanos y caliper 0.005).....	101
Anexo D: Tablas de test de medias de covariables entre alumnos con viajes largos y cortos (control) en la ronda 3 con los cortes de 25 y 30 minutos utilizando emparejamientos de 5 vecinos más cercanos (caliper de 0.005) y kernel epanechnikov (alumnos de primaria rurales y urbanos) .....	106
Anexo E: Tablas de test de medias de covariables entre alumnos con viajes largos y cortos (control) en la ronda 5 con los cortes de 25 y 30 minutos utilizando emparejamientos de 5 vecinos más cercanos (caliper de 0.005) y kernel epanechnikov (alumnos de secundaria rurales y urbanos).....	110
Anexo F: Tablas de estimaciones del ATT sin bootstrapping en la ronda 3 y 5 con emparejamientos de 5 vecinos más cercanos (caliper de 0.005) y kernel epanechnikov (alumnos de primaria y secundaria según zona de residencia).....	114
Anexo G: Tablas de test de medias de covariables entre alumnos con viajes largos y cortos (control) en las rondas 3 y 5 con los puntos de corte de 25 y 30 minutos, emparejamiento de 5 vecinos más cercanos (caliper de 0.005) y dividiendo los alumnos con viajes largos a partir del intervalo de 40 minutos (alumnos de primaria y secundaria).....	116
Anexo H: Tablas de estimaciones del ATT sin bootstrapping en las rondas 3 y 5 con emparejamiento de 5 vecinos más cercanos (caliper de 0.005) dividiendo los alumnos con viajes largos a partir del intervalo de 40 minutos (alumnos de primaria y secundaria).....	121

## Índice de tablas

Tabla 1: Cantidad de instituciones educativas nacionales de nivel primario y secundario según zona de residencia y gestión (2023).....	4
Tabla 2: Covariables de emparejamiento para la ronda 3 (alumnos de primaria) ....	23
Tabla 3: Covariables de emparejamiento para la ronda 5 (alumnos de secundaria) 25	
Tabla 4: Test de medias de covariables entre alumnos con viajes largos y cortos (control) en la ronda 3 con el punto de corte de 20 minutos (alumnos de primaria) .	55
Tabla 5: Soporte común en el emparejamiento en la ronda 3 con el punto de corte de 20 minutos (alumnos de primaria) .....	56
Tabla 6: Estimación del ATT en la ronda 3 con el punto de corte de 20 minutos, emparejamiento de 5 vecinos más cercanos y caliper de 0.005 (alumnos de primaria) .....	57
Tabla 7: Estimación del ATT en la ronda 3 con el punto de corte de 25 minutos, emparejamiento de 5 vecinos más cercanos y caliper de 0.005 (alumnos de primaria) .....	58
Tabla 8: Estimación del ATT en la ronda 3 con el punto de corte de 30 minutos, emparejamiento de 5 vecinos más cercanos y caliper de 0.005 (alumnos de primaria) .....	58
Tabla 9: Test de medias de covariables entre alumnos con viajes largos y cortos (control) en la ronda 5 con el punto de corte de 20 minutos (alumnos de secundaria) .....	59
Tabla 10: Soporte común en el emparejamiento en la ronda 5 con el punto de corte de 20 minutos (alumnos de secundaria).....	61
Tabla 11: Estimación del ATT en la ronda 5 con el punto de corte de 20 minutos, emparejamiento de 5 vecinos más cercanos y caliper de 0.005 (alumnos de secundaria).....	62
Tabla 12: Estimación del ATT en la ronda 5 con el punto de corte de 25 minutos, emparejamiento de 5 vecinos más cercanos y caliper de 0.005 (alumnos de secundaria).....	63
Tabla 13: Estimación del ATT en la ronda 5 con el punto de corte de 30 minutos, emparejamiento de 5 vecinos más cercanos y caliper de 0.005 (alumnos de secundaria).....	63
Tabla 14: Efecto del tiempo de desplazamiento en los alumnos de nivel primario de viajes largos (ATT) con los cortes de 20, 25 y 30 minutos en la ronda 3 (5 vecinos más cercanos y caliper 0.005) .....	64
Tabla 15: Efecto del tiempo de desplazamiento en los alumnos de nivel secundario de viajes largos (ATT) con los cortes de 20, 25 y 30 minutos en la ronda 5 (5 vecinos más cercanos y caliper 0.005) .....	64
Tabla 16: Estimación del ATT en la ronda 3 con emparejamientos de vecinos más cercanos, emparejamientos de kernel y caliper de 0.005 (alumnos de primaria) .....	67

Tabla 17: Estimación del ATT en la ronda 3 con emparejamientos de vecinos más cercanos, emparejamientos de kernel y caliper de 0.01 (alumnos de primaria) .....	67
Tabla 18: Estimación del ATT en la ronda 5 con emparejamientos de vecinos más cercanos, emparejamientos de kernel y caliper de 0.005 (alumnos de secundaria) ..	68
Tabla 19: Estimación del ATT en la ronda 5 con emparejamientos de vecinos más cercanos, emparejamientos de kernel y caliper de 0.01 (alumnos de secundaria) ..	68
Tabla 20: Estimación del ATT empleando bootstrapping en la ronda 3 (alumnos de primaria) y 5 (alumnos de secundaria) con emparejamiento de 5 vecinos más cercanos y caliper de 0.005 .....	71
Tabla 21: Test de medias de covariables entre alumnos con viajes largos y cortos (control) en la ronda 3 con el punto de corte de 20 minutos, emparejamiento de 5 vecinos más cercanos y caliper de 0.005 (alumnos de primaria rurales) .....	72
Tabla 22: Test de medias de covariables entre alumnos con viajes largos y cortos (control) en la ronda 3 con el punto de corte de 20 minutos, emparejamiento de 5 vecinos más cercanos y caliper de 0.005 (alumnos de primaria urbanos) .....	72
Tabla 23: Test de medias de covariables entre alumnos con viajes largos y cortos (control) en la ronda 3 con el punto de corte de 20 minutos y emparejamiento de kernel epanechnikov (alumnos de primaria rurales).....	73
Tabla 24: Test de medias de covariables entre alumnos con viajes largos y cortos (control) en la ronda 3 con el punto de corte de 20 minutos y emparejamiento de kernel epanechnikov (alumnos de primaria urbanos).....	73
Tabla 25: Estimación del ATT empleando bootstrapping en la ronda 3 con emparejamiento de 5 vecinos más cercanos y caliper de 0.005 (alumnos de primaria rurales y urbanos).....	74
Tabla 26: Estimación del ATT empleando bootstrapping en la ronda 3 con emparejamiento de kernel epanechnikov (alumnos de primaria rurales y urbanos) ..	74
Tabla 27: Test de medias de covariables entre alumnos con viajes largos y cortos (control) en la ronda 5 con el punto de corte de 20 minutos, emparejamiento de 5 vecinos más cercanos y caliper de 0.005 (alumnos de secundaria rurales).....	76
Tabla 28: Test de medias de covariables entre alumnos con viajes largos y cortos (control) en la ronda 5 con el punto de corte de 20 minutos, emparejamiento de 5 vecinos más cercanos y caliper de 0.005 (alumnos de secundaria urbanos).....	76
Tabla 29: Test de medias de covariables entre alumnos con viajes largos y cortos (control) en la ronda 5 con el punto de corte de 20 minutos y emparejamiento de kernel epanechnikov (alumnos de secundaria rurales) .....	77
Tabla 30: Test de medias de covariables entre alumnos con viajes largos y cortos (control) en la ronda 5 con el punto de corte de 20 minutos y emparejamiento de kernel epanechnikov (alumnos de secundaria urbanos) .....	77
Tabla 31: Estimación del ATT empleando bootstrapping en la ronda 5 con emparejamiento de 5 vecinos más cercanos y caliper de 0.005 (alumnos de secundaria rurales y urbanos) .....	78

Tabla 32: Estimación del ATT empleando bootstrapping en la ronda 5 con emparejamiento de kernel epanechnikov (alumnos de secundaria rurales y urbanos) .....	78
Tabla 33: Test de medias de covariables entre alumnos con viajes largos menores a 40 minutos y viajes cortos (control) en la ronda 3 con el punto de corte de 20 minutos, emparejamiento de 5 vecinos más cercanos y caliper de 0.005 (alumnos de primaria) .....	80
Tabla 34: Test de medias de covariables entre alumnos con viajes largos mayores o iguales a 40 minutos y viajes cortos (control) en la ronda 3 con el punto de corte de 20 minutos, emparejamiento de 5 vecinos más cercanos y caliper de 0.005 (alumnos de primaria) .....	81
Tabla 35: Test de medias de covariables entre alumnos con viajes largos menores a 40 minutos y viajes cortos (control) en la ronda 5 con el punto de corte de 20 minutos, emparejamiento de 5 vecinos más cercanos y caliper de 0.005 (alumnos de secundaria).....	81
Tabla 36: Test de medias de covariables entre alumnos con viajes largos mayores o iguales a 40 minutos y viajes cortos (control) en la ronda 5 con el punto de corte de 20 minutos, emparejamiento de 5 vecinos más cercanos y caliper de 0.005 (alumnos de secundaria).....	82
Tabla 37: Estimación del ATT empleando bootstrapping en la ronda 3 con emparejamiento de 5 vecinos más cercanos, caliper de 0.005 y dividiendo los alumnos con viajes largos a partir del intervalo de 40 minutos (alumnos de primaria) .....	82
Tabla 38: Estimación del ATT empleando bootstrapping en la ronda 5 con emparejamiento de 5 vecinos más cercanos, caliper de 0.005 y dividiendo los alumnos con viajes largos a partir del intervalo de 40 minutos (alumnos de secundaria).....	83
Tabla 39: Regresión logit para el Propensity score con el corte de 20 minutos en la ronda 3 (alumnos de primaria).....	92
Tabla 40: Regresión logit para el Propensity score con el corte de 25 minutos en la ronda 3 (alumnos de primaria).....	93
Tabla 41: Regresión logit para el Propensity score con el corte de 30 minutos en la ronda 3 (alumnos de primaria).....	93
Tabla 42: Regresión logit para el Propensity score con el corte de 20 minutos en la ronda 5 (alumnos de secundaria) .....	94
Tabla 43: Regresión logit para el Propensity score con el corte de 25 minutos en la ronda 5 (alumnos de secundaria) .....	94
Tabla 44: Regresión logit para el Propensity score con el corte de 30 minutos en la ronda 5 (alumnos de secundaria) .....	95
Tabla 45: Test de medias de covariables entre alumnos con viajes largos y cortos (control) en la ronda 3 con el punto de corte de 25 minutos (alumnos de primaria) .	96
Tabla 46: Soporte común en el emparejamiento en la ronda 3 con el punto de corte de 25 minutos (alumnos de primaria) .....	97

Tabla 47: Test de medias de covariables entre alumnos con viajes largos y cortos (control) en la ronda 3 con el punto de corte de 30 minutos (alumnos de primaria)	98
Tabla 48: Soporte común en el emparejamiento en la ronda 3 con el punto de corte de 30 minutos (alumnos de primaria)	99
Tabla 49: Test de medias de covariables entre alumnos con viajes largos y cortos (control) en la ronda 5 con el punto de corte de 25 minutos (alumnos de secundaria)	101
Tabla 50: Soporte común en el emparejamiento en la ronda 5 con el punto de corte de 25 minutos (alumnos de secundaria)	102
Tabla 51: Test de medias de covariables entre alumnos con viajes largos y cortos (control) en la ronda 5 con el punto de corte de 30 minutos (alumnos de secundaria)	103
Tabla 52: Soporte común en el emparejamiento en la ronda 5 con el punto de corte de 30 minutos (alumnos de secundaria)	104
Tabla 53: Test de medias de covariables entre alumnos con viajes largos y cortos (control) en la ronda 3 con el punto de corte de 25 minutos, emparejamiento de 5 vecinos más cercanos y caliper de 0.005 (alumnos de primaria rurales)	106
Tabla 54: Test de medias de covariables entre alumnos con viajes largos y cortos (control) en la ronda 3 con el punto de corte de 25 minutos, emparejamiento de 5 vecinos más cercanos y caliper de 0.005 (alumnos de primaria urbanos)	106
Tabla 55: Test de medias de covariables entre alumnos con viajes largos y cortos (control) en la ronda 3 con el punto de corte de 30 minutos, emparejamiento de 5 vecinos más cercanos y caliper de 0.005 (alumnos de primaria rurales)	107
Tabla 56: Test de medias de covariables entre alumnos con viajes largos y cortos (control) en la ronda 3 con el punto de corte de 30 minutos, emparejamiento de 5 vecinos más cercanos y caliper de 0.005 (alumnos de primaria urbanos)	107
Tabla 57: Test de medias de covariables entre alumnos con viajes largos y cortos (control) en la ronda 3 con el punto de corte de 25 minutos y emparejamiento de kernel epanechnikov (alumnos de primaria rurales)	108
Tabla 58: Test de medias de covariables entre alumnos con viajes largos y cortos (control) en la ronda 3 con el punto de corte de 25 minutos y emparejamiento de kernel epanechnikov (alumnos de primaria urbanos)	108
Tabla 59: Test de medias de covariables entre alumnos con viajes largos y cortos (control) en la ronda 3 con el punto de corte de 30 minutos y emparejamiento de kernel epanechnikov (alumnos de primaria rurales)	109
Tabla 60: Test de medias de covariables entre alumnos con viajes largos y cortos (control) en la ronda 3 con el punto de corte de 30 minutos y emparejamiento de kernel epanechnikov (alumnos de primaria urbanos)	109
Tabla 61: Test de medias de covariables entre alumnos con viajes largos y cortos (control) en la ronda 5 con el punto de corte de 25 minutos, emparejamiento de 5 vecinos más cercanos y caliper de 0.005 (alumnos de secundaria rurales)	110

Tabla 62: Test de medias de covariables entre alumnos con viajes largos y cortos (control) en la ronda 5 con el punto de corte de 25 minutos, emparejamiento de 5 vecinos más cercanos y caliper de 0.005 (alumnos de secundaria urbanos).....	110
Tabla 63: Test de medias de covariables entre alumnos con viajes largos y cortos (control) en la ronda 5 con el punto de corte de 30 minutos, emparejamiento de 5 vecinos más cercanos y caliper de 0.005 (alumnos de secundaria rurales).....	111
Tabla 64: Test de medias de covariables entre alumnos con viajes largos y cortos (control) en la ronda 5 con el punto de corte de 30 minutos, emparejamiento de 5 vecinos más cercanos y caliper de 0.005 (alumnos de secundaria urbanos).....	111
Tabla 65: Test de medias de covariables entre alumnos con viajes largos y cortos (control) en la ronda 5 con el punto de corte de 25 minutos y emparejamiento de kernel epanechnikov (alumnos de secundaria rurales) .....	112
Tabla 66: Test de medias de covariables entre alumnos con viajes largos y cortos (control) en la ronda 5 con el punto de corte de 25 minutos y emparejamiento de kernel epanechnikov (alumnos de secundaria urbanos) .....	112
Tabla 67: Test de medias de covariables entre alumnos con viajes largos y cortos (control) en la ronda 5 con el punto de corte de 30 minutos y emparejamiento de kernel epanechnikov (alumnos de secundaria rurales) .....	113
Tabla 68: Test de medias de covariables entre alumnos con viajes largos y cortos (control) en la ronda 5 con el punto de corte de 30 minutos y emparejamiento de kernel epanechnikov (alumnos de secundaria urbanos) .....	113
Tabla 69: Estimación del ATT sin bootstrapping en la ronda 3 con emparejamiento de 5 vecinos más cercanos y caliper de 0.005 (alumnos de primaria rurales y urbanos) .....	114
Tabla 70: Estimación del ATT sin bootstrapping en la ronda 3 con emparejamiento de kernel epanechnikov (alumnos de primaria rurales y urbanos) .....	114
Tabla 71: Estimación del ATT sin bootstrapping en la ronda 5 con emparejamiento de 5 vecinos más cercanos y caliper de 0.005 (alumnos de secundaria rurales y urbanos) .....	115
Tabla 72: Estimación del ATT sin bootstrapping en la ronda 5 con emparejamiento de kernel epanechnikov (alumnos de secundaria rurales y urbanos).....	115
Tabla 73: Test de medias de covariables entre alumnos con viajes largos menores a 40 minutos y viajes cortos (control) en la ronda 3 con el punto de corte de 25 minutos, emparejamiento de 5 vecinos más cercanos y caliper de 0.005 (alumnos de primaria) .....	116
Tabla 74: Test de medias de covariables entre alumnos con viajes largos mayores o iguales a 40 minutos y viajes cortos (control) en la ronda 3 con el punto de corte de 25 minutos, emparejamiento de 5 vecinos más cercanos y caliper de 0.005 (alumnos de primaria) .....	117
Tabla 75: Test de medias de covariables entre alumnos con viajes largos menores a 40 minutos y viajes cortos (control) en la ronda 3 con el punto de corte de 30 minutos,	

emparejamiento de 5 vecinos más cercanos y caliper de 0.005 (alumnos de primaria) .....	117
Tabla 76: Test de medias de covariables entre alumnos con viajes largos mayores o iguales a 40 minutos y viajes cortos (control) en la ronda 3 con el punto de corte de 30 minutos, emparejamiento de 5 vecinos más cercanos y caliper de 0.005 (alumnos de primaria) .....	118
Tabla 77: Test de medias de covariables entre alumnos con viajes largos menores a 40 minutos y viajes cortos (control) en la ronda 5 con el punto de corte de 25 minutos, emparejamiento de 5 vecinos más cercanos y caliper de 0.005 (alumnos de secundaria).....	118
Tabla 78: Test de medias de covariables entre alumnos con viajes largos mayores o iguales a 40 minutos y viajes cortos (control) en la ronda 5 con el punto de corte de 25 minutos, emparejamiento de 5 vecinos más cercanos y caliper de 0.005 (alumnos de secundaria).....	119
Tabla 79: Test de medias de covariables entre alumnos con viajes largos menores a 40 minutos y viajes cortos (control) en la ronda 5 con el punto de corte de 30 minutos, emparejamiento de 5 vecinos más cercanos y caliper de 0.005 (alumnos de secundaria).....	119
Tabla 80: Test de medias de covariables entre alumnos con viajes largos mayores o iguales a 40 minutos y viajes cortos (control) en la ronda 5 con el punto de corte de 30 minutos, emparejamiento de 5 vecinos más cercanos y caliper de 0.005 (alumnos de secundaria).....	120
Tabla 81: Estimación del ATT sin bootstrapping en la ronda 3 con emparejamiento de 5 vecinos más cercanos, caliper de 0.005 y dividiendo los alumnos con viajes largos a partir del intervalo de 40 minutos (alumnos de primaria).....	121
Tabla 82: Estimación del ATT empleando bootstrapping en la ronda 5 con emparejamiento de 5 vecinos más cercanos, caliper de 0.005 y dividiendo los alumnos con viajes largos a partir del intervalo de 40 minutos (alumnos de secundaria).....	121

## Índice de gráficos

Gráfico 1: Matrícula en la Educación Básica Regular en Perú (2013-2023, en millones de alumnos).....	1
Gráfico 2: Cantidad de servicios educativos en el Perú (2013-2023).....	5
Gráfico 3: Ratio de servicios educativos de nivel primario por cada secundario en el Perú (2013-2023) .....	6
Gráfico 4: Distribución por sexo según el tiempo de desplazamiento (alumnos de primaria) .....	29
Gráfico 5: Distribución por la lengua materna según el tiempo de desplazamiento (alumnos de primaria).....	30
Gráfico 6: Porcentaje de alumnos con presencia de retraso en crecimiento en la ronda 2 según el tiempo de desplazamiento (alumnos de primaria).....	31
Gráfico 7: Distribución del modo de transporte según el tiempo de desplazamiento (alumnos de primaria).....	31
Gráfico 8: Distribución de traslado de zona de residencia urbana a rural entre la ronda 1 y 3 según el tiempo de desplazamiento (alumnos de primaria).....	33
Gráfico 9: Distribución de traslado de zona de residencia rural a urbana entre la ronda 1 y 3 según el tiempo de desplazamiento (alumnos de primaria).....	33
Gráfico 10: Distribución de la permanencia de la residencia en una zona urbana entre la ronda 1 y 3 según el tiempo de desplazamiento (alumnos de primaria) .....	33
Gráfico 11: Distribución de la permanencia de la residencia en una zona rural entre la ronda 1 y 3 según el tiempo de desplazamiento (alumnos de primaria).....	33
Gráfico 12: Distribución de la educación de la madre a partir de si posee una educación secundaria o mayor según el tiempo de desplazamiento (alumnos de primaria) .....	34
Gráfico 13: Distribución del número de familiares residentes en el hogar según el tiempo de desplazamiento (alumnos de primaria) .....	35
Gráfico 14: Distribución del índice de riqueza de la familia en la ronda 1 según el tiempo de desplazamiento (alumnos de primaria) .....	36
Gráfico 15: Distribución del índice de riqueza de la familia en la ronda 2 según el tiempo de desplazamiento (alumnos de primaria) .....	36
Gráfico 16: Distribución del índice de riqueza de la familia en la ronda 3 según el tiempo de desplazamiento (alumnos de primaria) .....	37
Gráfico 17: Distribución de la institución educativa matriculada a partir de la gestión alternativa según el tiempo de desplazamiento (alumnos de primaria) .....	38
Gráfico 18: Distribución de la institución educativa matriculada a partir de la gestión privada según el tiempo de desplazamiento (alumnos de primaria).....	38
Gráfico 19: Distribución de la institución educativa matriculada a partir de la gestión pública según el tiempo de desplazamiento (alumnos de primaria) .....	38

Gráfico 20: Distribución del puntaje obtenido en el examen CDA en la ronda 2 según el tiempo de desplazamiento (alumnos de primaria).....	39
Gráfico 21: Distribución por el sexo según el tiempo de desplazamiento (alumnos de secundaria).....	40
Gráfico 22: Distribución por la lengua materna según el tiempo de desplazamiento (alumnos de secundaria) .....	41
Gráfico 23: Porcentaje de alumnos con presencia de retraso en crecimiento en la ronda 2 según el tiempo de desplazamiento (alumnos de secundaria).....	42
Gráfico 24: Distribución del modo de transporte según el tiempo de desplazamiento (alumnos de secundaria) .....	42
Gráfico 25: Distribución de traslado de zona de residencia urbana a rural entre la ronda 1 y 5 según el tiempo de desplazamiento (alumnos de secundaria) .....	44
Gráfico 26: Distribución de traslado de zona de residencia urbana a rural entre la ronda 1 y 5 según el tiempo de desplazamiento (alumnos de secundaria) .....	44
Gráfico 27: Distribución de la permanencia de la residencia en una zona urbana entre la ronda 1 y 5 según el tiempo de desplazamiento (alumnos de secundaria) .....	44
Gráfico 28: Distribución de la permanencia de la residencia en una zona rural entre la ronda 1 y 5 según el tiempo de desplazamiento (alumnos de secundaria) .....	44
Gráfico 29: Distribución de la educación de la madre a partir de si posee una educación secundaria o mayor según el tiempo de desplazamiento (alumnos de secundaria). 45	
Gráfico 30: Distribución del número de familiares residentes en el hogar según el tiempo de desplazamiento (alumnos de secundaria) .....	46
Gráfico 31: Distribución del índice de riqueza de la familia en la ronda 1 según el tiempo de desplazamiento (alumnos de secundaria) .....	47
Gráfico 32: Distribución del índice de riqueza de la familia en la ronda 4 según el tiempo de desplazamiento (alumnos de secundaria) .....	47
Gráfico 33: Distribución del índice de riqueza de la familia en la ronda 5 según el tiempo de desplazamiento (alumnos de secundaria) .....	48
Gráfico 34: Distribución de la institución educativa matriculada a partir de la gestión alternativa según el tiempo de desplazamiento (alumnos de secundaria) .....	49
Gráfico 35: Distribución de la institución educativa matriculada a partir de la gestión privada según el tiempo de desplazamiento (alumnos de secundaria) .....	49
Gráfico 36: Distribución de la institución educativa matriculada a partir de la gestión pública según el tiempo de desplazamiento (alumnos de secundaria).....	49
Gráfico 37: Distribución del puntaje obtenido en el examen CDA en la ronda 2 según el tiempo de desplazamiento (alumnos de secundaria).....	50
Gráfico 38: Distribución del tiempo de desplazamiento en la ronda 3 (alumnos de primaria) .....	51
Gráfico 39: Distribución del tiempo de desplazamiento en la ronda 5 (alumnos de secundaria).....	52

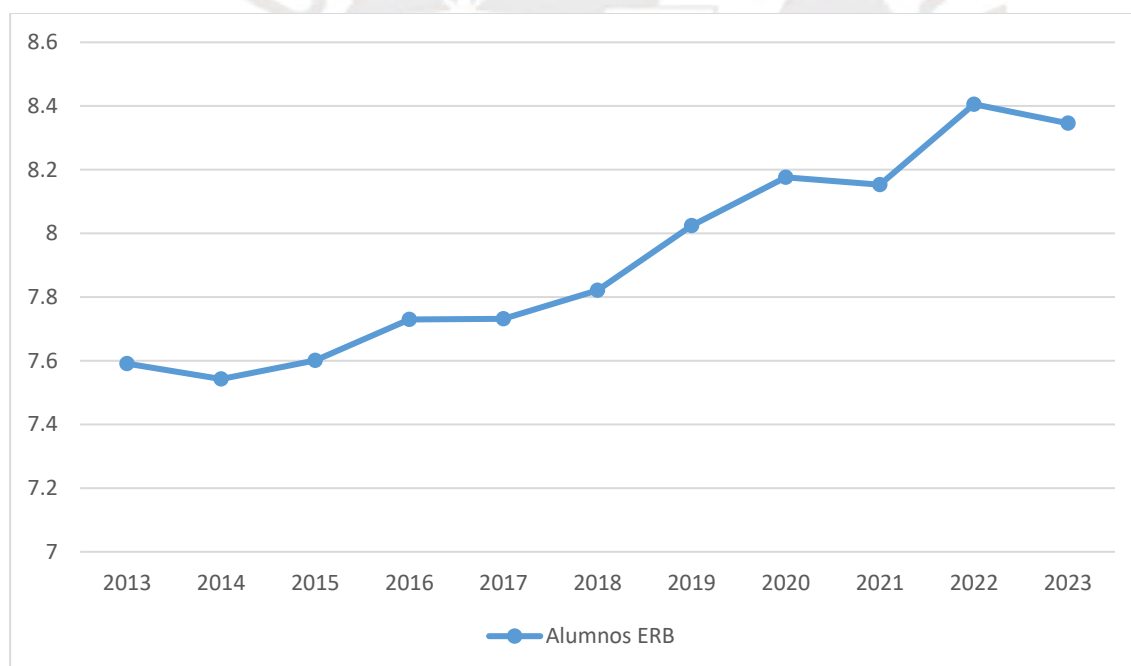
Gráfico 40: Puntaje obtenido en la prueba de matemática en la ronda 3 según el tiempo de desplazamiento (alumnos de primaria).....	53
Gráfico 41: Puntaje obtenido en la prueba de matemática en la ronda 5 según el tiempo de desplazamiento (alumnos de secundaria) .....	53
Gráfico 42: Porcentaje de sesgo estandarizado entre las covariables de la ronda 3 con el punto de corte de 20 minutos de los alumnos de viajes cortos al ser emparejados y no emparejados (alumnos de primaria) .....	55
Gráfico 43: Soporte común de la ronda 3 con el punto de corte de 20 minutos según el Propensity Score (alumnos de primaria).....	56
Gráfico 44: Porcentaje de sesgo estandarizado entre las covariables de la ronda 5 con el punto de corte de 20 minutos de los alumnos de viajes cortos al ser emparejados y no emparejados (alumnos de secundaria) .....	60
Gráfico 45: Soporte común de la ronda 5 con el punto de corte de 20 minutos según el Propensity Score (alumnos de secundaria).....	61
Gráfico 46: Porcentaje de sesgo estandarizado entre las covariables de la ronda 3 con el punto de corte de 25 minutos de los alumnos de viajes cortos al ser emparejados y no emparejados (alumnos de primaria) .....	97
Gráfico 47: Soporte común de la ronda 3 con el punto de corte de 25 minutos según el Propensity Score (alumnos de primaria).....	98
Gráfico 48: Porcentaje de sesgo estandarizado entre las covariables de la ronda 3 con el punto de corte de 30 minutos de los alumnos de viajes cortos al ser emparejados y no emparejados (alumnos de primaria) .....	99
Gráfico 49: Soporte común de la ronda 3 con el punto de corte de 30 minutos según el Propensity Score (alumnos de primaria).....	100
Gráfico 50: Porcentaje de sesgo estandarizado entre las covariables de la ronda 5 con el punto de corte de 25 minutos de los alumnos de viajes cortos al ser emparejados y no emparejados (alumnos de secundaria) .....	102
Gráfico 51: Soporte común de la ronda 5 con el punto de corte de 25 minutos según el Propensity Score (alumnos de secundaria) .....	103
Gráfico 52: Porcentaje de sesgo estandarizado entre las covariables de la ronda 5 con el punto de corte de 30 minutos de los alumnos de viajes cortos al ser emparejados y no emparejados (alumnos de secundaria) .....	104
Gráfico 53: Soporte común de la ronda 5 con el punto de corte de 30 minutos según el Propensity Score (alumnos de secundaria) .....	105

## Introducción

De la mano de un crecimiento económico sostenido durante la década del 2010, el Perú experimentó un aumento constante de la asistencia escolar tanto a nivel primario como secundario, acercándose el país al objetivo de una cobertura universal de la Educación Básica Regular (EBR), que abarca desde el nivel de educación inicial hasta el secundario (Portocarrero, 2023). Según los datos del número de servicios educativos de la Unidad Estadística de Calidad Educativa (ESCALE) del Ministerio de Educación del Perú (MINEDU), la cantidad de estudiantes matriculados en los niveles que cubre el EBR desde hace una década ha continuado incrementando. Como se muestra en el gráfico 1, la matrícula en la EBR de alumnos ha seguido una tendencia positiva, siendo 7.59 millones de matriculados en el año 2013, los cuales subieron a una cifra de 8.34 millones de alumnos en el 2023 (ESCALE, 2023).

Gráfico 1

Matrícula en la Educación Básica Regular en Perú (2013-2023, en millones de alumnos)



Fuente: ESCALE (2023). Elaboración propia.

Sin embargo, las desigualdades socioeconómicas siguen persistiendo en la sociedad peruana, cuya influencia en el acceso a una educación de calidad a la población vulnerable del país se vuelve innegablemente notoria en el tiempo de viaje de los estudiantes en este grupo poblacional. La información provista por ESCALE añade que la cantidad de instituciones educativas de nivel primaria en Perú en el año 2023 fue más del doble que las de nivel secundario, llegando a las cifras de 38128 y

15214 respectivamente. Asimismo, la oferta educativa del sector privado es casi inexistente en las zonas rurales, mientras que en las zonas urbanas casi dos de cada tres escuelas son de gestión privada (ESCALE, 2023).

La reducida y desigual oferta educativa de nivel secundario en comparación con el primario genera que una gran porción de los alumnos peruanos tenga que realizar trayectos más largos para asistir a sus clases cuando transicionan entre los niveles educativos debido a una precaria conectividad entre estos. Asimismo, este tiempo para desplazarse a la institución educativa, si ocupa una considerable parte del horario diario de las familias, puede limitar el acceso a la educación provista y mermar el proceso de aprendizaje de los alumnos peruanos (Córdova & Ruiz, 2014; Gutiérrez, 2022).

La literatura que analiza el efecto del tiempo de desplazamiento a la institución educativa en el rendimiento académico de los alumnos encuentra, en su mayoría, una relación inversa entre las variables. Tanto a nivel internacional (Contreras et al., 2018; Tigre, 2017; Tigre et al., 2014) como nacional (Villanueva, 2023; Córdova & Ruiz, 2014), diversas investigaciones señalan que un mayor tiempo de desplazamiento es perjudicial para el rendimiento académico de los estudiantes, pero también hay estudios (García Hermoso et al., 2017) que señalan que este efecto negativo es heterogéneo de acuerdo con el tiempo de desplazamiento.

No obstante, la significancia y la duración necesarias para que el tiempo de desplazamiento afecte el rendimiento académico de los alumnos dependen del país de las investigaciones, las zonas de residencia y el nivel de educación. Esta última desagregación por niveles de educación, la cual ha sido escasamente estudiada tanto a nivel internacional como nacional, podría reflejar un impacto más diferenciado debido a los contrastes entre la oferta educativa a nivel primario y secundario, las diferentes preferencias en la educación de las familias según la zona de residencia, y la precaria conectividad entre los niveles educativos en el Perú. Investigaciones previas nacionales como la de Gutiérrez (2022) encuentran que una reducida conectividad terrestre entre las instituciones educativas primarias y secundarias fomenta la deserción escolar de los alumnos en la transición entre estos niveles educativos, así como el trabajo de Córdova & Ruiz (2014) demuestra que un mayor tiempo de desplazamiento incrementa la probabilidad de que los alumnos en zonas

rurales experimenten atraso escolar. Sin embargo, este impacto es mayor y heterogéneo a nivel regional para las instituciones primarias, mientras que a nivel secundario el efecto es menor, pero homogéneo. De esta manera, se espera que el efecto negativo del tiempo de desplazamiento en el rendimiento académico de los estudiantes peruanos varíe de acuerdo con el nivel educativo (Córdova & Ruiz, 2014; Gutiérrez, 2022).

Por lo tanto, el presente estudio se enfocará en desagregar cómo afecta este tiempo dedicado al desplazamiento desde la residencia hacia la institución educativa al rendimiento académico de los estudiantes de nivel primario y secundario en el Perú. Esto se logrará a través del uso de la base de datos de Niños del Milenio, un proyecto longitudinal internacional de recolección de información sobre el estilo de vida de los niños con la meta de recomendar políticas públicas dirigidas a las causas y consecuencias de la pobreza infantil, donde la educación juega un rol importante. Este recolecta datos de los mismos niños a través de los años (denominadas rondas), divididos en dos cohortes de acuerdo a su edad (Villanueva, 2023). Así, mediante los datos de la cohorte menor de los alumnos peruanos y la metodología propuesta de Propensity Score Matching por Tigre (2014), se busca resaltar la relevancia en la relación propuesta.

De este modo, los resultados del estudio permitirán recomendar la implementación de políticas públicas que busquen reducir el tiempo de viaje para que no sea nocivo para el proceso de aprendizaje de los estudiantes, mejorar la calidad del servicio de la oferta educativa, y optimizar la conectividad entre las instituciones educativas primarias y secundarias para facilitar el acceso y reducir el tiempo de viaje de los alumnos hacia sus instituciones educativas. Por tanto, el análisis del posible impacto negativo que genera el tiempo de desplazamiento para los estudiantes del país permite reevaluar el esfuerzo de esta población y sus familiares para acceder a la educación, e implementar medidas que favorezcan la accesibilidad a los servicios educativos, elevar la calidad de vida de los estudiantes y que la juventud peruana experimente favorablemente su vida escolar

## 1. Marco teórico y Revisión de Literatura

### 1.1 La Oferta educativa

Según la información provista por ESCALE (2023), la oferta educativa de nivel primario resulta más abundante en las zonas rurales que en las urbanas. Como se puede apreciar en la tabla 1, el 67% de las escuelas de primaria en el Perú se ubican en las zonas rurales, mientras que la mayoría de instituciones de nivel secundario se encuentran en las áreas urbanas, representando el 53%. Asimismo, se puede comprobar que las instituciones educativas privadas predominan en las áreas urbanas, representando el 63% y 62% del nivel primario y secundario respectivamente. Mientras tanto, en el área rural, las mayores porciones de las escuelas son públicas, llegando a ser el 99% de las instituciones tanto a nivel primario como secundario (ESCALE, 2023). La gran cantidad de escuelas primarias en las zonas rurales se debe al deseo del Estado de cubrir la necesidad de educación de este nivel en todas las comunidades residentes en esta área, pues estas se encuentran dispersas por factores geográficos y culturales, mientras que en el área urbana el sector privado se concentra debido a la alta demanda de educación bajo esta modalidad de gestión (Chaman & Moali, 2020; Contreras et al., 2018).

Tabla 1:

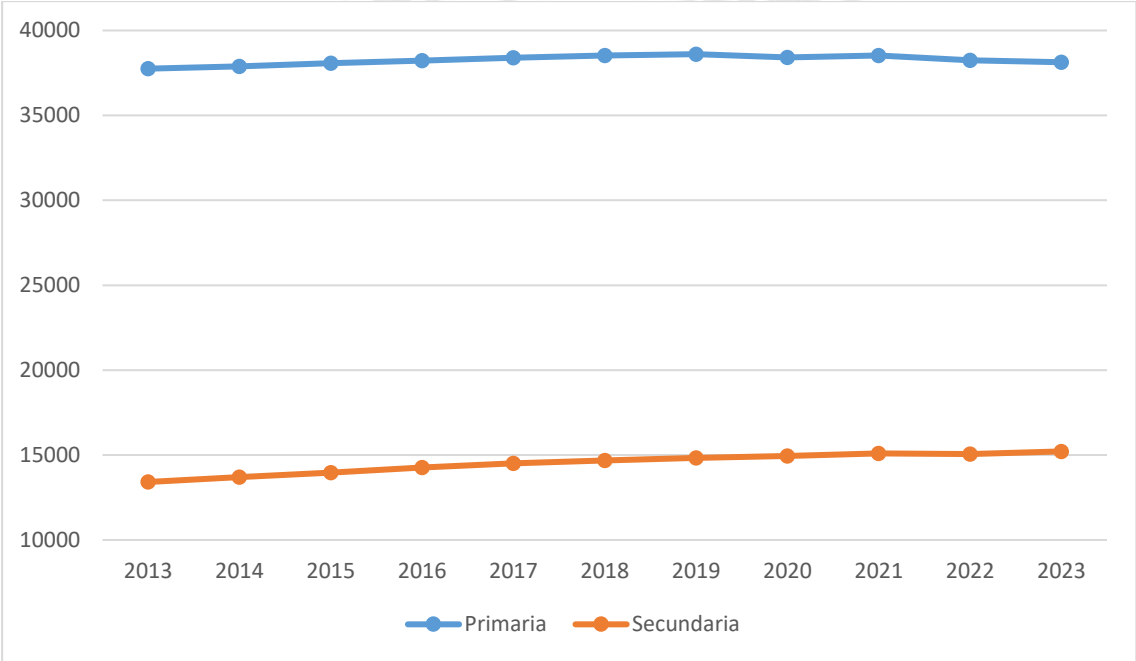
Cantidad de instituciones educativas nacionales de nivel primario y secundario según zona de residencia y gestión (2023)

NACIONAL			
	TOTAL	URBANO	RURAL
PRIMARIA	38128	33%	67%
SECUNDARIA	15214	53%	47%
URBANO			
	TOTAL	PÚBLICO	PRIVADO
PRIMARIA	12620	37%	63%
SECUNDARIA	8029	38%	62%
RURAL			
	TOTAL	PÚBLICO	PRIVADO
PRIMARIA	25508	99%	1%
SECUNDARIA	7185	99%	1%

Fuente: ESCALE (2023). Elaboración propia.

Del mismo modo, al analizar la cantidad de instituciones educativas por año desde el 2013 hasta el 2023 en el gráfico 2, se puede apreciar que la cantidad de escuelas primarias se ha mantenido constante en el tiempo, alrededor de las 38,000 instituciones, mientras que los colegios de nivel secundaria la cifra ha aumentado levemente desde los 13,000 hasta las 15,000 instituciones. No obstante, como se ve en el gráfico 3, la ratio de instituciones primarias por cada secundaria no ha logrado bajar de 2.5, lo que podría indicar una vigente baja conectividad entre los niveles educativos, la cual es aún más acentuada en las zonas rurales, a pesar de su mejoría comparada a la zona urbana (ESCALE,2024).

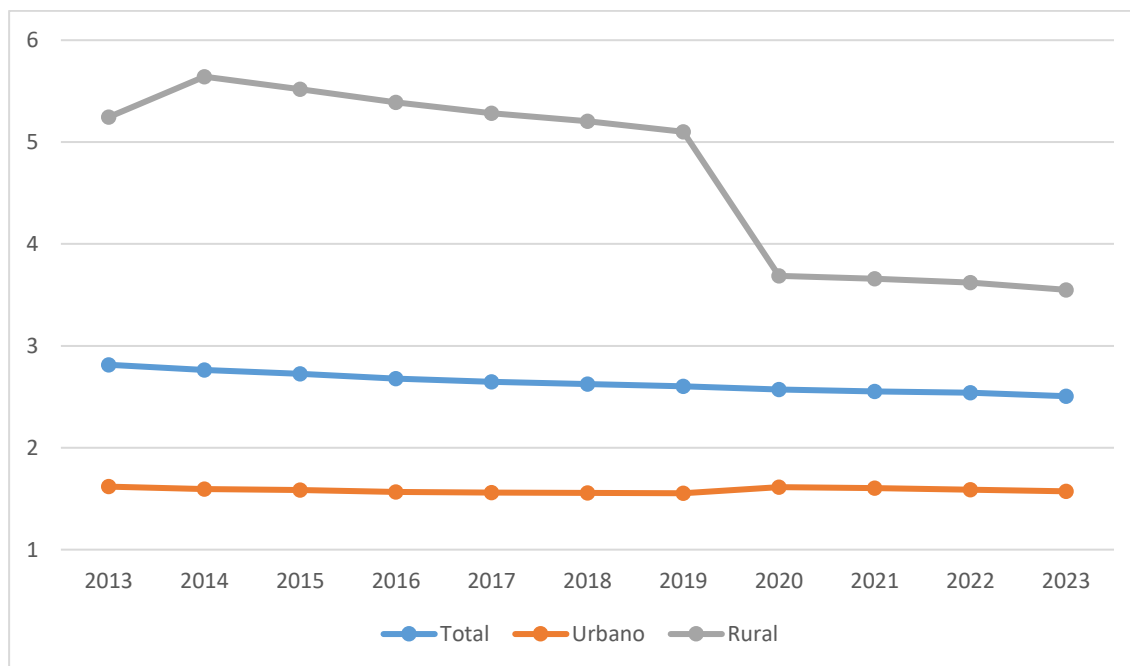
Gráfico 2  
Cantidad de servicios educativos en el Perú (2013-2023)



Fuente: ESCALE (2023). Elaboración propia.

Gráfico 3

Ratio de servicios educativos de nivel primario por cada secundario en el Perú (2013-2023)



Fuente: ESCALE (2023). Elaboración propia.

Si bien el aumento de más instituciones educativas a nivel general surge como una respuesta directa para aumentar la cobertura educativa, esta no necesariamente podría resultar en una mejora de la calidad de la educación en el país. Como argumentan Duraisamy et al. (1997) mediante sus estudios del contexto educativo del Estado de Tamil Nadu en India, existe un “trade-off” entre la cantidad y calidad en la oferta educativa, donde la ampliación de la cobertura por parte de la gestión pública puede también generar instituciones con una calidad precaria de infraestructura, docentes y recursos, pues el presupuesto en la educación se concentra solo en generar más instituciones educativas.

Por lo tanto, diversos autores como Contreras et al. (2018), Córdova & Ruiz (2014) y Gutiérrez (2002) proponen que se debe priorizar la facilidad de acceso a las instituciones educativas a nivel nacional, donde el tiempo de desplazamiento a dichas instituciones posee un rol importante en la calidad de la educación provista. Asimismo, este último autor destaca cómo la transición escolar de los estudiantes del nivel primaria a secundaria es la principal etapa donde hay más deserciones especialmente en las áreas rurales y semiurbanas, y donde el rendimiento académico y la cercanía de las instituciones educativas secundarias a las primarias juegan un rol importante en la decisión de las familias para dejar que sus hijos continúen en sus estudios, por lo que

es relevante analizar cómo es la relación entre estas dos variables para minimizar las deserciones y aumentar la calidad educativa.

### 1.2 Tiempo de desplazamiento

Como el tiempo de desplazamiento del alumno hacia su institución educativa constituye la variable explicativa principal en el trabajo de investigación, se define claramente este factor en el presente documento como el tiempo total que implementan los estudiantes en sus viajes de ida desde su domicilio hacia la institución educativa a la cual asisten. Diversos autores utilizan este tiempo, así como el tiempo del viaje de regreso a su residencia, para analizar el costo de oportunidad que esta variable implica en las decisiones de los alumnos y sus familias, pues refleja ingresos potenciales perdidos en una actividad necesaria para acceder a la educación y es capaz de asimilar otros costos relacionados al desplazamiento como la congestión vehicular y gastos propios según el modo de transporte utilizado (Tigre, 2014; Tigre et al., 2017). Asimismo, como comentan Ding & Feng (2022), se ha evidenciado en diferentes estudios un impacto negativo del tiempo de desplazamiento en el bienestar psicológico en las personas, ya sean jóvenes estudiantes o adultos empleados, por lo que los autores aluden a que esta reducción en la felicidad en los alumnos podría también impactar su rendimiento académico.

De esta manera, el cálculo de este tiempo es esencial para comprender e interpretar claramente su impacto en el proceso de aprendizaje y bienestar de los alumnos, pues podrían utilizarlo en otras actividades más beneficiosas para ellos, como estudiar, actividades recreacionales o mayor tiempo de descanso (Contreras et al., 2018). Asimismo, el tiempo de desplazamiento se prefiere sobre la distancia de viaje debido a diferentes factores de acuerdo a las zonas de residencia de los estudiantes. En las zonas rurales, el tiempo es preferido debido a la geografía irregular y la falta de vías adecuadas hacia el destino, por lo que un alumno podría emplear más tiempo según cuán accidentado sea el trayecto. Por otro lado, en las áreas urbanas, la congestión vehicular y los distintos modos de viaje que pueden elegir los estudiantes generan diferentes tiempos de viaje aún cuando la distancia es la misma para todos los individuos (Córdova & Ruiz, 2014).

### 1.3 Rendimiento académico

El rendimiento académico escolar, la variable dependiente del presente estudio, es definido por Steinmayr et al. (2015) como el resultado del desempeño que un

individuo alcanza en objetivos específicos formulados usualmente por instituciones educativas como escuelas y universidades. Además, estos autores añaden que el rendimiento académico puede ser representado por distintos indicadores, siendo el más conocido el de una evaluación calificada con el fin de reflejar las capacidades aprendidas del individuo (Steinmayr et al., 2015). No obstante, el rendimiento académico puede representar varias competencias de una persona en diferentes áreas de aprendizaje. Por lo tanto, para la realización del presente estudio, se buscará enfocar en un área esencial para el desarrollo académico y laboral de la persona, que son las matemáticas. Como comenta Guhl (2019), una educación matemática adecuada en la edad temprana durante la educación inicial y primaria provoca que el alumno desarrolle facultades como la solución de problemas, el conteo y la clasificación de imágenes y números. Estas facultades complementan al desarrollo de diversas habilidades como la comprensión lectora y el pensamiento crítico, las cuales también son importantes para el rendimiento del individuo en las demás áreas de aprendizaje (Guhl, 2019).

Esta vinculación entre la educación matemática y el rendimiento en las demás áreas de aprendizaje también se evidencia en los siguientes niveles de educación, así como en el ambiente laboral. Por un lado, Cabuquin & Abocejo (2023) encuentran una correlación significativa y positiva en el desempeño de los estudiantes en sus evaluaciones en matemáticas con su puntaje general de los demás cursos del currículo en diversas escuelas públicas en Filipinas. Por otra parte, el Comité Asesor sobre Educación Matemática de la Sociedad Real del Reino Unido (2011), mediante la realización de una encuesta nacional en Inglaterra para analizar la importancia de las habilidades matemáticas en el ambiente laboral, comenta que los empleadores buscan que los potenciales empleados sepan aplicar como mínimo conocimientos matemáticos generales, y que sean capaces de emplearlos en diferentes contextos y de manera eficaz, como en el uso de dispositivos electrónicos y de los programas esenciales en las principales labores de oficina. Si bien diferentes campos laborales exigen distintos niveles de educación matemática, el autor añade que existe una tendencia global en la que los oficios de trabajo manual y de bajo nivel académico están disminuyendo, mientras que los trabajos que requieren habilidades de gestión de recursos y habilidades de solución de problemas son cada vez más ofertados, los cuales exigen conocimientos relacionados a las matemáticas.

De esta manera, para el presente estudio, el rendimiento académico de los alumnos de nivel primaria y secundaria del Perú se medirá a partir de los resultados de los alumnos peruanos de la cohorte menor en las evaluaciones de matemáticas obtenidas de la base de datos Niños del Milenio (NdM) en las rondas 3 y 5 de este estudio. Este examen en la ronda 3 cubre temáticas como interpretación de datos, solución de problemas numéricos, medición y geometría básica, además de que las preguntas varían de complejidad de acuerdo a la cohorte examinada. Para las siguientes rondas, la dificultad de los exámenes se adecúa a la edad de los estudiantes y varias preguntas nuevas se adaptaron tomando de base diferentes tests cognitivos internacionales con temáticas similares (Villanueva, 2023). De esta manera, en la ronda 3 el examen de matemáticas tenía 29 preguntas mientras que para la ronda 5 se evaluaba 31, cada una con un valor de un punto. Adicionalmente, se utilizará de manera auxiliar la evaluación de desarrollo cognitivo (CDA en inglés) para representar las habilidades innatas de los alumnos en la ronda 3 y 5 para la evaluación de matemáticas, ya que esta prueba cubre ejercicios de nociones cuantitativas, empleando 12 preguntas con el valor de un punto cada una (Espinoza & Douglas, 2022).

Por último, Villanueva (2023), al utilizar datos longitudinales de estudiantes cuando cursaban el nivel educativo primario y secundario, complementa en la literatura del rendimiento académico al constituirlo como un producto de diversos insumos de la demanda y oferta educativa que se van acumulando en el tiempo. Así, el tiempo diario destinado a los viajes, al interpretarse como una variable intertemporal que influye en la educación de los estudiantes peruanos, resulta en un aumento progresivo de horas perdidas que podrían haber sido utilizadas en el proceso de aprendizaje de los alumnos, el cual con el tiempo se ve aún más degradado (Villanueva, 2023).

#### 1.4 Modelo Económico

La relación del rendimiento académico y el tiempo de desplazamiento de estudiantes puede interpretarse bajo distintas funciones; sin embargo, la que se implementará y profundizará en el estudio es la función de producción educativa propuesta por Levin (1974) y Hanushek (1979), la cual se presenta en la siguiente ecuación.

$$A_{ij} = f(S_j, B_i, P_{ij}, I_i)$$

De esta manera,  $A_{ij}$  representa el resultado del proceso de aprendizaje del alumno (su rendimiento académico)  $i$  en la institución educativa  $j$ ,  $S_j$  es el conjunto de insumos educativos e infraestructura disponible de la institución  $j$ ,  $B_i$  denota las diversas características socioeconómicas de la familia del estudiante  $i$ ,  $P_{ij}$  refleja los diferentes canales de influencia externos en el rendimiento del alumno  $i$  en la institución  $j$ , y  $I_i$  indica las capacidades innatas del estudiante  $i$ . Por tanto, como argumentan Córdova & Ruiz (2014), el rendimiento académico de un alumno se encuentra sujeto a una combinación de variables pertenecientes tanto de la oferta como la demanda del mercado educativo. Respecto al tiempo de desplazamiento, los autores detallan que este revela información en ambos sectores del mercado al reflejar la accesibilidad de la educación de las familias de acuerdo a su nivel socioeconómico, así como la provisión de las instituciones educativas y el correcto abastecimiento de estas de herramientas educativas para ofrecer correctamente el servicio en las comunidades respectivas (Córdova & Ruiz, 2014).

No obstante, Hanushek (1992) añade que la asignación óptima del tiempo de las familias y, por ende, del alumno, varía en el tiempo debido a cambios en las decisiones familiares, en la cantidad de integrantes y en la oferta educativa. En consecuencia, el autor concluye que la función de producción del rendimiento educativo del alumno debe acumular todos los insumos pasados y actuales que brindan la demanda y oferta del mercado educativo. Asimismo, Córdova y Ruiz (2014) sustentan que el tiempo de desplazamiento (ya sea desde o hacia la institución educativa) afecta negativamente la eficiencia en que los insumos son aprovechados por el estudiante, ya sea por razones como la fatiga o menor disponibilidad de tiempo para el estudio. Estos aportes de los autores se ven reflejados en la siguiente función de producción acumulada. Ahora, las variables de la función del proceso educativo indicado previamente varían en el tiempo denotado por el superíndice  $t$ , y se agrega el factor  $D$  que irrumpe en el efecto de los insumos de la función en el proceso educativo, y que depende directamente e inversamente del tiempo de desplazamiento  $c$ .

$$A_{ij} = D(c)^t * f(S_j^t, B_i^t, P_{ij}^t, I_i^t)$$

De este modo, se obtiene que el rendimiento académico de los alumnos depende de los insumos acumulados en el tiempo brindados tanto por la oferta como

la demanda educativa, cuyo aprovechamiento por parte de los estudiantes se ve afectado por el tiempo de desplazamiento en viajes que involucren su residencia y la institución educativa que acude.

#### 1.5 El efecto del tiempo de desplazamiento en el rendimiento académico

En su mayor parte, la literatura indica un impacto negativo del tiempo de desplazamiento sobre el rendimiento académico, el cual gira en torno a la distribución del tiempo de los alumnos. Córdova & Ruiz (2014) agregan que el tiempo de desplazamiento de un estudiante influye en su rendimiento académico en 3 aspectos principales en la asignación de su tiempo disponible diario. Primero, un mayor tiempo de desplazamiento podría significar menores horas dedicadas a actividades académicas, aumenta el costo de oportunidad de la asistencia a las instituciones educativas y genera mayor interés en participar en actividades laborales apoyando a la familia. Segundo, más horas dedicadas a viajes hacia la institución educativa podrían generar cansancio y posibles distracciones en los estudiantes, afectando negativamente su capacidad de aprendizaje tanto dentro como fuera de la institución educativa. Por último, también podría reducir el tiempo de interacción y supervisión de los padres de familia hacia sus hijos, disminuyendo su participación en la vida escolar de estos como el apoyo en sus deberes y tareas académicas (Córdova & Ruiz, 2014; Li & Liu, 2009).

Por otro lado, como sustenta Vuri (2007), las diversidades regionales, culturales y socioeconómicas entre las diferentes poblaciones podrían afectar la influencia del tiempo de desplazamiento en el rendimiento académico y asistencia escolar de los estudiantes. Adicionalmente, otro factor que puede alterar esta relación en las variables de estudio es el modo de viaje de los alumnos, el cual es analizado por algunos autores que señalan que el tiempo de desplazamiento puede ser aprovechado para el ejercicio físico y el correcto desarrollo de los estudiantes, y que este tiempo recién resulta perjudicial para el proceso de aprendizaje de los alumnos cuando se supera cierto umbral de duración del viaje (García Hermoso et al., 2017; Yeung & Nguyen-Hoang, 2020).

Por tanto, en la revisión de literatura empírica, se han encontrado trabajos que buscan esclarecer el impacto del tiempo de viaje tanto de ida como de ida y vuelta en el rendimiento académico de estudiantes a nivel nacional e internacional, los cuales en su mayoría se enfocan en ciertos grupos de la población estudiantil, como en los

niveles de educación primario y secundario, o según si los estudiantes residen en zonas urbanas o rurales. En primer lugar, Contreras et al. (2018) utiliza los datos de las evaluaciones más recientes de matemáticas y de lenguaje del Sistema de Medición de la Calidad de la Educación (SIMCE) en la ciudad de Santiago de Chile para analizar el impacto del tiempo de viaje desde y hacia la institución educativa en los estudiantes de secundaria de la ciudad, en el cual sus resultados varían de acuerdo al modelo que se utiliza. Primero, al utilizar un modelo de mínimos cuadrados ordinarios (MCO) los resultados muestran un efecto positivo y significativo del tiempo de viaje en el rendimiento de los alumnos. Por otro lado, al utilizar modelos que implementan la variable instrumental de la distancia del hogar del individuo a las 2 escuelas más cercanas para estimar correctamente el efecto del tiempo de viaje y efectos fijos a nivel de instituciones educativas, los autores encuentran que en este modelo esta última variable afecta negativamente al rendimiento académico del alumno. No obstante, el impacto de esta variable es más reducido y no significativo en comparación con el modelo MCO.

Este mismo ejercicio, pero en instituciones educativas primarias y utilizando exclusivamente el tiempo de viaje de ida es realizado por Tigre et al. (2017), con datos de una encuesta multinivel realizada por el instituto de investigación social Joaquim Nabuco en la ciudad de Recife en 2013 que recolecta calificaciones de alumnos en el curso de matemáticas y datos de características objetivas y subjetivas de las instituciones educativas y familias. Así, mediante el uso de la misma variable instrumental que Contreras et al. para reajustar la estimación de la variable explicativa del tiempo de desplazamiento, los autores hallan que el tiempo de desplazamiento de ida genera un efecto negativo y significativo en el rendimiento académico de los estudiantes de primaria. Además, si bien ambos trabajos utilizan variables de control relacionadas a características de la familia del alumno y de la institución educativa, el de Tigre et al. implementa también el modo de viaje que utilizan los estudiantes para desplazarse desde y hacia sus instituciones educativas.

Del mismo modo, Villanueva (2023) sigue la metodología de variables instrumentales como Contreras et al. para analizar el impacto del tiempo de desplazamiento en el rendimiento académico en las zonas rurales del Perú. No obstante, él utiliza como instrumento una variable dicotómica que indica si la comunidad donde reside el alumno existe una institución educativa en la que pueda

estudiar de acuerdo si se encuentra en primaria o secundaria. Además, a diferencia de los autores pasados, Villanueva implementa datos en diferentes años para realizar un modelo panel de efectos fijos con variables instrumentales. Utilizando los datos de las rondas 3,4 y 5 del estudio Niños del Milenio en el cual emplea las calificaciones de los exámenes de matemáticas, comprensión lectora y el test de imágenes de Peabody, el autor encuentra que el tiempo de desplazamiento de ida y vuelta afecta negativamente al rendimiento académico, pero este es reducido y no significativo. Cabe añadir que el autor asume que el tiempo de ida y vuelta equivale a la duplicación del tiempo de viaje de ida bajo el supuesto propuesto por él mismo que en zonas rurales hay reducidas modalidades de transporte disponibles para los alumnos. Además, omite variables en su modelo que inciden en esta relación como la educación de la madre y lengua materna porque no varían en el tiempo y podrían causar sesgo, pero que considera de igual manera relevantes en el análisis.

Adicionalmente, Tigre (2014) analiza la misma relación entre el tiempo de desplazamiento desde el domicilio hacia la institución educativa y el rendimiento académico de los alumnos en el nivel de educación primaria en Brasil, utilizando la misma encuesta que el trabajo de Tigre et al. (2017), pero realiza su análisis a partir del método de Propensity Score Matching (PSM). Diferenciando los dos grupos dentro de la población a partir del tiempo de desplazamiento, asumiendo el supuesto de independencia condicional y utilizando distintos procedimientos de emparejamiento, el autor encuentra que los alumnos de nivel de educación primaria que se demoran más de 1 hora en sus viajes desde y hacia la institución educativa presentan un menor rendimiento académico que los que viajan en un tiempo menor a este.

Por otro lado, Córdova & Ruiz (2014) buscan realizar el mismo análisis del tiempo de desplazamiento de ida hacia la institución educativa en el rendimiento de los estudiantes, pero implementan el atraso escolar como una variable dicotómica dependiente que refleja el rendimiento académico de los alumnos en las zonas rurales del Perú. De esta manera, mediante modelos logit binomiales, los autores hallan diferentes resultados de acuerdo a los niveles de educación escolar. A nivel primario, el tiempo de desplazamiento presenta un efecto positivo y significativo en la probabilidad de que el alumno se encuentre atrasado. Asimismo, este efecto es heterogéneo según las regiones del Perú, siendo la costa donde el impacto del tiempo de desplazamiento es más notorio. Mientras tanto, en la sierra y en la selva este efecto

es menor debido a los altos tiempos de viaje en general en todos los estudiantes y una mayor importancia de otros factores en la oferta educativa sobre el rendimiento del alumno respectivamente. A nivel secundario, el impacto del tiempo de desplazamiento en el atraso escolar sigue siendo positivo y significativo; sin embargo, en esta ocasión el efecto es homogéneo para todas las regiones en el país, lo que podría indicar una escasez general de estas instituciones en las zonas rurales del país.

Por otra parte, García-Hermoso et al. (2017) contribuyen en la literatura expandiendo el impacto del tiempo de desplazamiento sobre el rendimiento académico de los estudiantes a través de su estudio y encuesta realizada a estudiantes en zonas rurales de secundaria en Chile al enfocarse en el tiempo de desplazamiento activo (TDA). Los autores definen el TDA como el tiempo que emplea un individuo para desplazarse mediante el ejercicio físico de su cuerpo, como caminar o el uso de bicicleta. Así, mediante diversas fuentes ellos sustentan que la actividad física adecuada promovería el correcto desarrollo del hipocampo, estructura cerebral que participa en el proceso de aprendizaje y memoria del ser humano, así como aportaría a un crecimiento físico adecuado.

Mediante modelos logit donde la variable dependiente divide a la población estudiantil según si poseen una nota de matemáticas o lenguaje superior al punto de corte propuesto, ellos hallan que estudiantes con TDA de entre 30 minutos a una hora en sus viajes desde y hacia la institución educativa son más propensos a poseer un rendimiento académico mayor a quienes no se desplazan activamente. Adicionalmente, los alumnos con un TDA fuera de este rango no presentan una relación significativa con el rendimiento académico. Por estos resultados, los autores infieren que sólo un específico rango de TDA afectaría positivamente al rendimiento académico de los estudiantes, por lo que un tiempo de desplazamiento muy corto o largo no beneficiaría a los alumnos, e incluso podría perjudicar su rendimiento académico potencial (García-Hermoso, 2017).

Finalmente, un análisis similar al del previo trabajo, pero enfocado en cómo el tiempo de desplazamiento afecta de manera distinta al rendimiento académico de los estudiantes de acuerdo al modo de transporte es realizado por Yeung & Nguyen-Hoang (2020). A través de un estudio que recolecta información sobre el rendimiento académico a partir de las calificaciones de matemáticas y lectura y las características

del modo de transporte a las instituciones educativas de los alumnos de Estados Unidos, los autores utilizan los datos de alumnos de nivel primario y un modelo de variables instrumentales para realizar su investigación. Estos instrumentos indican la cantidad de adultos trabajadores que utilizan medios de transporte activos en la vecindad de los alumnos, bajo el fundamento de que mientras más personas se desplacen activamente, los alumnos tenderán también a implementarlo. Así, los resultados señalan que los alumnos que se desplazan activamente y los que viajan en transporte privado poseen un rendimiento académico mayor que los que viajan en el servicio de bus de los colegios, debido a la mayor actividad física y la mayor disponibilidad de tiempo del alumno respectivamente.

En resumen, en la literatura empírica se puede hallar una relativa predominancia de resultados que sugieren un efecto negativo del tiempo de desplazamiento sobre el rendimiento académico. No obstante, el impacto podría ser bastante sensible a las especificaciones de cada investigación, ya sea el país donde se realizó, las evaluaciones usadas para medir el rendimiento académico, la zona de residencia, la metodología empleada, entre otras. Por último, cabe añadir que hay escasa literatura que compara este efecto en la población de alumnos de nivel primario y secundario, lo que podría contribuir al hallazgo de efectos heterogéneos del tiempo de desplazamiento sobre el rendimiento académico de acuerdo con el nivel educativo de los alumnos.

## 2. Objetivos e hipótesis de la investigación

El presente trabajo de investigación posee como objetivo principal identificar cómo afecta el tiempo de desplazamiento al rendimiento académico de los estudiantes en el nivel primario y secundario en el Perú. Por tanto, la pregunta de investigación es la siguiente: ¿Cuál es el impacto del tiempo de desplazamiento en el rendimiento académico reflejado en las calificaciones de las evaluaciones de matemática de los estudiantes de primaria y secundaria del Perú? De esta manera, la hipótesis correspondiente propone que el aumento del tiempo de desplazamiento en viajes desde la residencia hacia la institución educativa generaría un menor rendimiento académico de los estudiantes en el Perú, siendo este efecto más profundo para los estudiantes de secundaria. Los mecanismos detrás de este efecto yacen en el menor tiempo dedicado al estudio, mayor cansancio y distracciones que irrumpen en el proceso de aprendizaje, y una menor capacidad de supervisión de los padres de familia en la vida escolar de los estudiantes del Perú (Córdova y Ruiz, 2014).

Por tanto, la comprobación de esta relación inversa entre el tiempo de desplazamiento y el rendimiento académico se logrará con la adaptación de la metodología implementada por Tigre (2014) de Propensity Score Matching con la información de las características de los estudiantes y sus calificaciones en exámenes realizados por el proyecto Niños del Milenio, lo que permitirá hallar a partir de qué duración el tiempo de viaje puede afectar de manera significativa el proceso de aprendizaje de los estudiantes peruanos. Asimismo, se busca cumplir los objetivos específicos, que se centran comprobar si existen efectos heterogéneos en los estudiantes según su zona de residencia rural o urbana debido a la desigual oferta dependiendo del nivel educativo y analizar la linealidad del efecto del tiempo de desplazamiento en el rendimiento académico de los alumnos de primaria y secundaria.

### 3. Metodología

#### 3.1 Descripción de la Base de Datos

Como se mencionó previamente, la base de datos utilizada para el presente trabajo proviene del proyecto longitudinal Niños del Milenio, el cual posee como objetivo recolectar información de individuos desde su infancia y monitorear su crecimiento en el entorno que vive para entender las causas y consecuencias de la pobreza infantil. Este estudio le otorga bastante importancia a la educación en el desarrollo personal y cognitivo de las personas, por lo que se enfoca en identificar diversos factores que inciden en el proceso de aprendizaje de los individuos encuestados. La población total encuestada a nivel internacional es de 12,000 individuos, recolectados en los países de Vietnam, India, Etiopía y Perú. Hasta la actualidad se han realizado 6 rondas en los años 2002, 2006, 2009, 2013, 2016 y 2020-2021. Así, se busca encuestar a los mismos individuos durante el estudio, lo que permitirá monitorear las trayectorias de los individuos encuestados durante la niñez, adolescencia y adultez (Cueto et al., 2018).

Entre los encuestados, la población se divide en 2 cohortes, una mayor y otra menor. En la cohorte mayor, los individuos nacieron en el año 1994 o alrededor de este, mientras que en la cohorte menor los encuestados nacieron entre los años 2001 y 2002. A nivel nacional, la población de la cohorte mayor en la primera ronda era de aproximadamente 750 encuestados, mientras que en la menor era de 2000. Además, el grupo de la cohorte mayor empezó a participar en las encuestas del estudio cuando tenían cerca de 8 años, y la población de la menor comenzó alrededor del año de edad. Por último, la selección de los encuestados fue aleatoria, con la exclusión del 5 por ciento de los distritos más adinerados de cada país, debido al énfasis del estudio en investigar los causantes de la pobreza en estas localidades (Cueto et al., 2018).

A partir de esta base de datos, se plantea utilizar la información de la cohorte menor de los alumnos peruanos de las rondas 3 y 5, así como las rondas que preceden a estas para analizar el impacto del tiempo de desplazamiento en el rendimiento académico de los estudiantes de primaria y secundaria del Perú, considerando los insumos de periodos anteriores. Las rondas 3 y 5 son utilizadas ya que en estas la mayoría de la población de alumnos se encuentra en los niveles educativos de primaria y secundaria. De esta manera, para el examen de matemáticas

del estudio NdM, se han seleccionado 1524 alumnos con información suficiente de diversas características consideradas en el modelo que se encontraban en los niveles educativos primario y secundario en las rondas 3 y 5 respectivamente. Esta muestra ha sido elaborada a partir de la base de datos de la ronda 2 del estudio NdM formada por Boyden (2022a), la base de la ronda 3 realizada por Boyden (2022b), la base de la ronda 5 hecha por Sanchez et al. (2022) y la base construida de las rondas 1,2,3,4 y 5 del estudio NdM elaborada por Boyden (2022c). Finalmente, se han descartado alumnos que posean un tiempo de desplazamiento mayor a 2 horas, debido a la falta de credibilidad y su potencial de distorsionar la verdadera relación entre el tiempo de viaje y el rendimiento académico. Así, se tendrían 1524 y 1518 observaciones para los niveles educativos de primaria y secundaria respectivamente.

### 3.2 Descripción del modelo de estimación

Como en la investigación de Tigre (2014), la estrategia realizada para analizar el efecto del tiempo de viaje de ida en el rendimiento académico de los estudiantes de primaria y secundaria en el Perú se basará en la metodología de evaluación de impacto, cuyo objetivo es determinar el resultado de un programa o política pública sobre los individuos o entidades involucrados en su aplicación. Esto se consigue aislando el efecto de las políticas implementadas mediante la comparación de los resultados de la población que recibió el tratamiento y un escenario contrafactual donde se tienen los resultados de la población no tratada como si hubieran recibido las políticas efectuadas.

Así, como se ve en la ecuación presentada, el estudio buscará dividir la población estudiantil según su tiempo de desplazamiento de ida desde su residencia hasta su institución educativa bajo ciertos intervalos de tiempo para simular los casos de evaluación de impacto que separan la población entre quienes recibieron tratamiento o no. Los individuos cuyo tiempo de viaje exceda el intervalo serán considerados como “alumnos con viajes largos” que en la terminología de la evaluación de impacto sería la población tratada, y los que se demoran menos del intervalo se catalogarán como “alumnos con viajes cortos”, que representaría la población no tratada.

$$y_i = \begin{cases} y_{1i}, & \text{si } T = 1 \\ y_{0i}, & \text{si } T = 0 \end{cases}$$

De esta manera,  $y_{1i}$  denota el rendimiento académico dado que sea del alumno con viajes largos ( $T = 1$ ) y  $y_{0i}$  es el rendimiento del alumno con viajes cortos ( $T = 0$ ). El efecto causal de esta diferencia en el tiempo de viaje de ida se brinda con la resta de los potenciales rendimientos de los grupos de estudiantes,  $\beta_i = y_{1i} - y_{0i}$ . No obstante, como se mencionaba anteriormente, se desea aislar el efecto del tiempo de viaje en el rendimiento académico de los estudiantes, por lo que desea calcular en lo que se denomina en la evaluación de impacto como el efecto promedio del tratamiento sobre los tratados (ATT en inglés), el cual calcula el impacto de cierto programa o política en la población tratada. En este caso, el tratamiento ( $T = 1$ ) indicaría si el alumno se demora más o igual del intervalo propuesto en sus viajes de ida, y si se tardara menos del intervalo no sería descrito como beneficiario del tratamiento ( $T = 0$ ). Con lo explicado previamente, la ecuación planteada adelante denota cómo se calculará el ATT para hallar el efecto del tiempo de viaje de ida sobre el rendimiento académico de los alumnos de primaria y secundaria del Perú (Tigre, 2014).

$$\beta_{ATT} = E[\beta_i | T = 1] = E[y_{1i} - y_{0i} | T = 1]$$

El minuendo de la sustracción es una variable observable ( $y_{1i} | T = 1$ ), pues denota el rendimiento académico de los alumnos con viajes largos; sin embargo, el sustraendo es inobservable ( $y_{0i} | T = 1$ ) ya que indica el rendimiento académico potencial de los alumnos con viajes cortos si hubiesen realizado viajes largos. Para calcular este segundo escenario hipotético se empleará la técnica de Propensity Score Matching (PSM), la cual consiste en la construcción de un grupo de control a partir de la población que no ha recibido tratamiento, que en este caso son los alumnos con viajes cortos, para que estos puedan ser comparados con la población tratada que serían los alumnos con viajes largos. Este grupo control construido servirá para hallar el rendimiento potencial del escenario contrafactual, pues los alumnos con viajes cortos seleccionados serían semejantes a quienes realizan viajes largos (Tigre, 2014).

La elaboración de este grupo de control consiste en calcular la probabilidad de que el alumno sea uno con viajes largos a partir de variables observables (también denominadas covariables) que determinen sus tendencias en la duración de sus viajes. La obtención de esta probabilidad se consigue mediante modelos de elección discreta que utilizan las variables observables mencionadas, ya sea con modelos probit o logit. Con esta probabilidad o "Propensity Score" (PS), se busca emparejar alumnos que

realizan viajes largos con los que tienen viajes cortos que tengan un PS semejante, y así conseguir un grupo de control adecuado a partir de los alumnos con viajes cortos emparejados (Aguilar, 2017).

Adicionalmente, como argumenta Aguilar (2017), para que la implementación del PSM sea viable, se deben cumplir dos supuestos, el de independencia condicional y la existencia de un soporte común. El primero consiste en que la asignación del tratamiento es independiente de los resultados potenciales al condicionar por el vector de covariables utilizadas en la generación del PS, mientras que el segundo señala que, para toda covariable, debe existir tanto población tratada como no tratada, pues sino no se podría generar los emparejamientos. De esta manera, la siguiente ecuación expuesta demuestra el cálculo del ATT con la técnica de Propensity Score Matching, el cual para la presente investigación representaría el efecto del tiempo de desplazamiento de los alumnos con viajes largos en su rendimiento académico.

$$\beta_{ATT} = E[\beta_i | T = 1, p(X_i)] = E[y_{1i} - y_{0i} | T = 1, p(X_i)]$$

Por lo tanto, como sustenta Tigre (2014), el condicionamiento sobre el Propensity Score  $p(X_i)$  en el cálculo del ATT genera que la distribución de las covariables  $X_i$  para los alumnos con viajes cortos sea balanceada de tal manera que sea similar a la distribución de los alumnos con viajes largos. Por lo tanto, el emparejamiento, bajo el supuesto de independencia condicional, descarta el sesgo ocasionado por la asignación no aleatorizada del tratamiento; es decir, si el alumno realiza viajes largos o cortos, por lo que el ATT se puede interpretar de manera causal.

Por último, la creación del escenario contrafactual se realizará inicialmente con el emparejamiento con el k-vecino más cercano con reemplazo. Este método emparejará cada alumno de viajes largos con un número “k” de alumnos de viajes cortos que posea el PS más cercano a este primero dentro del grupo de control. Asimismo, que el emparejamiento sea con reemplazo significa que un alumno con viajes cortos dentro del grupo control puede ser emparejado con más de un alumno con viajes largos debido a que su PS es el más parecido para todos estos últimos alumnos (Tigre, 2014). Además, se buscará calibrar los emparejamientos modificando la diferencia mínima entre los PS de los alumnos con viajes largos y cortos. Esto se realizará modificando el valor del caliper, definido como una fracción de la desviación estándar de la estimación de la probabilidad estimada mediante el modelo logit o probit

(Leite, 2017). Cabe añadir que se propondrán distintos puntos de corte en el tiempo de desplazamiento además del que se usará inicialmente para comprobar la sensibilidad de su efecto en el rendimiento académico de los alumnos, así como se realizarán otros métodos de emparejamiento dentro de la sección de análisis de sensibilidad de los resultados.

### 3.3 Descripción de las variables del modelo de estimación

En esta sección se procederá a nombrar y explicar la variable de tratamiento, las variables de resultados y las covariables que serán implementadas en el emparejamiento entre el grupo de alumnos con viajes largos y el grupo control de alumnos con viajes cortos para hallar el efecto del tiempo de desplazamiento en el rendimiento académico en los alumnos de primaria y secundaria en el Perú. En primer lugar, como se menciona en la subsección anterior, la variable  $T$  de tratamiento denota si el alumno realiza viajes largos o cortos, tomando el valor de 1 si es un alumno con viajes largos y 0 si es un alumno con viajes cortos de acuerdo a un corte de tiempo específico utilizado. Esta variable se consigue a partir del tiempo de desplazamiento de ida desde el domicilio hasta la institución educativa que se registra en la base de datos Niños del Milenio. Este es reportado por el guardián del alumno hasta la ronda 3, mientras que de la ronda 4 en adelante este es reportado por él mismo, el cual es registrado en minutos. Respecto a la variable de resultados  $y$ , esta representa el rendimiento académico de los alumnos en una evaluación en particular, que en este caso sería la de matemáticas. Tanto para los alumnos de primaria como de secundaria se implementará el puntaje estandarizado en la evaluación de matemática de las rondas 3 y 5 respectivamente.

Por tanto, se efectuará dos veces la técnica de PSM para cada corte de tiempo utilizado que divida a los alumnos con viajes largos y cortos, pues se realizarán emparejamientos individuales para los alumnos cuando se encontraban en el nivel educativo primario y secundario respectivamente. Asimismo, en cada emparejamiento también se utilizan covariables de periodos anteriores para aprovechar la característica longitudinal de la base de datos; es decir, para los emparejamientos en la ronda 3 (alumnos de primaria) se implementarán variables observables de los alumnos en la ronda 2 y 1, así como en los emparejamientos de la ronda 5 (alumnos de secundaria) se usarán variables de la ronda 4, 2 y 1. Por último, cabe añadir que las covariables de los emparejamientos entre los niveles educativos son las mismas

pero son configuradas respecto a las rondas utilizadas. A continuación, se presentan las covariables a utilizar en el emparejamiento de los alumnos cuando cursaban primaria y secundaria en sus instituciones educativas en las tablas 2 y 3 respectivamente. La razón de la selección de estas covariables yace en la función de producción educativa acumulada previamente descrita, donde insumos de la demanda educativa, oferta educativa, el entorno del alumno y las habilidades innatas de este en el tiempo inciden en la relación entre el tiempo de viaje y el rendimiento educativo (Hanushek, 1992).



Tabla 2

Covariables de emparejamiento para la ronda 3 (alumnos de primaria)

Variable	Ronda o rondas que utiliza la variable	Denominación y características de la variable	Explicación de la variable
Sexo del alumno	3	Nombre: chsex Promedio: 0.499 Min: 0 Max: 1	Variable dummy que toma el valor de 1 si el alumno es mujer y 0 si es hombre.
Lengua materna español	3	Nombre: NOhablaesp3 Promedio: 0.125 Min: 0 Max: 1	Variable dummy que toma el valor de 1 si la lengua materna del alumno es español y 0 si es cualquier otra lengua.
Retraso en el crecimiento del alumno en la ronda 2	2	Nombre: stuntdummy2 Promedio: 0.312 Min: 0 Max: 1	Variable dummy que toma el valor de 1 si el alumno presentaba retraso en su crecimiento físico en la ronda 2 y 0 si es el caso contrario. Se considera que el alumno presenta retraso si su altura es menor para su edad según el estándar de la Organización Mundial de la Salud.
Modo de transporte utilizado basado en la actividad física	3	Nombre: TDA3 Promedio: 0.748 Min: 0 Max: 1	Variable dummy que toma el valor de 1 si utiliza un modo de transporte activo (caminando o bicicleta) y 0 si emplea un modo de transporte que no involucre ejercicio físico; es decir, inactivo.
Cambio de zona de residencia urbana a rural entre las rondas 1 y 3	1 y 3	Nombre: urbanoArural1a3 Promedio: 0.022 Min: 0 Max: 1	Variable dummy que toma el valor de 1 si la residencia del alumno se trasladó de una zona urbana a una rural entre las rondas 1 y 3, y toma el valor de 0 si es cualquier otro.
Cambio de zona de residencia rural a urbana entre las rondas 1 y 3	1 y 3	Nombre: ruralAurbano1a3 Promedio: 0.055 Min: 0 Max: 1	Variable dummy que toma el valor de 1 si la residencia del alumno se trasladó de una zona rural a una urbana entre las rondas 1 y 3, y toma el valor de 0 si es cualquier otro caso.
Permanencia en zona de residencia urbana entre las rondas 1 y 3	1 y 3	Nombre: solourbano1a3 Promedio: 0.676 Min: 0 Max: 1	Variable dummy que toma el valor de 1 si la residencia del alumno se mantuvo en una zona de residencia urbana entre las rondas 1 y 3, y toma el valor de 0 si es cualquier otro caso.

Educación de la madre del alumno	3	Nombre: momeduSECU3 Promedio: 0.377 Min: 0 Max: 1	Variable dummy que toma el valor de 1 si la madre del alumno posee una educación secundaria o superior, y 0 si es el caso contrario.
Número de familiares en la residencia	3	Nombre: hhsiz3 Promedio: 5.399 Min: 2 Max: 16	Variable discreta que representa el número de familiares que conviven en la misma residencia que la del alumno.
Índice de riqueza en la ronda 1	1	Nombre: wi1 Promedio: 0.438 Min: 0.02 Max: 0.910	Variable que representa el nivel de riqueza de la familia del alumno en la ronda 1. Construida a partir de los índices de acceso a servicios, calidad del hogar y el consumo de bienes durables. Variable continua cuyo valor máximo es 1 y mínimo 0.
Índice de riqueza en la ronda 2	2	Nombre: wi2 Promedio: 0.480 Min: 0.02 Max: 0.922	Variable que representa el nivel de riqueza de la familia del alumno en la ronda 2. Construida a partir de los índices de acceso a servicios, calidad del hogar y el consumo de bienes durables. Variable continua cuyo valor máximo es 1 y mínimo 0.
Índice de riqueza en la ronda 3	3	Nombre: wi3 Promedio: 0.553 Min: 0.031 Max: 0.927	Variable que representa el nivel de riqueza de la familia del alumno en la ronda 3. Construida a partir de los índices de acceso a servicios, calidad del hogar y el consumo de bienes durables. Variable continua cuyo valor máximo es 1 y mínimo 0.
Matriculado en Institución educativa de gestión alterna	3	Nombre: entype3otros Promedio: 0.028 Min: 0 Max: 1	Variable dummy que toma el valor de 1 si el alumno está matriculado en una institución educativa de gestión alterna (que no sea privada o pública) y 0 si está matriculado en una con diferente gestión.
Matriculado en Institución educativa de gestión privada	3	Nombre: entype3privado Promedio: 0.136 Min: 0 Max: 1	Variable dummy que toma el valor de 1 si el alumno está matriculado en un instituto educativo privado y 0 si está matriculado en uno con diferente gestión.
Puntaje obtenido en la evaluación de desarrollo cognitivo (CDA) en la ronda 2	2	Nombre: CDA Promedio: 8.503 Min: 0 Max: 12	Variable discreta que representa el puntaje obtenido en la prueba CDA de los alumnos en la ronda 2. El puntaje de la prueba posee como mínimo 0 puntos y máximo 12 puntos. Utilizada en el emparejamiento para hallar el resultado potencial del rendimiento académico del examen de matemática en la ronda 3 como habilidades innatas.

Fuente: Boyden (2022a, 2022b, 2022c) y Sanchez et al. (2022). Elaboración propia.

Tabla 3

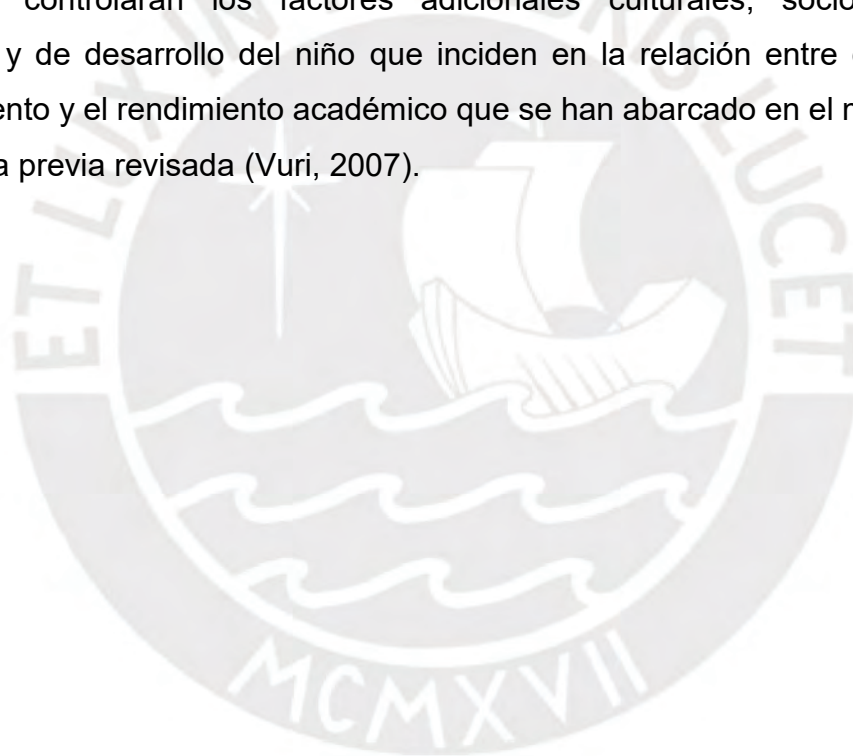
Covariables de emparejamiento para la ronda 5 (alumnos de secundaria)

Variable	Ronda o rondas que utiliza la variable	Denominación y características de la variable	Explicación de la variable
Sexo del alumno	5	Nombre: chsex Promedio: 0.5 Min: 0 Max: 1	Variable dummy que toma el valor de 1 si el alumno es mujer y 0 si es hombre.
Lengua materna español	5	Nombre: NOhablaesp5 Promedio: 0.126 Min: 0 Max: 1	Variable dummy que toma el valor de 1 si la lengua materna del alumno es español y 0 si es cualquier otra lengua.
Retraso en el crecimiento del alumno en la ronda 2	2	Nombre: stuntdummy2 Promedio: 0.314 Min: 0 Max: 1	Variable dummy que toma el valor de 1 si el alumno presentaba retraso en su crecimiento físico en la ronda 2 y 0 si es el caso contrario. Se considera que el alumno presenta retraso si su altura es menor para su edad según el estándar de la Organización Mundial de la Salud.
Modo de transporte utilizado basado en la actividad física	5	Nombre: TDA5 Promedio: 0.603 Min: 0 Max: 1	Variable dummy que toma el valor de 1 si utiliza un modo de transporte activo que requiera actividad física (caminando o bicicleta) y 0 si emplea un modo de transporte que no requiere ejercicio físico; es decir, inactivo.
Cambio de zona de residencia urbana a rural entre las ronda 1 y 5	1 y 5	Nombre: urbanoArural1a5 Promedio: 0.024 Min: 0 Max: 1	Variable dummy que toma el valor de 1 si la residencia del alumno se trasladó de una zona urbana a una rural entre las rondas 1 y 3, y toma el valor de 0 si es cualquier otro.
Cambio de zona de residencia rural a urbana entre las rondas 1 y 5	1 y 5	Nombre: ruralAurbano1a5 Promedio: 0.079 Min: 0 Max: 1	Variable dummy que toma el valor de 1 si la residencia del alumno se trasladó de una zona rural a una urbana entre las rondas 1 y 5, y toma el valor de 0 si es cualquier otro caso.
Permanencia en zona de residencia urbana entre las rondas 1 y 5	1 y 5	Nombre: solourbano1a5 Promedio: 0.674 Min: 0 Max: 1	Variable dummy que toma el valor de 1 si la residencia del alumno se mantuvo en una zona de residencia urbana entre las rondas 1 y 5, y toma el valor de 0 si es cualquier otro caso.

Educación de la madre del alumno	5	Nombre: momeduSECU5 Promedio: 0.412 Min: 0 Max: 1	Variable dummy que toma el valor de 1 si la madre del alumno posee una educación secundaria o superior, y 0 si es el caso contrario.
Número de familiares en la residencia	5	Nombre: hhsiz5 Promedio: 5.238 Min: 2 Max: 16	Variable discreta que representa el número de familiares que conviven en la misma residencia que la del alumno.
Índice de riqueza en la ronda 1	1	Nombre: wi1 Promedio: 0.437 Min: 0.02 Max: 0.910	Variable que representa el nivel de riqueza de la familia del alumno en la ronda 1. Construida a partir de los índices de acceso a servicios, calidad del hogar y el consumo de bienes durables. Variable continua cuyo valor máximo es 1 y mínimo 0.
Índice de riqueza en la ronda 4	4	Nombre: wi4 Promedio: 0.603 Min: 0.01 Max: 0.953	Variable que representa el nivel de riqueza de la familia del alumno en la ronda 4. Construida a partir de los índices de acceso a servicios, calidad del hogar y el consumo de bienes durables. Variable continua cuyo valor máximo es 1 y mínimo 0.
Índice de riqueza en la ronda 5	5	Nombre: wi5 Promedio: 0.637 Min: 0.07 Max: 0.936	Variable que representa el nivel de riqueza de la familia del alumno en la ronda 5. Construida a partir de los índices de acceso a servicios, calidad del hogar y el consumo de bienes durables. Variable continua cuyo valor máximo es 1 y mínimo 0.
Matriculado en Institución educativa de gestión alterna	5	Nombre: entype5otros Promedio: 0.029 Min: 0 Max: 1	Variable dummy que toma el valor de 1 si el alumno está matriculado en una institución educativa de gestión alterna (que no sea privada o pública) y 0 si está matriculado en una con diferente gestión.
Matriculado en Institución educativa de gestión privada	5	Nombre: entype5privado Promedio: 0.146 Min: 0 Max: 1	Variable dummy que toma el valor de 1 si el alumno está matriculado en una institución educativa privada y 0 si está matriculado en una con diferente gestión.
Puntaje obtenido en la evaluación de desarrollo cognitivo (CDA) en la ronda 2	2	Nombre: CDA Promedio: 8.497 Min: 0 Max: 12	Variable discreta que representa el puntaje obtenido en la prueba CDA de los alumnos en la ronda 2. El puntaje de la prueba posee como mínimo 0 puntos y máximo 12 puntos. Utilizada en el emparejamiento para hallar el resultado potencial del rendimiento académico del examen de matemática en la ronda 5 como habilidades innatas.

Fuente: Boyden (2022a, 2022b, 2022c) y Sanchez et al. (2022). Elaboración propia.

Cabe añadir que las covariables en la ronda 3 y 5 controlan los efectos externos provocados por estas en el mismo periodo. Asimismo, las que se encuentran en periodos anteriores buscan aislar el efecto del tiempo de desplazamiento en el rendimiento académico de la acumulación de los diversos insumos de las habilidades innatas del alumno y la oferta y demanda educativa. De esta manera, las variables de zona de residencia buscan controlar los cambios geográficos de la residencia; la de retraso de crecimiento, monitorear las implicancias de un desarrollo físico inadecuado; y los de índice de riqueza pretenden controlar el cambio de potenciales mejoras o empeoramiento del nivel socioeconómico de la familia al inicio de la vida de los alumnos y cuando se encuentran antes de ingresar a los niveles educativos de primaria y secundaria en las rondas 3 y 5, utilizando los índices de las rondas 1, 2, y 4. Así, se controlarán los factores adicionales culturales, socioeconómicos, geográficos y de desarrollo del niño que inciden en la relación entre el tiempo de desplazamiento y el rendimiento académico que se han abarcado en el marco teórico y la literatura previa revisada (Vuri, 2007).



#### 4. Hechos estilizados

En esta sección se analizarán las covariables en el emparejamiento de los alumnos en las rondas 3 y 5 del estudio Niños del Milenio cuando se encontraban en el nivel de primaria y secundaria respectivamente, así como se presentará la relación preliminar entre el rendimiento académico y el tiempo de desplazamiento. El corte inicial utilizado para diferenciar los alumnos es de 20 minutos, en el que los alumnos que se demoren igual o más que este tiempo en sus viajes de ida a la institución educativa serán considerados alumnos con viajes largos, mientras los que se tarden menos de 20 minutos se clasificarán como alumnos con viajes cortos. Este corte ha sido inicialmente utilizado como referencia para evaluar el efecto del tiempo de desplazamiento en el rendimiento académico en un periodo de tiempo determinado, para después comparar sus resultados con otros puntos de corte de tiempo de viaje mayores en la estimación de PSM, específicamente con los cortes de 25 y 30 minutos.

Asimismo, la elección de estos cortes de tiempo se encuentra respaldado por el tiempo de viaje de ida promedio encontrado para los alumnos de primaria y secundaria en la encuesta nacional del uso del tiempo (ENUT) del 2024. Estas encuestas realizadas por el Instituto Nacional de Estadística e Informática (INEI) brindan una amplia perspectiva de cómo las personas de 12 años o mayores distribuyen su tiempo diario en las actividades que realizan cotidianamente. La información se recoge principalmente utilizando un horario ficticio, el cual el entrevistado completa mencionando las actividades que ha realizado el día anterior. Así, en la versión del 2024, se entrevistó un total de 6056 viviendas a nivel nacional, de los cuales se utilizaron 107 estudiantes de primaria y 1227 estudiantes de secundaria<sup>1</sup>. De esta manera, se halló que los alumnos de primaria tardan en promedio 22.52 minutos en trasladarse a su institución educativa, mientras que los de secundaria se demoran 23.77 minutos. Por lo tanto, los cortes de tiempo propuestos para este trabajo podrían brindar una mejor perspectiva de si alumnos con tiempos de viaje mayores pueden tener un efecto más agudo en su rendimiento académico que quienes realizan trayectos de ida más breves (INEI, 2024b).

---

<sup>1</sup> La encuesta nacional del uso del tiempo del 2010 no fue utilizada al implementar la variable de tiempo de viaje como la suma de ida y vuelta al instituto educativo.

#### 4.1 Análisis de covariables utilizadas en la ronda 3

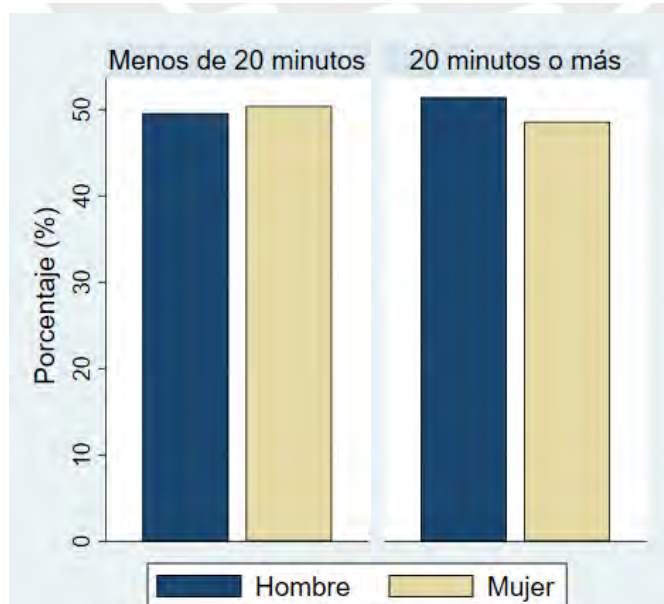
En la tercera ronda, el 23.43% de la población de alumnos en primaria (357 observaciones) es conformado por alumnos con viajes largos de ida con una duración igual o mayor que 20 minutos hacia sus instituciones educativas, mientras que el 76.57% restante (1167 observaciones) está compuesto por los alumnos con viajes cortos de ida con una duración menor que 20 minutos. A partir de la información mencionada previamente, se analizarán los diferentes subgrupos de covariables usadas en esta ronda según el tiempo de desplazamiento.

##### 4.1.1 Covariables relacionadas a las características del alumno

Las distribuciones de las covariables de la demanda educativa que caracterizan al alumno según el tiempo de desplazamiento se pueden observar en los gráficos del 4 al 7. Estas son el sexo, la lengua materna, si el alumno padece de retraso en su crecimiento físico y el modo de transporte que utiliza el alumno. Respecto al sexo, se puede apreciar que la distribución se mantiene equitativa entre los alumnos con viajes cortos y largos, con una leve representación mayor de los hombres en el segundo grupo de alumnos, pasando a representar del 49.5% al 51.5% de la población estudiantil de primaria. En contraste, los alumnos cuya lengua materna no es el español son más representativos en el grupo que realizan viajes largos comparado al de viajes cortos, conformando el 23% y 9% de las respectivas poblaciones.

Gráfico 4

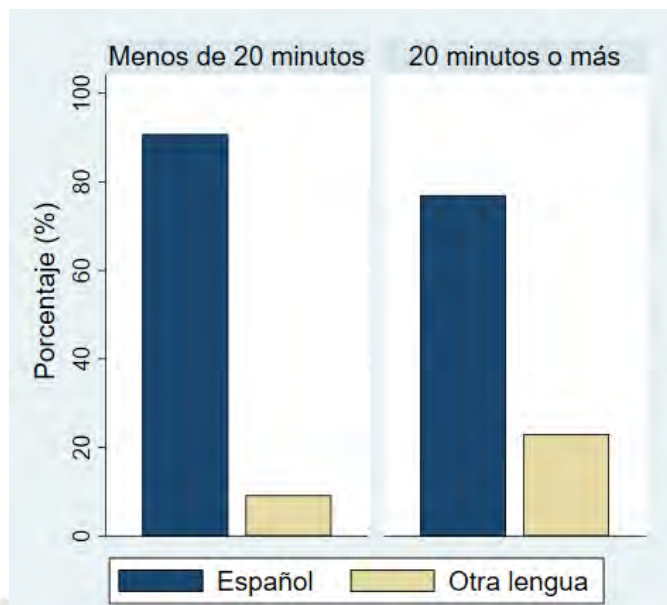
Distribución por sexo según el tiempo de desplazamiento (alumnos de primaria)



Fuente: Boyden (2022c). Elaboración propia.

Gráfico 5

Distribución por la lengua materna según el tiempo de desplazamiento (alumnos de primaria)

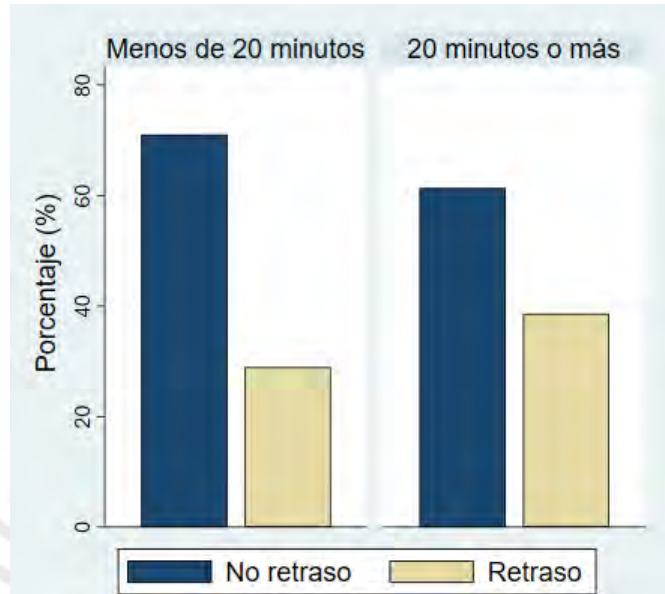


Fuente: Boyden (2022c). Elaboración propia.

Por último, en el caso del retraso de crecimiento del alumno, se puede observar que la proporción que lo padece es de 29% si es un alumno con viajes cortos y 38% si es un alumno con viajes largos; mientras que en lo que respecta al modo de transporte, el 76% y 69% de los alumnos que realizan viajes cortos y largos utilizan un modo de transporte activo respectivamente.

Gráfico 6

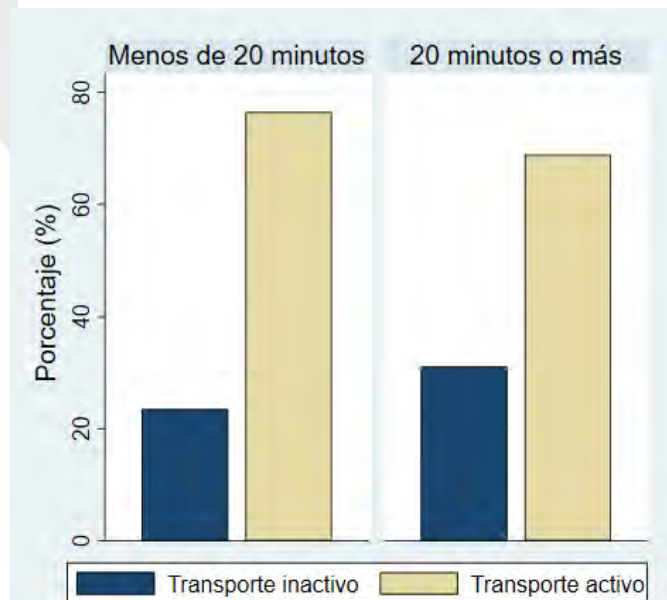
Porcentaje de alumnos con presencia de retraso en crecimiento en la ronda 2 según el tiempo de desplazamiento (alumnos de primaria)



Fuente: Boyden (2022a, 2022c). Elaboración propia.

Gráfico 7

Distribución del modo de transporte según el tiempo de desplazamiento (alumnos de primaria)



Fuente: Fuente: Boyden (2022b). Elaboración propia. \*A partir de García-Hermoso et al. (2017), un modo de transporte activo se define como un medio para movilizarse que requiera actividad física, mientras que un modo inactivo carece de esta característica.

#### 4.1.2 Covariables relacionadas a características de la familia del alumno

Las covariables de la demanda educativa que caracterizan la familia del alumno son comprendidas por las que detallan la zona de residencia de la familia en el tiempo, la educación de la madre, el número de familiares viviendo en el mismo hogar que el alumno y el nivel de riqueza de la familia a través del tiempo. Desde los gráficos 8 al 11 se puede apreciar la distribución de las distintas configuraciones de la zona geográfica de residencia en el tiempo. En primer lugar, la distribución de alumnos que se han desplazado entre la ronda 1 y 3 de una zona urbana a rural es casi idéntica en los grupos de alumnos, conformando el 2% de los que realizan viajes cortos y 3% de los de viajes largos. En segundo lugar, la proporción de alumnos que se han trasladado de residencia de una zona rural a urbana es similar entre los que realizan viajes cortos y largos, donde su representatividad alcanza el 5% y 8% respectivamente. En tercer lugar, los alumnos que se han mantenido en una zona de residencia urbana entre las rondas forman la mayoría de los alumnos en ambos grupos, siendo el 73% y 52% de los que realizan viajes cortos y largos. Por último, el 21% y el 38% de los estudiantes de viajes cortos y largos han permanecido en una zona de residencia rural entre las rondas 1 y 3.

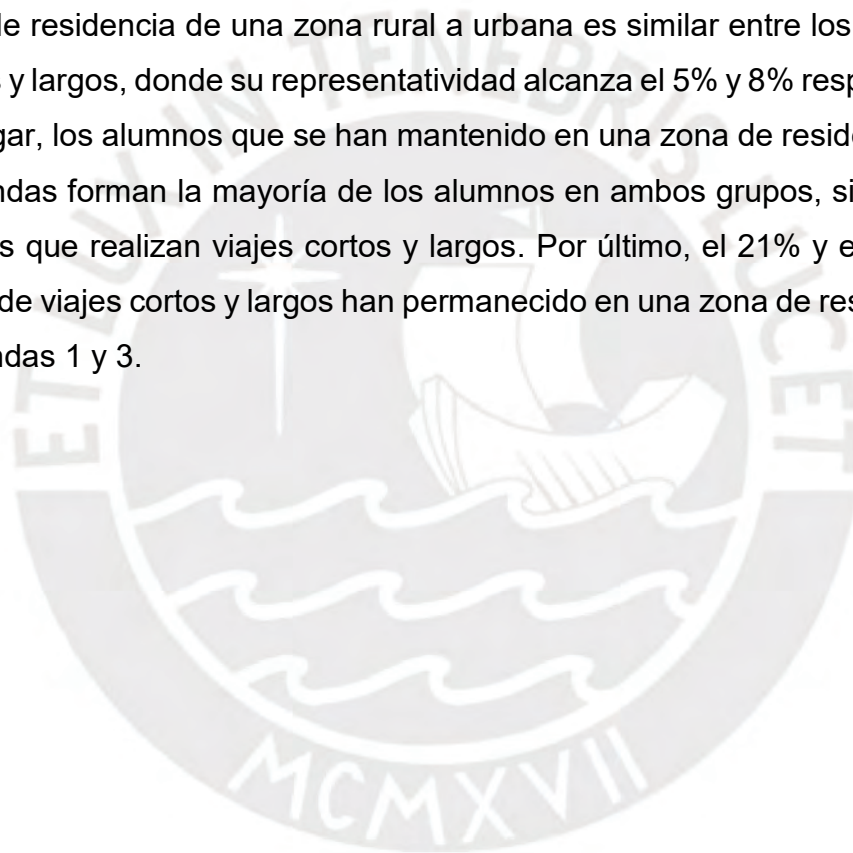
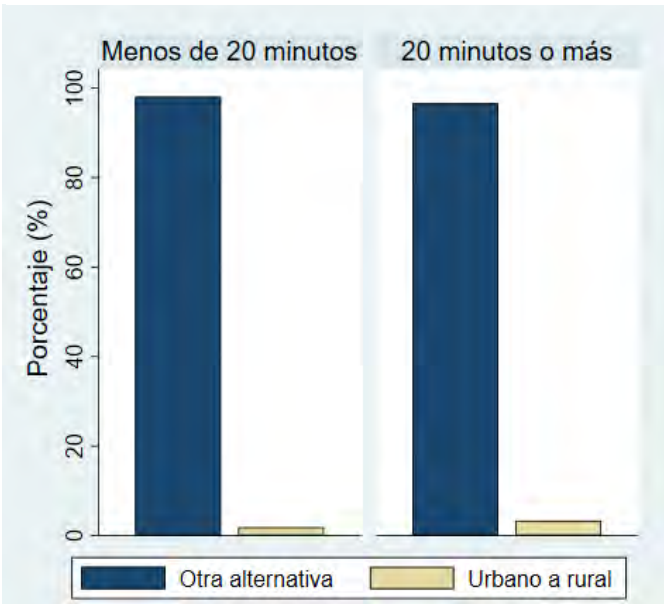


Gráfico 8

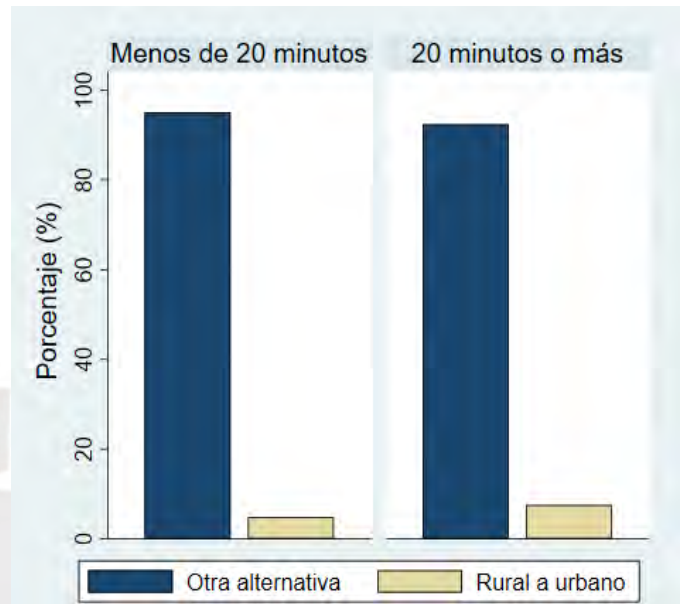
Distribución de traslado de zona de residencia urbana a rural entre la ronda 1 y 3 según el tiempo de desplazamiento (alumnos de primaria)



Fuente: Boyden (2022c). Elaboración propia.

Gráfico 9

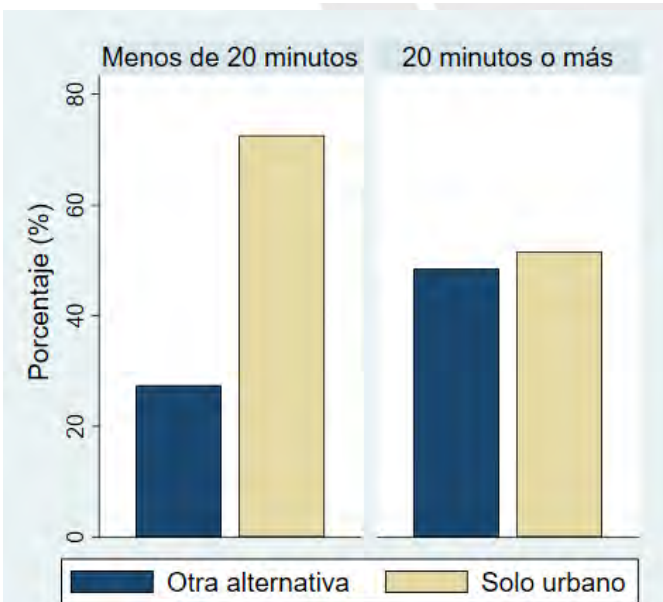
Distribución de traslado de zona de residencia rural a urbana entre la ronda 1 y 3 según el tiempo de desplazamiento (alumnos de primaria)



Fuente: Boyden (2022c). Elaboración propia.

Gráfico 10

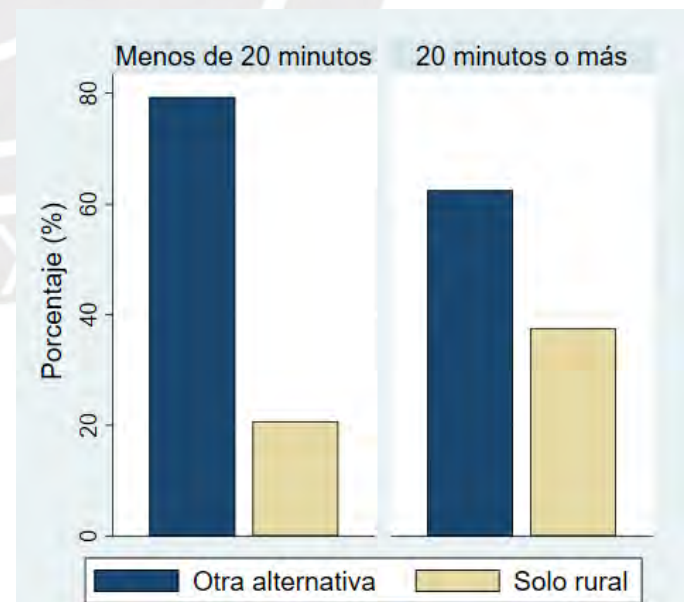
Distribución de la permanencia de la residencia en una zona urbana entre la ronda 1 y 3 según el tiempo de desplazamiento (alumnos de primaria)



Fuente: Boyden (2022c). Elaboración propia.

Gráfico 11

Distribución de la permanencia de la residencia en una zona rural entre la ronda 1 y 3 según el tiempo de desplazamiento (alumnos de primaria)

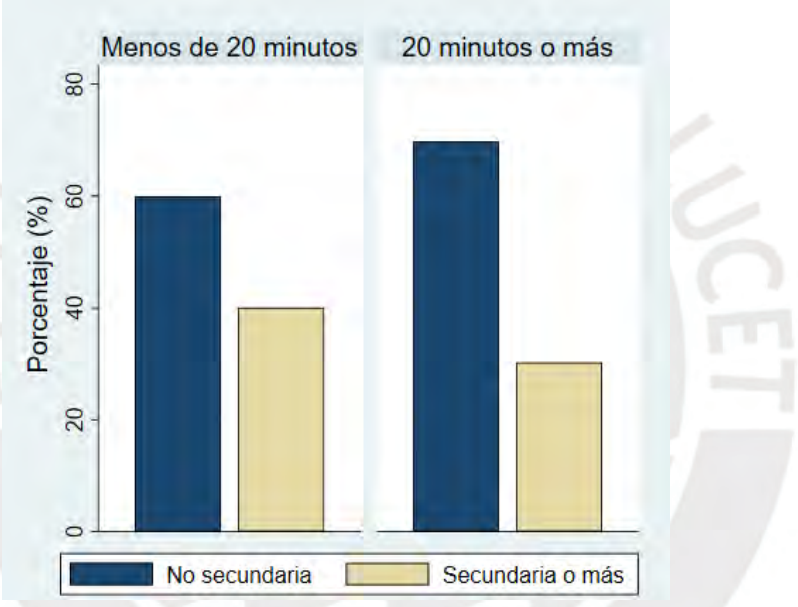


Fuente: Boyden (2022c). Elaboración propia.

Por otra parte, la distribución a partir de la educación de la madre se puede observar en el gráfico 12, en la cual el 40% de los alumnos de viajes cortos son criados por una madre con una educación secundaria o mayor, mientras que solo el 30% de los alumnos de viajes largos tienen a una madre con este nivel de educación. De manera contraria, como se aprecia en el gráfico 13, la distribución de la cantidad de familiares residentes en el hogar es bastante similar entre los grupos de alumnos, siendo el promedio de ambos alrededor de 5.4 familiares.

Gráfico 12

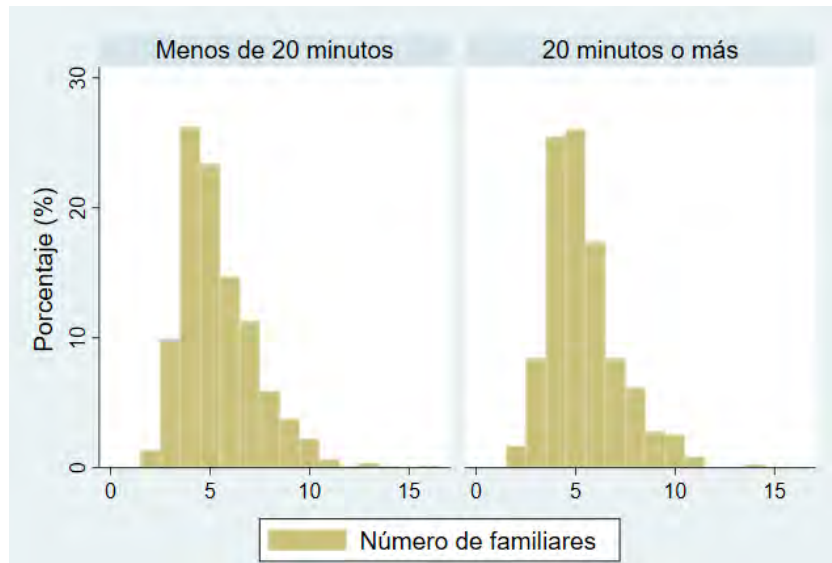
Distribución de la educación de la madre a partir de si posee una educación secundaria o mayor según el tiempo de desplazamiento (alumnos de primaria)



Fuente: Boyden (2022c). Elaboración propia.

Gráfico 13

Distribución del número de familiares residentes en el hogar según el tiempo de desplazamiento (alumnos de primaria)

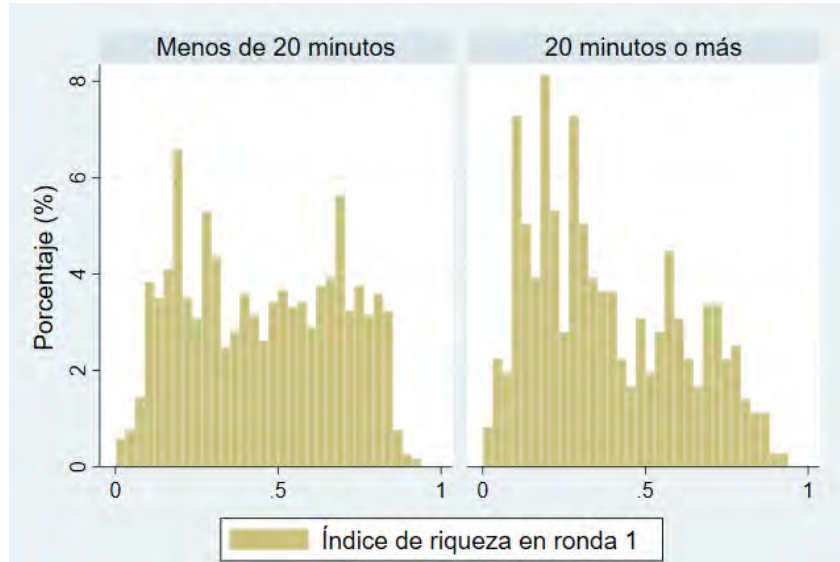


Fuente: Boyden (2022c). Elaboración propia.

Por último, los gráficos 14, 15 y 16 muestran las distribuciones de los índices de riqueza de las familias de los alumnos en las rondas 1,2 y 3 respectivamente. Como se mencionaba en la sección de descripción de las covariables, este índice es construido a partir de subíndices que evalúan el acceso a servicios, calidad del hogar y el consumo de bienes durables de la familia del alumno (Boyden, 2022c). En todas las rondas, los alumnos con viajes cortos poseen un mayor índice de riqueza promedio que los que realizan viajes largos, siendo la diferencia entre los promedios de 0.072, 0.068 y 0.079 puntos en las rondas 1,2 y 3 respectivamente. Asimismo, se puede apreciar en las 3 rondas que la distribución del índice de riqueza de los alumnos con viajes cortos es equitativa, mientras la de los que realizan viajes largos es más desequilibrada, concentrándose la población entre los valores bajos y altos del índice.

Gráfico 14

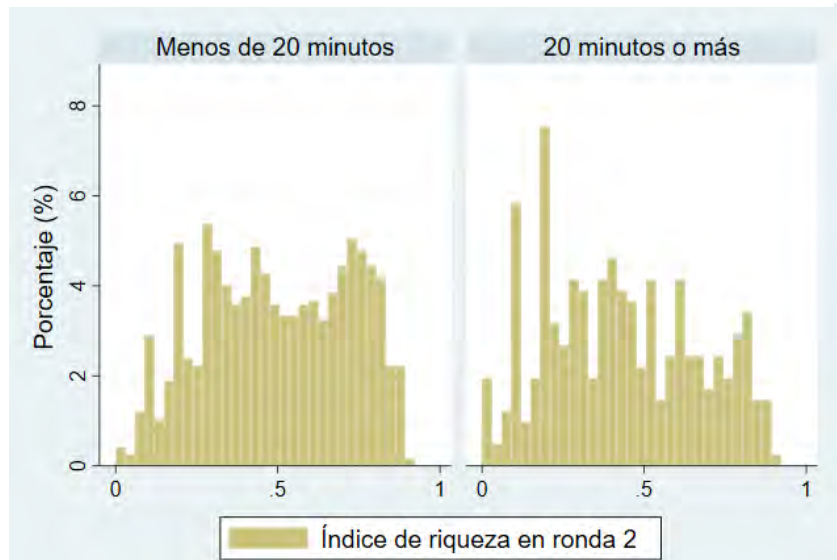
Distribución del índice de riqueza de la familia en la ronda 1 según el tiempo de desplazamiento (alumnos de primaria)



Fuente: Boyden (2022c). Elaboración propia.

Gráfico 15

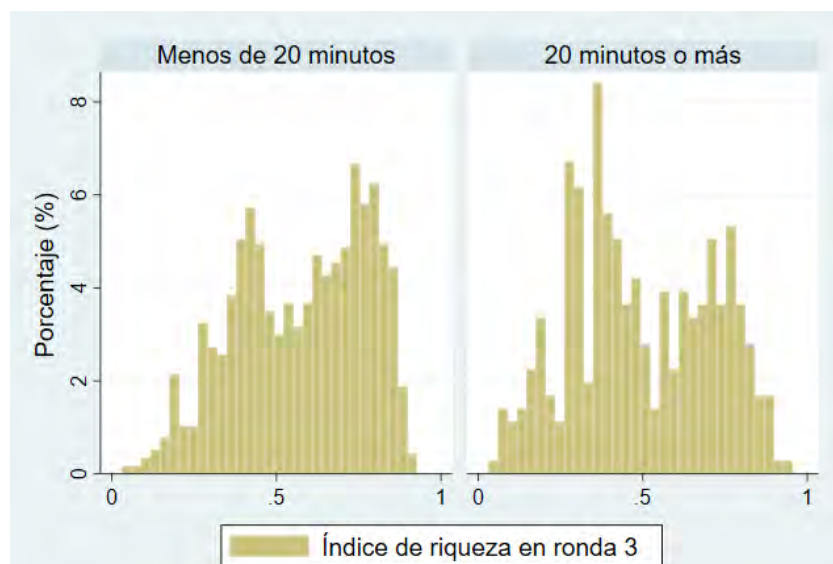
Distribución del índice de riqueza de la familia en la ronda 2 según el tiempo de desplazamiento (alumnos de primaria)



Fuente: Boyden (2022c). Elaboración propia.

Gráfico 16

Distribución del índice de riqueza de la familia en la ronda 3 según el tiempo de desplazamiento (alumnos de primaria)



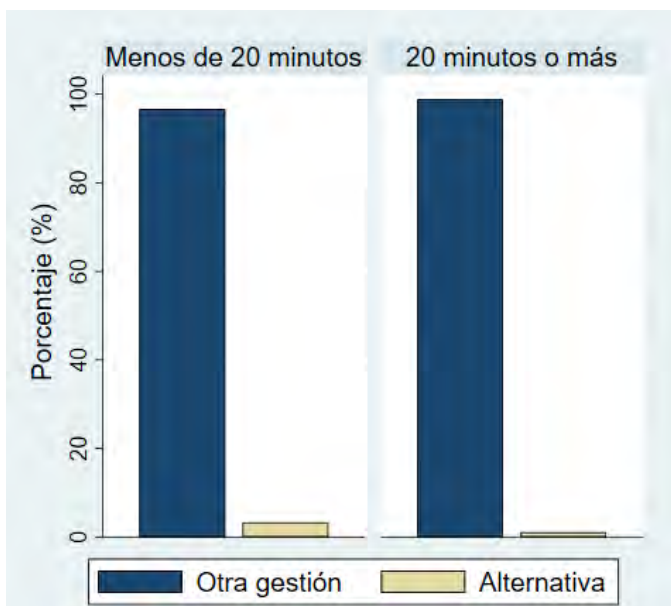
Fuente: Boyden (2022c). Elaboración propia.

#### 4.1.3 Covariables de la oferta educativa y habilidades innatas del alumno

Las covariables que caracterizan la oferta educativa son las que identifican el tipo de gestión que se rige en las instituciones educativas a las que acuden los alumnos de primaria en la ronda 3. Así, en los gráficos 17, 18 y 19 se pueden observar las distribuciones de los alumnos a partir de si se encuentran matriculados en una institución educativa con una gestión alternativa, privada o pública. La proporción de alumnos matriculados en instituciones con gestión alterna es reducida para ambos grupos, representando el 3% para los alumnos con viajes cortos y el 1% para los alumnos con viajes largos. Mientras tanto, el 15% de los que realizan viajes cortos se encuentran matriculados en instituciones con gestión privada, así como el 9% para los alumnos que realizan viajes largos. Por tanto, el 82% de los alumnos de viajes cortos se encontraban matriculados en una institución con gestión pública, y para el caso de los que realizan viajes largos, esta proporción es de 90%.

Gráfico 17

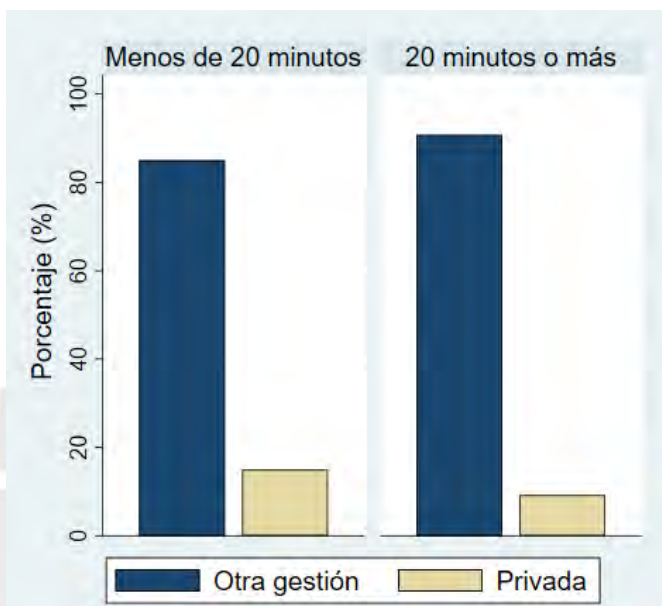
Distribución de la institución educativa matriculada a partir de la gestión alternativa según el tiempo de desplazamiento (alumnos de primaria)



Fuente: Boyden (2022c). Elaboración propia.

Gráfico 18

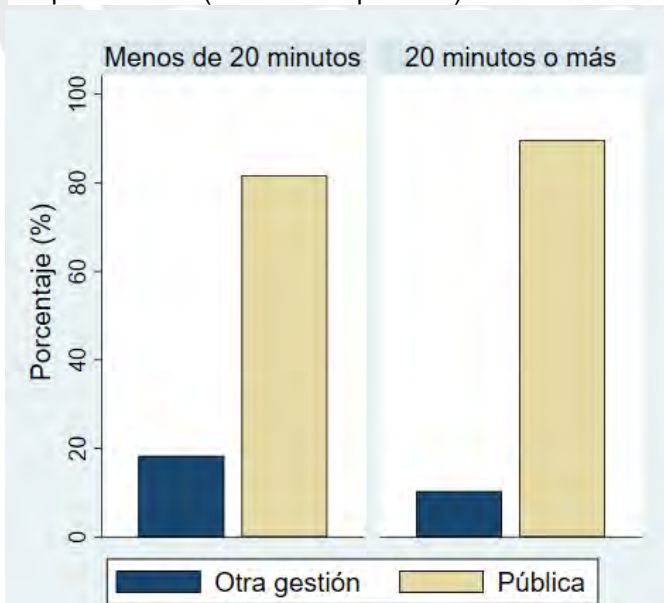
Distribución de la institución educativa matriculada a partir de la gestión privada según el tiempo de desplazamiento (alumnos de primaria)



Fuente: Boyden (2022c). Elaboración propia.

Gráfico 19

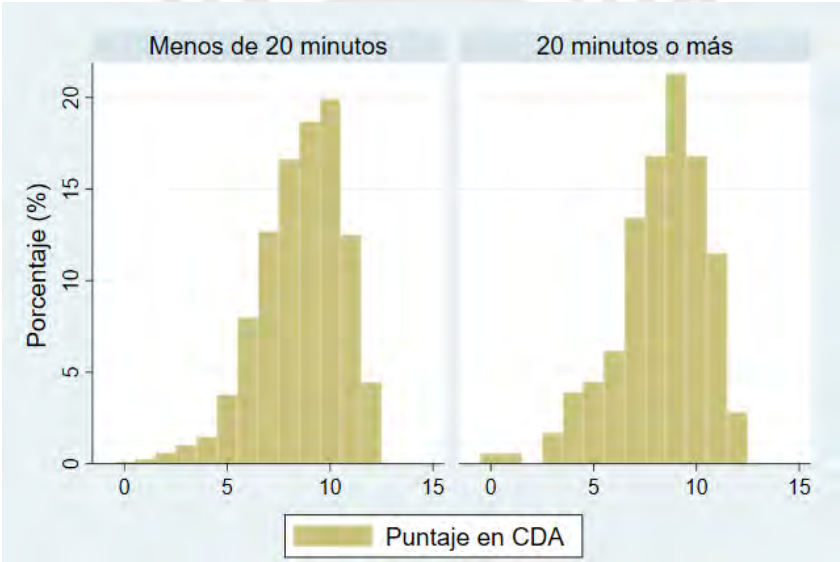
Distribución de la institución educativa matriculada a partir de la gestión pública según el tiempo de desplazamiento (alumnos de primaria)



Fuente: Boyden (2022c). Elaboración propia.

Finalmente, la covariable que define las habilidades innatas de los alumnos es su puntaje obtenido en la evaluación de desarrollo cognitivo realizada en la ronda 2. Como se puede observar en el gráfico 20, la distribución del puntaje en el examen CDA es similar entre los grupos de alumnos de viajes cortos y largos, cuyos promedios del puntaje obtenido son de 8.56 y 8.28 puntos respectivamente, siendo el máximo obtenible 12. Adicionalmente, se puede apreciar que los que realizan viajes largos poseen valores menores más extremos, así como una menor proporción de alumnos que hayan conseguido calificaciones mayores a 10 puntos.

Gráfico 20  
 Distribución del puntaje obtenido en el examen CDA en la ronda 2 según el tiempo de desplazamiento (alumnos de primaria)



Fuente: Boyden (2022a). Elaboración propia.

4.2 Análisis de covariables utilizadas en la ronda 5

En el caso de la quinta ronda, con una muestra de 1518 alumnos de secundaria y con el punto de corte de 20 minutos, el 31.03% de la población (471 observaciones) es conformado por alumnos con viajes largos de ida con una duración mayor o igual a 20 minutos hacia sus instituciones educativas, mientras que el 68.97% restante (1047 observaciones) es formado por los alumnos con viajes cortos de ida con una duración menor de 20 minutos. De manera análoga, se utilizan las mismas covariables y subgrupos que en el caso de los alumnos de primaria en la ronda 3, pero ahora se ajustan al periodo de su etapa secundaria en la ronda 5. A partir de la información

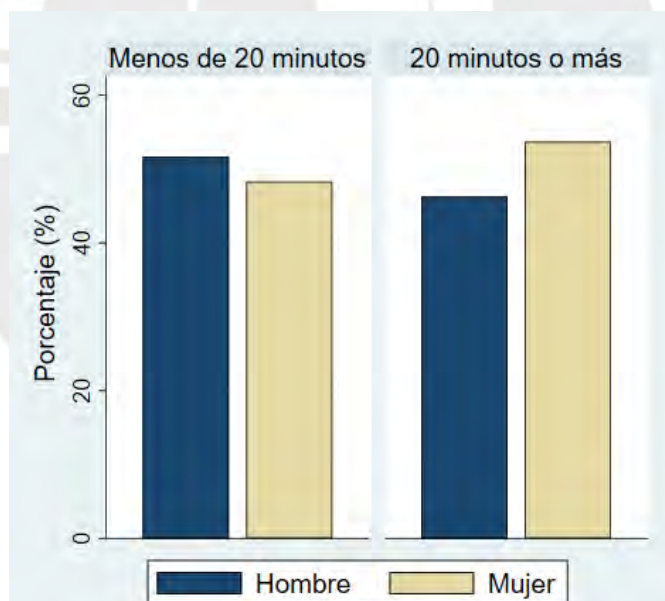
mencionada previamente, se analizarán los diferentes subgrupos de covariables usadas en esta ronda según el tiempo de desplazamiento.

#### 4.2.1 Covariables relacionadas a las características del alumno

Las distribuciones de las covariables que caracterizan al alumno según el tiempo de desplazamiento se pueden observar en los gráficos del 21 al 24. En lo que concierne al sexo, se puede apreciar que la distribución se mantiene equitativa entre los alumnos con viajes cortos y largos, con una leve representación mayor de las mujeres en el segundo grupo de alumnos, pasando a representar del 46% al 52% de la población estudiantil. Por otra parte, la diferencia de la población de alumnos que no poseen como lengua materna el español es más pronunciada entre los grupos, ya que estos representan el 9% para los que realizan viajes cortos y 21% para los estudiantes con viajes largos.

Gráfico 21

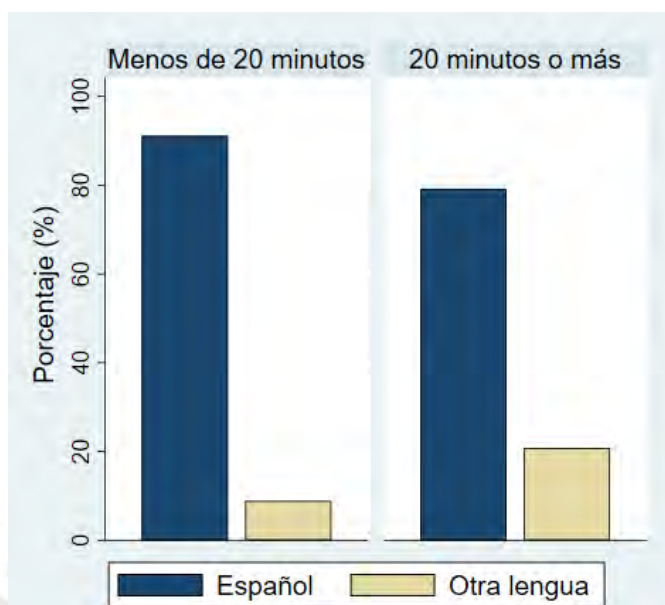
Distribución por el sexo según el tiempo de desplazamiento (alumnos de secundaria)



Fuente: Boyden (2022c). Elaboración propia.

Gráfico 22

Distribución por la lengua materna según el tiempo de desplazamiento (alumnos de secundaria)

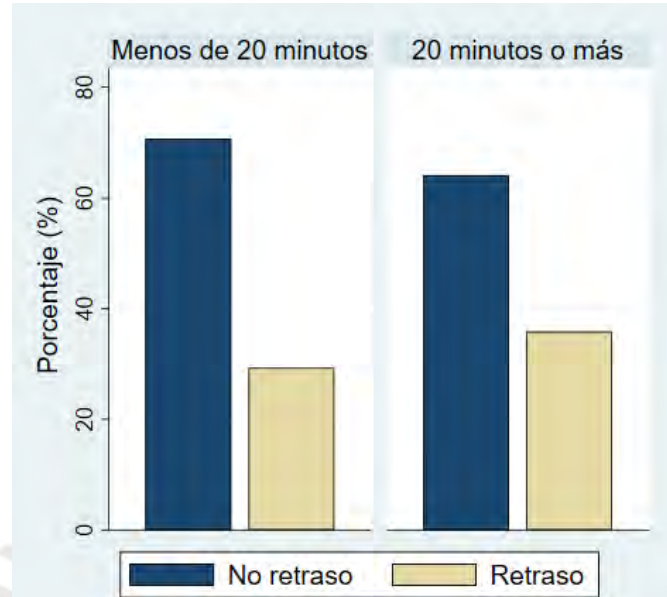


Fuente: Boyden (2022c). Elaboración propia.

En el caso del retraso de crecimiento del alumno, se puede observar que la proporción que lo padece es de 29% si es un alumno con viajes cortos y 36% si es un alumno con viajes largos. Finalmente, la diferencia en la distribución del uso un modo de transporte activo es remarcable, ya que el 64% de los alumnos de viajes cortos optan por un transporte activo ya sea caminando o con bicicleta, mientras que el 51% de los que realizan viajes largos eligen este modo de transporte que requiere ejercicio físico.

Gráfico 23

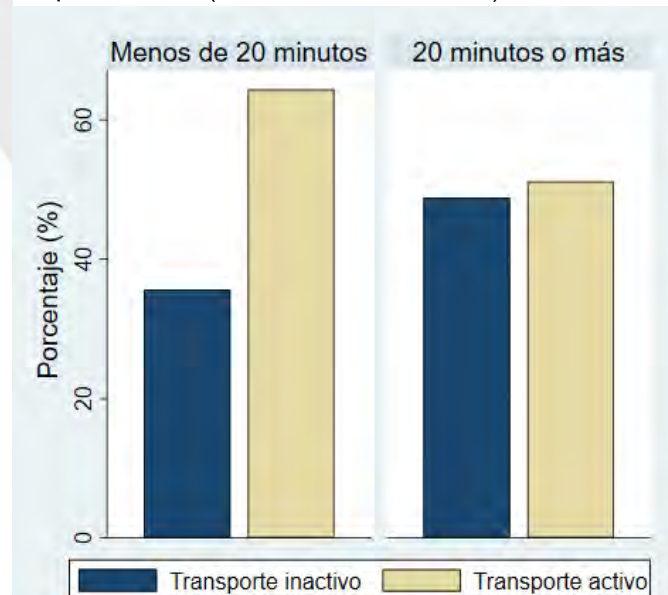
Porcentaje de alumnos con presencia de retraso en crecimiento en la ronda 2 según el tiempo de desplazamiento (alumnos de secundaria)



Fuente: Boyden (2022a, 2022c). Elaboración propia.

Gráfico 24

Distribución del modo de transporte según el tiempo de desplazamiento (alumnos de secundaria)



Fuente: Fuente: Sanchez et al. (2022). Elaboración propia. \*A partir de García-Hermoso et al. (2017), un modo de transporte activo se define como un medio para moverse que requiera actividad física, mientras que un modo inactivo carece de esta característica.

#### 4.2.2 Covariables relacionadas a características de la familia del alumno

De manera similar a los alumnos de primaria, los gráficos del 25 al 28 detallan las diferentes configuraciones de permanencia o cambio de zona de residencia entre las rondas 1 y 5 para los alumnos en el nivel secundario. La proporción de alumnos que han alternado de una residencia urbana a una rural es bastante similar, siendo el 2% para el grupo de alumnos con viajes cortos y 4% para los que realizan viajes largos. Por otro lado, la población que ha cambiado de una zona rural a una urbana representa una proporción ligeramente mayor para los alumnos de viajes cortos que para los que realizan viajes largos, formando el 9% y 6% de estos grupos respectivamente. La porción de alumnos que se han mantenido en una zona urbana también es mayor para los que realizan viajes cortos, pero por un margen más notorio, conformando el 74% para estos y el 52% para los alumnos que realizan viajes largos. Por último, la proporción de la población que ha permanecido en zonas rurales es mayor para los alumnos que realizan viajes largos, siendo el 38%, mientras que en el grupo que realizan viajes cortos representan el 15%.

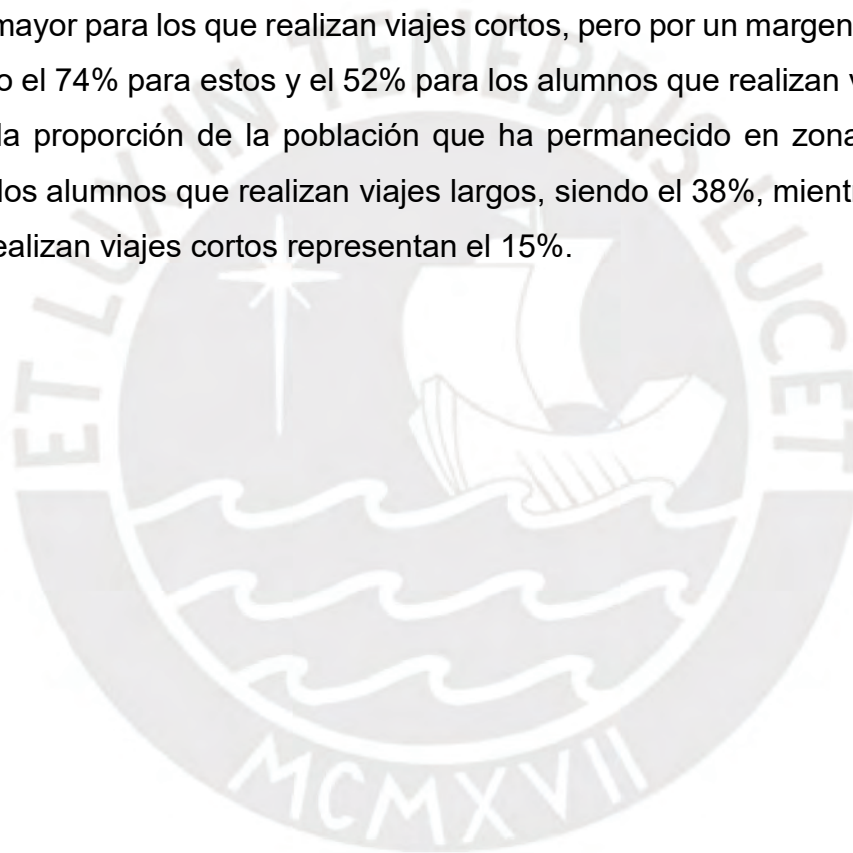
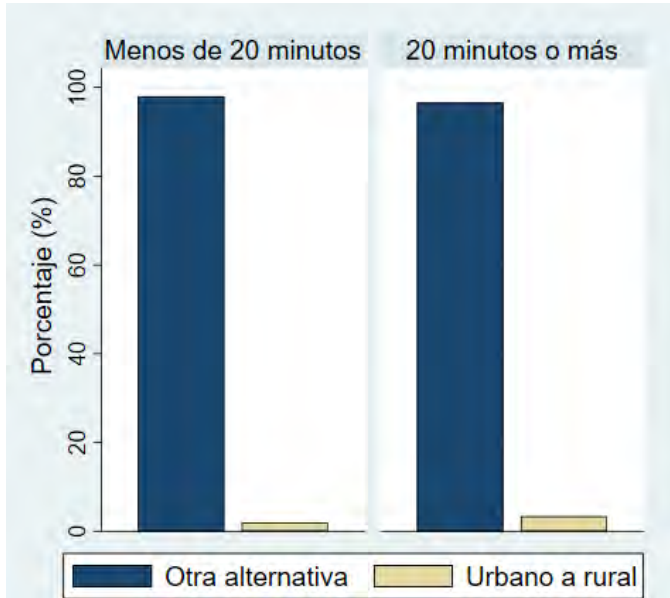


Gráfico 25

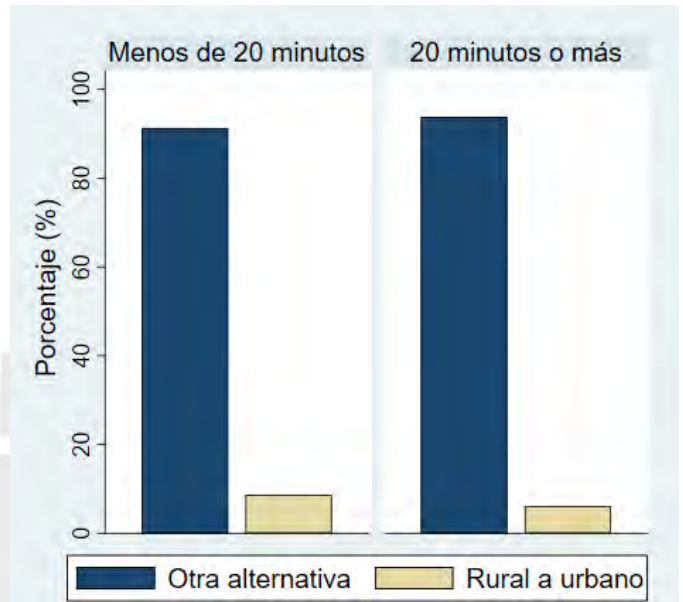
Distribución de traslado de zona de residencia urbana a rural entre la ronda 1 y 5 según el tiempo de desplazamiento (alumnos de secundaria)



Fuente: Boyden (2022c). Elaboración propia.

Gráfico 26

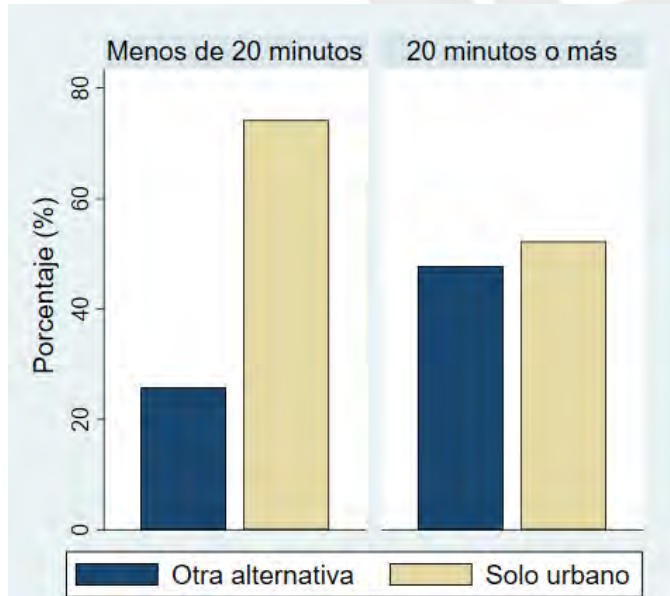
Distribución de traslado de zona de residencia urbana a rural entre la ronda 1 y 5 según el tiempo de desplazamiento (alumnos de secundaria)



Fuente: Boyden (2022c). Elaboración propia.

Gráfico 27

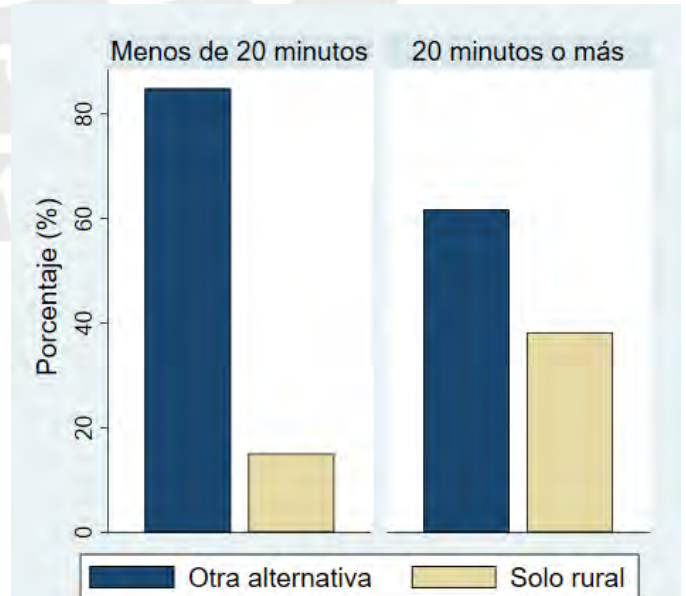
Distribución de la permanencia de la residencia en una zona urbana entre la ronda 1 y 5 según el tiempo de desplazamiento (alumnos de secundaria)



Fuente: Boyden (2022c). Elaboración propia.

Gráfico 28

Distribución de la permanencia de la residencia en una zona rural entre la ronda 1 y 5 según el tiempo de desplazamiento (alumnos de secundaria)

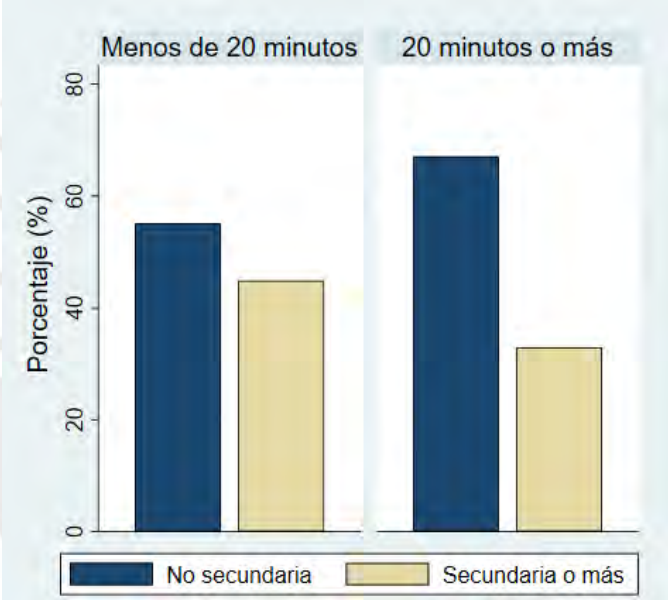


Fuente: Boyden (2022c). Elaboración propia.

Respecto a la distribución de los grupos de acuerdo a la educación de la madre, esta se puede apreciar en el gráfico 29, en el que los alumnos cuya madre posee una educación igual o mayor a una secundaria constituyen el 45% del grupo que realiza viajes cortos y el 33% para los alumnos cuyos viajes son largos. Además, en el gráfico 30 se denota cómo los grupos poseen una distribución similar del número de familiares convivientes en la residencia, siendo el promedio 5.17 y 5.39 familiares para los alumnos con viajes cortos y largos de manera correspondiente.

Gráfico 29

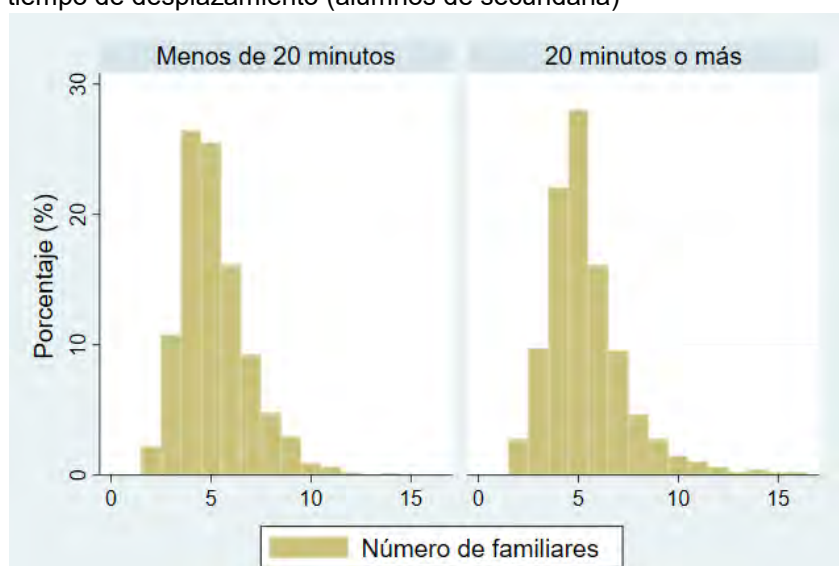
Distribución de la educación de la madre a partir de si posee una educación secundaria o mayor según el tiempo de desplazamiento (alumnos de secundaria)



Fuente: Boyden (2022c). Elaboración propia.

Gráfico 30

Distribución del número de familiares residentes en el hogar según el tiempo de desplazamiento (alumnos de secundaria)

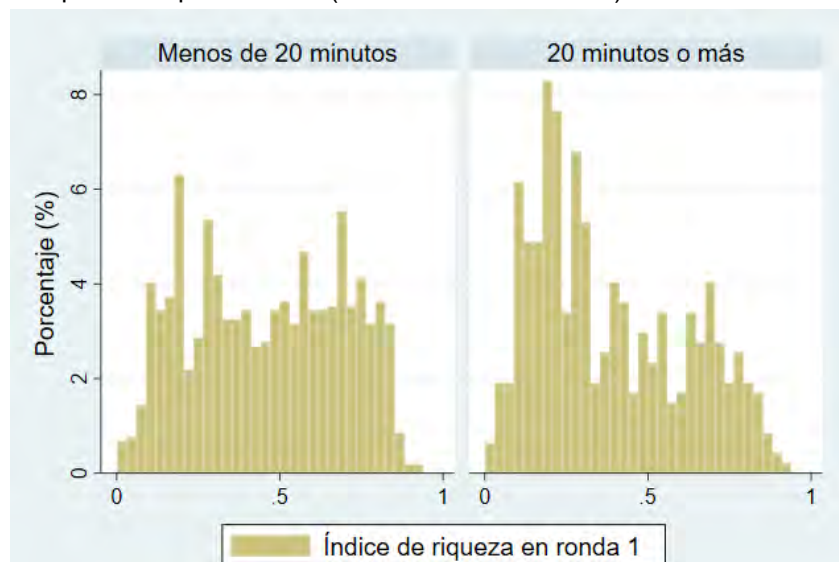


Fuente: Boyden (2022c). Elaboración propia.

Finalmente, los gráficos 31, 32 y 33 detallan cómo se distribuyen los índices de riqueza en las rondas 1, 4 y 5 respectivamente. En los 3 casos, el índice de riqueza promedio de los alumnos con viajes cortos es mayor al de los que realizan viajes largos, siendo la diferencia entre estos aproximadamente igual a 0.077 puntos para las rondas 1 y 4 e igual a 0.075 para la ronda 5. Adicionalmente, así como en el caso de los alumnos de primaria, los alumnos con viajes largos poseen una distribución concentrada en los valores más bajos y altos del índice de riqueza en las 3 rondas mencionadas, con excepción en la ronda 1 donde se concentra particularmente en los valores más bajos. No obstante, los alumnos de viajes cortos solo poseen una distribución equitativa en la ronda 1, mientras que en la 4 y 5 la distribución del índice sigue un patrón ascendente de acuerdo al aumento del valor del índice, con excepción en particulares valores alrededor de 0.5 y los valores mayores extremos cercanos a 1.

Gráfico 31

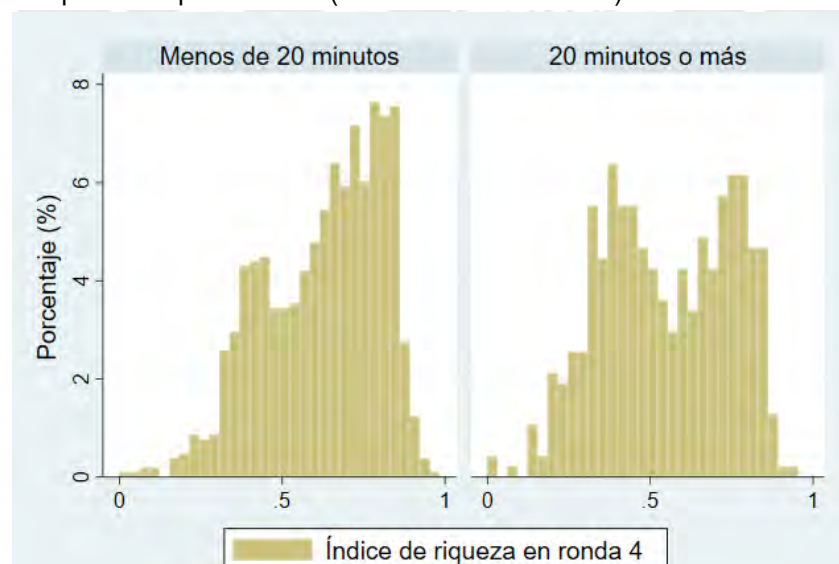
Distribución del índice de riqueza de la familia en la ronda 1 según el tiempo de desplazamiento (alumnos de secundaria)



Fuente: Boyden (2022c). Elaboración propia.

Gráfico 32

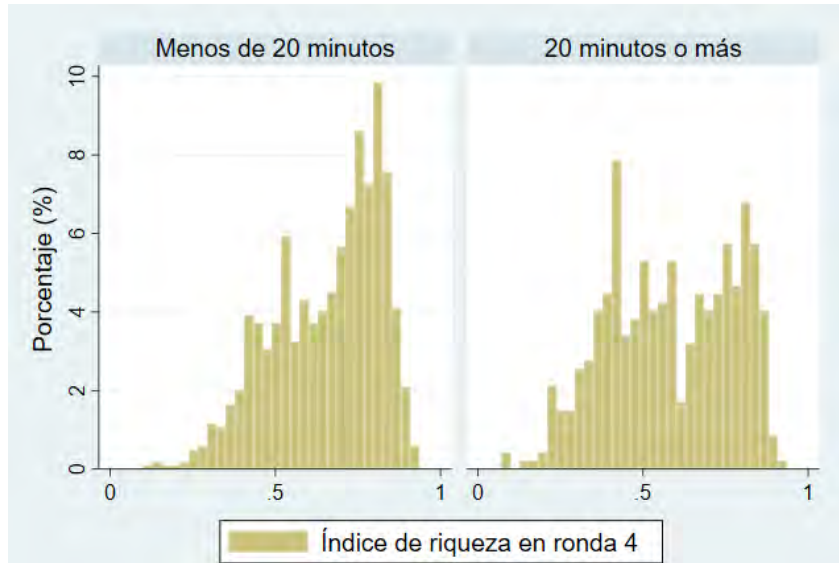
Distribución del índice de riqueza de la familia en la ronda 4 según el tiempo de desplazamiento (alumnos de secundaria)



Fuente: Boyden (2022c). Elaboración propia.

Gráfico 33

Distribución del índice de riqueza de la familia en la ronda 5 según el tiempo de desplazamiento (alumnos de secundaria)



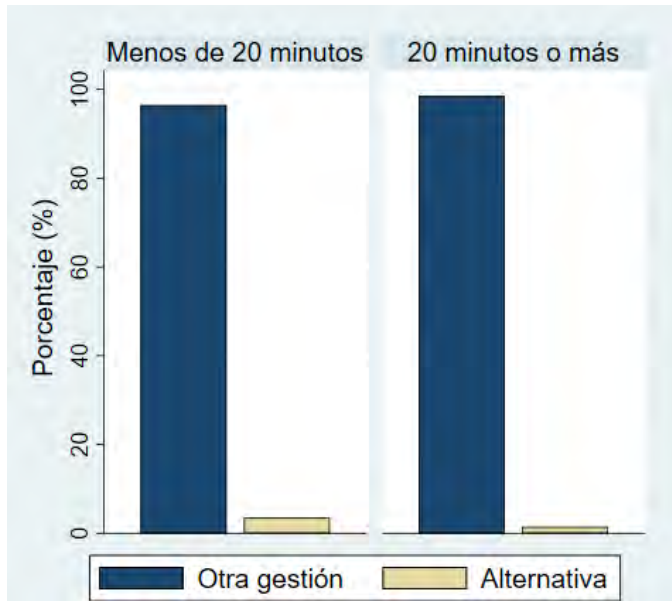
Fuente: Boyden (2022c). Elaboración propia.

#### 4.2.3 Covariables de la oferta educativa y habilidades innatas del alumno

En lo que concierne a las covariables de la oferta educativa, los gráficos 34, 35 y 36 indican cómo se distribuyen la gestión de las instituciones educativas en las que se encuentran los alumnos según los grupos de acuerdo al tiempo de viaje. Así, el 4% de los alumnos de viajes cortos se encuentran matriculados en instituciones educativas con gestión alterna, mientras que esta cifra es el 1% para los que realizan viajes largos. Para el caso de alumnos matriculados en instituciones de gestión privada, estos conforman el 15% y 13% de quienes realizan viajes cortos y largos respectivamente. De esta manera, el 81% de los alumnos que realizan viajes cortos están en una institución educativa pública, siendo esta cifra el 86% para los que realizan viajes largos.

Gráfico 34

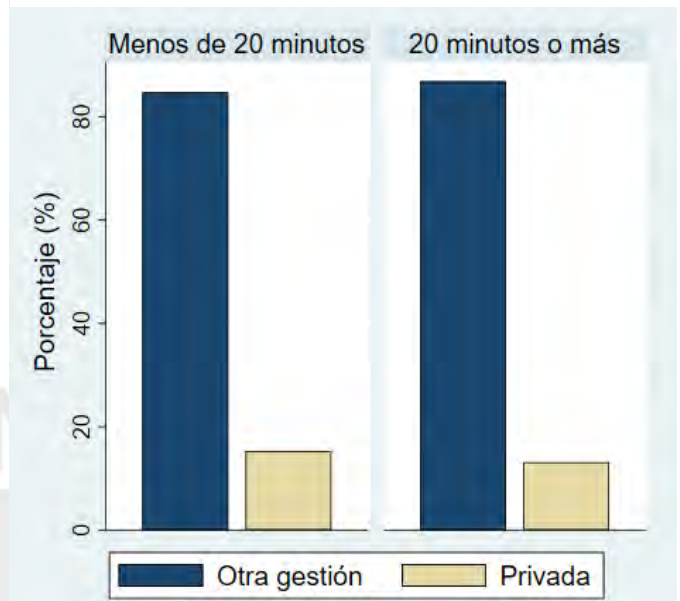
Distribución de la institución educativa matriculada a partir de la gestión alternativa según el tiempo de desplazamiento (alumnos de secundaria)



Fuente: Boyden (2022c). Elaboración propia.

Gráfico 35

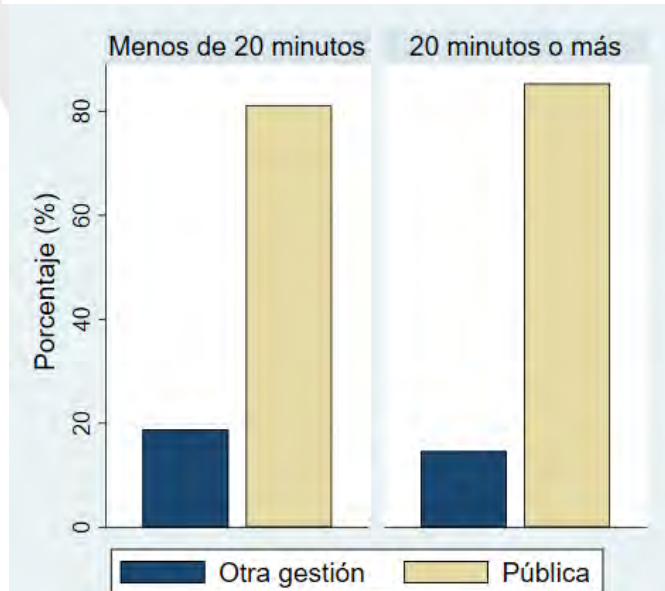
Distribución de la institución educativa matriculada a partir de la gestión privada según el tiempo de desplazamiento (alumnos de secundaria)



Fuente: Boyden (2022c). Elaboración propia.

Gráfico 36

Distribución de la institución educativa matriculada a partir de la gestión pública según el tiempo de desplazamiento (alumnos de secundaria)

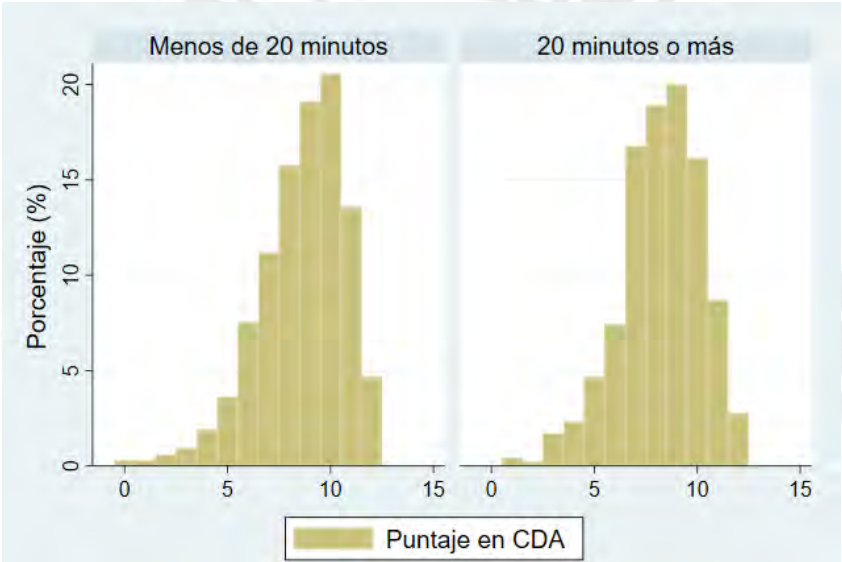


Fuente: Boyden (2022c). Elaboración propia.

Para finalizar, en el gráfico 37 se puede observar la distribución del puntaje obtenido en el examen CDA rendido en la ronda 2. Esta covariable, que representa las habilidades innatas del alumno, es similar para ambos grupos, siendo el promedio del puntaje 8.62 para los alumnos que realizan viajes cortos y 8.22 para quienes realizan viajes largos. Asimismo, los alumnos que realizan viajes largos poseen una menor proporción de observaciones mayores a 10 puntos comparados a los que hacen viajes cortos desde su residencia hacia la institución educativa.

Gráfico 37

Distribución del puntaje obtenido en el examen CDA en la ronda 2 según el tiempo de desplazamiento (alumnos de secundaria)



Fuente: Boyden (2022a). Elaboración propia.

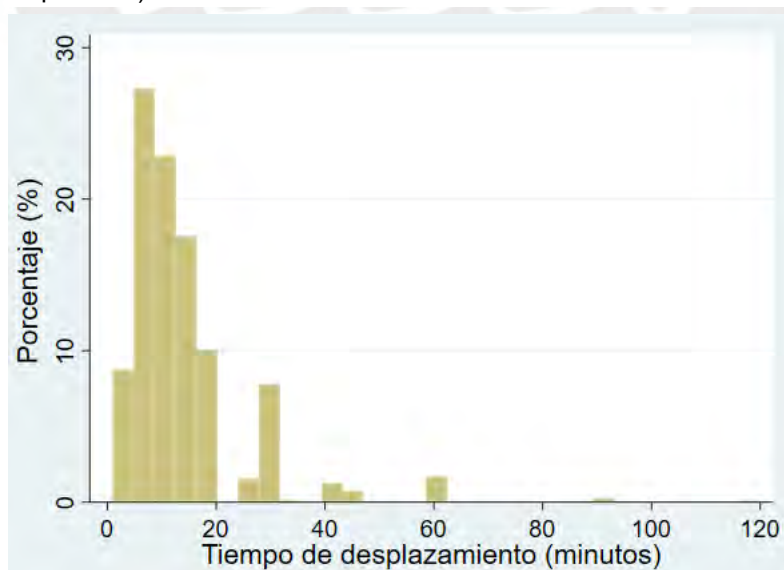


#### 4.3 Análisis previo del tiempo de desplazamiento y rendimiento académico

Antes de realizar este análisis previo, primero se presenta en los gráficos 38 y 39 las distribuciones del tiempo de desplazamiento de los alumnos de primaria en la ronda 3 y secundaria en la ronda 5 respectivamente. Como se puede apreciar, la distribución de este tiempo para los alumnos de primaria se concentra en duraciones menores o iguales a 20 minutos, con algunos valores mayores a esta, siendo su promedio de 13.54 minutos. Mientras que para el grupo de alumnos de secundaria el promedio del tiempo de desplazamiento es mayor, siendo su valor de 17.54 minutos. Si bien la distribución de estos alumnos es parecida al de los estudiantes de primaria al concentrarse en valores menores a 20 minutos, esta posee una mayor cantidad de observaciones atípicas mayores que este intervalo. Esto corrobora lo mencionado respecto a la oferta educativa, pues como hay una escasa cantidad de instituciones educativas secundarias comparadas a las primarias, los alumnos de secundaria tienden a realizar recorridos más largos que los de primaria hacia sus respectivos colegios (ESCALE, 2023). Por tanto, de acuerdo al marco teórico presentado, los estudiantes de secundaria deberían ver afectado en mayor medida su rendimiento académico.

Gráfico 38

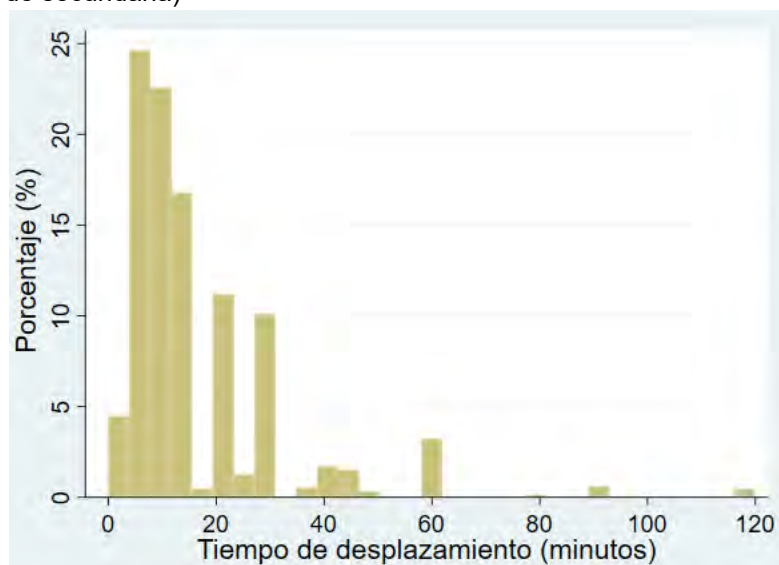
Distribución del tiempo de desplazamiento en la ronda 3 (alumnos de primaria)



Fuente: Boyden (2022c). Elaboración propia.

Gráfico 39

Distribución del tiempo de desplazamiento en la ronda 5 (alumnos de secundaria)



Fuente: Boyden (2022c). Elaboración propia.

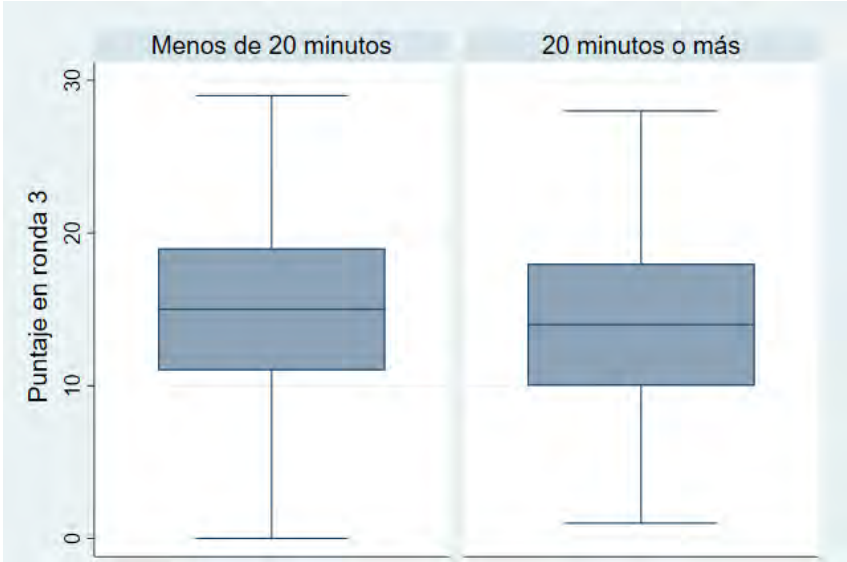
En este contexto, en los gráficos 40 y 41 se puede apreciar el rendimiento académico en el examen de matemáticas de los alumnos de primaria y secundaria en las rondas 3 y 5 respectivamente de acuerdo al punto de corte de 20 minutos. Además, para este análisis preliminar se ha utilizado el puntaje obtenido sin estandarizar en las evaluaciones de matemáticas rendidas. Como se mencionaba en el marco teórico, el puntaje máximo obtenible en los exámenes de las rondas 3 y 5 es de 29 y 31 respectivamente, siendo el mínimo para ambos 0. Así, la mediana y promedio de la nota de los alumnos de primaria en la ronda 3 con viajes cortos es de 15 y 14.94, mientras que para los de viajes largos es de 14 y 13.93 puntos respectivamente. Asimismo, el rango intercuartil de ambos grupos es bastante proporcional, siendo más notoria la mayor dispersión de las observaciones en el grupo de alumnos de viajes cortos. En general, la distribución de las notas en los alumnos de viajes largos ligeramente se inclina a un menor rendimiento académico comparado al otro grupo.

En contraste, la distribución del rendimiento académico entre los grupos de alumnos de secundaria en la ronda 5 es más desigual. Los que realizan viajes cortos poseen una mediana y promedio de puntaje en el examen de matemáticas de 12 y 12.07 puntos, y los alumnos con viajes largos poseen 10 y 10.89 respectivamente. En esta ronda, se puede observar que los alumnos con viajes largos poseen una mayor

concentración de observaciones antes de la mediana, mientras que en el caso de los que realizan viajes cortos la distribución es más equilibrada.

Gráfico 40

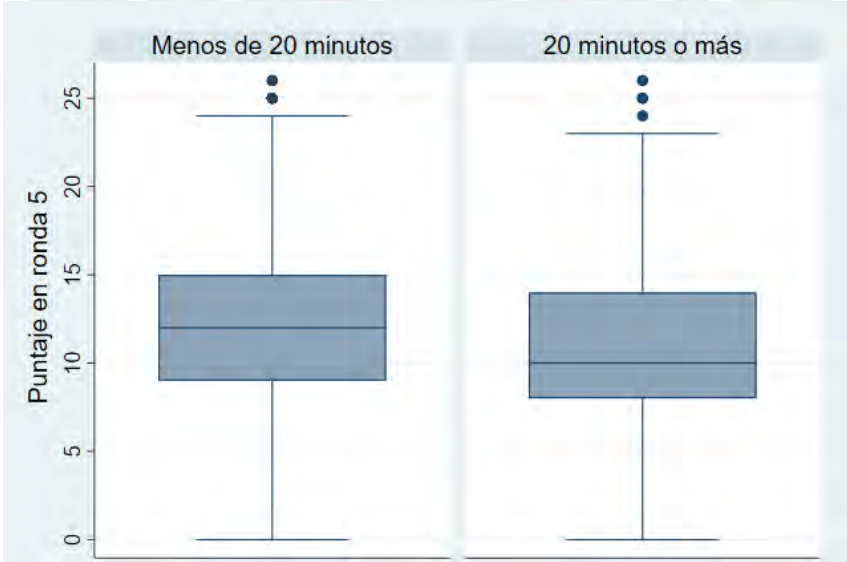
Puntaje obtenido en la prueba de matemática en la ronda 3 según el tiempo de desplazamiento (alumnos de primaria)



Fuente: Fuente: Boyden (2022b, 2022c). Elaboración propia.

Gráfico 41

Puntaje obtenido en la prueba de matemática en la ronda 5 según el tiempo de desplazamiento (alumnos de secundaria)



Fuente: Boyden (2022c) y Sanchez et al. (2022). Elaboración propia.

## 5. Resultados de las estimaciones principales

En esta sección se brindan los resultados de la estimación del efecto del tiempo de desplazamiento de los alumnos de primaria y secundaria en su rendimiento académico en la ronda 3 y 5 respectivamente. Como se mencionaba previamente, el cálculo se realiza mediante la técnica de Propensity score matching para aislar el efecto del tiempo de desplazamiento y poder evaluar bajo un determinado punto de corte de tiempo, aislando variables confundidoras. Cabe reiterar que la variable que representa el tiempo de desplazamiento es el tiempo de viaje de ida desde la residencia hasta la institución educativa que asiste el alumno, mientras que el rendimiento académico se captura con el puntaje obtenido estandarizado de los alumnos en la evaluación de matemática en sus respectivas rondas. Estos resultados se han realizado bajo el emparejamiento de 5 vecinos más cercanos y un caliper de 0.005 desviaciones estándar del PS para un ajuste estricto en el proceso. Adicionalmente, para calcular el PS de cada corte de tiempo en las rondas 3 y 5 se han realizado regresiones logit, las cuales se encuentran en el anexo A.

### 5.1 Prueba de matemáticas de los alumnos de primaria (ronda 3)

#### 5.1.1 Resultados con el punto de corte de 20 minutos

La tabla 4 indica la calidad del emparejamiento mostrando la diferencia en los valores promedio de las covariables en el grupo tratado de alumnos con viajes largos y el grupo control de los alumnos con viajes cortos cuando cursaban la educación primaria en la ronda 3 con el punto de corte de 20 minutos. El porcentaje de sesgo entre el promedio de las covariables de los grupos emparejados es de 3%, mientras que los t-tests de medias demuestran que las diferencias entre los promedios de las covariables en los grupos de alumnos no son significativas. Asimismo, en el gráfico 42, se puede evidenciar la reducción del sesgo al construir el grupo de control, pues el porcentaje estandarizado de este al comparar los valores promedio de las covariables entre el grupo tratado de alumnos con viajes largos y el no tratado con viajes cortos previo al emparejamiento es de 28.4%.

Tabla 4

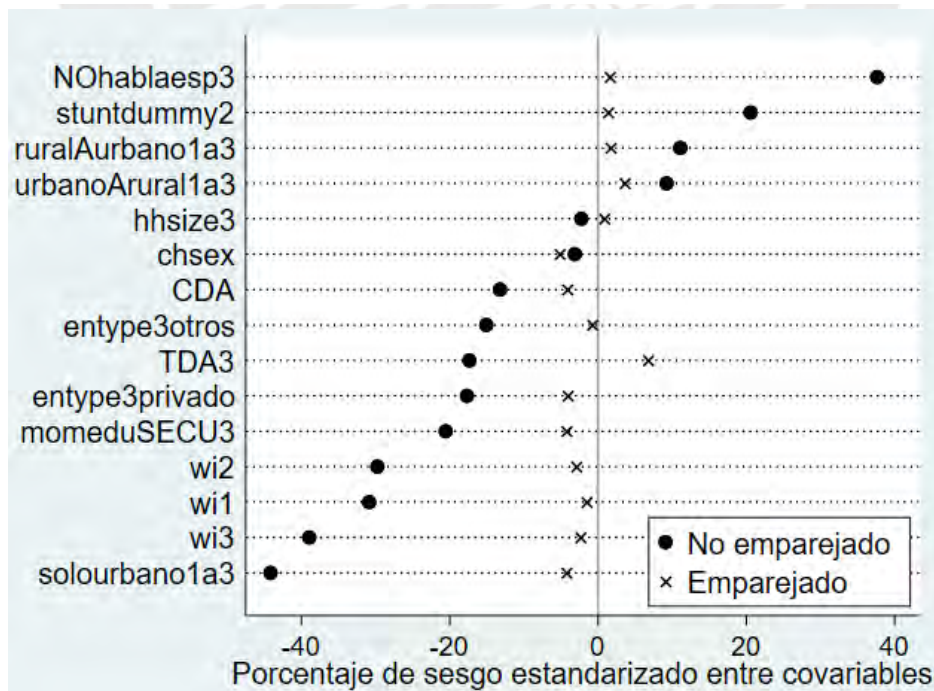
Test de medias de covariables entre alumnos con viajes largos y cortos (control) en la ronda 3 con el punto de corte de 20 minutos (alumnos de primaria)

Variable	Promedio			t-test	
	Viajes Largos	Viajes Cortos (Control)	%sesgo	t	p>t
chsex	0.487	0.513	-5.200	-0.680	0.497
CDA	8.306	8.393	-4.100	-0.550	0.583
NOhablaesp3	0.213	0.207	1.600	0.190	0.852
ruralAurbano1a3	0.075	0.071	1.700	0.210	0.834
urbanoArural1a3	0.032	0.026	3.600	0.450	0.651
solourbano1a3	0.530	0.550	-4.300	-0.530	0.595
entype3otros	0.012	0.013	-0.800	-0.140	0.890
entype3privado	0.095	0.108	-4.100	-0.580	0.564
stuntdummy2	0.383	0.377	1.400	0.170	0.862
hhsiz3	5.398	5.382	0.900	0.110	0.910
momeduSECU3	0.308	0.329	-4.200	-0.570	0.569
wi1	0.388	0.392	-1.500	-0.200	0.839
wi2	0.430	0.437	-2.900	-0.380	0.700
wi3	0.500	0.505	-2.300	-0.300	0.761
TDA3	0.697	0.667	6.800	0.850	0.397

Fuente: Boyden (2022a, 2022b, 2022c) y Sanchez et al. (2022). Elaboración propia.

Gráfico 42

Porcentaje de sesgo estandarizado entre las covariables de la ronda 3 con el punto de corte de 20 minutos de los alumnos de viajes cortos al ser emparejados y no emparejados (alumnos de primaria)



Fuente: Boyden (2022a, 2022b, 2022c) y Sanchez et al. (2022).  
Elaboración propia.

Asimismo, en la tabla 5 se puede apreciar que sí existe un soporte común de 1514 observaciones de los 1524 totales, por lo que hay 10 observaciones fuera del soporte común, los cuales pertenecen al grupo de alumnos con viajes largos. El gráfico 43 señala que los alumnos con viajes largos con un PS bajo tienen más alumnos del grupo de viajes cortos con los cuales pueden ser emparejados, mientras quienes poseen un puntaje más alto tienen menos posibilidades de ser emparejados e incluso algunos que se encuentran fuera del soporte común.

Tabla 5

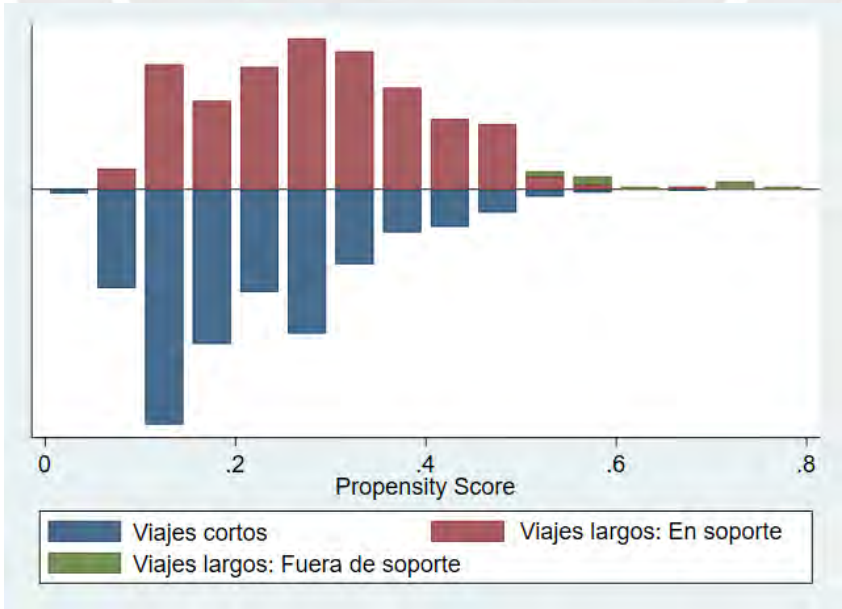
Soporte común en el emparejamiento en la ronda 3 con el punto de corte de 20 minutos (alumnos de primaria)

Asignación	Soporte Común		Total
	Fuera	Adentro	
Viajes cortos (<20 minutos)	0	1167	1167
Viajes largos (>=20 minutos)	10	347	357
Total	10	1514	1524

Fuente: Boyden (2022a, 2022b, 2022c) y Sanchez et al. (2022). Elaboración propia.

Gráfico 43

Soporte común de la ronda 3 con el punto de corte de 20 minutos según el Propensity Score (alumnos de primaria)



Fuente: Boyden (2022a, 2022b, 2022c) y Sanchez et al. (2022). Elaboración propia.

Finalmente, en la tabla 6, se puede observar el efecto aislado del tiempo de desplazamiento en el rendimiento académico en la prueba de matemáticas de los alumnos de primaria mediante el cálculo del efecto del tratamiento en los tratados (ATT) con el emparejamiento seleccionado y el punto de corte de 20 minutos. Asimismo, la variable “std\_math3” representa el puntaje obtenido estandarizado en el examen de matemáticas en la ronda 3. Este efecto, que representa la diferencia en la calificación en el grupo de alumnos de viajes largos con el del grupo control que realiza viajes cortos, es de -0.014 desviaciones estándar. Por tanto, los alumnos que realizan viajes largos mayores o iguales a 20 minutos consiguen en promedio un puntaje menor en 0.014 desviaciones estándar. Sin embargo, este efecto no es significativo, dado la prueba T de Student (T-stat) de -0.190.

Tabla 6

Estimación del ATT en la ronda 3 con el punto de corte de 20 minutos, emparejamiento de 5 vecinos más cercanos y caliper de 0.005 (alumnos de primaria)

Variable	Muestra	Viajes Largos	Viajes Cortos (Control)	Diferencia	Error std.	T-stat
std_math3	No emparejado	-0.142	0.041	-0.182	0.060	-3.020
	ATT	-0.124	-0.110	-0.014	0.071	-0.190

Fuente: Boyden (2022a, 2022b, 2022c) y Sanchez et al. (2022). Elaboración propia.

### 5.1.2 Resultados con los puntos de corte de 25 y 30 minutos

Para comprender de manera más profunda el efecto del tiempo de desplazamiento en el rendimiento académico de los alumnos de primaria a nivel nacional, se han realizado estimaciones PSM con las mismas covariables y especificaciones del emparejamiento, pero con puntos de corte de tiempo mayores, particularmente de 25 y 30 minutos, donde los alumnos de viajes largos son los que se demoran igual o más que estos intervalos, mientras que los de viajes cortos representarían la parte complementaria. En ambos casos, todos los promedios de las covariables entre los alumnos de viajes largos y el grupo control de los que realizan viajes cortos no son significativamente diferentes, así como se confirma la existencia en los dos de un soporte común con 7 observaciones en el grupo de alumnos con viajes largos fuera de este. Igual que en el caso con el corte de 20 minutos, los alumnos de viajes largos fuera del soporte común son los que poseen los PS más

altos. En el anexo B se puede apreciar más a detalle los resultados de estos procedimientos con estos puntos de corte para los alumnos de primaria en la ronda 3.

Así, en las tablas 7 y 8 se puede observar el ATT con los puntos de corte de 25 y 30 minutos. En el primer caso, la diferencia es positiva y equivalente a 0.125 desviaciones estándar, mientras que para el segundo el efecto también es positivo, pero igual a 0.188 desviaciones estándar del puntaje en el examen de matemáticas. Estos resultados indican que, para los puntos de corte de 25 y 30 minutos, los alumnos con viajes largos poseen en promedio un puntaje mayor a los que realizan viajes cortos si se aísla el efecto del tiempo de desplazamiento en el rendimiento académico de los alumnos de primaria en la ronda 3. No obstante, este impacto positivo del tiempo de viaje en el rendimiento es solo significativo al 5% con el punto de corte de 30 minutos al poseer un T-stat de 2.010.

Tabla 7

Estimación del ATT en la ronda 3 con el punto de corte de 25 minutos, emparejamiento de 5 vecinos más cercanos y caliper de 0.005 (alumnos de primaria)

Variable	Muestra	Viajes Largos	Viajes Cortos (Control)	Diferencia	Error std.	T-stat
std_math3	No emparejado	-0.144	0.019	-0.164	0.075	-2.190
	ATT	-0.129	-0.254	0.125	0.092	1.360

Fuente: Boyden (2022a, 2022b, 2022c) y Sanchez et al. (2022). Elaboración propia.

Tabla 8

Estimación del ATT en la ronda 3 con el punto de corte de 30 minutos, emparejamiento de 5 vecinos más cercanos y caliper de 0.005 (alumnos de primaria)

Variable	Muestra	Viajes Largos	Viajes Cortos (Control)	Diferencia	Error std.	T-stat
std_math3	No emparejado	-0.166	0.020	-0.186	0.079	-2.350
	ATT	-0.149	-0.337	0.188	0.093	2.010

Fuente: Boyden (2022a, 2022b, 2022c) y Sanchez et al. (2022). Elaboración propia.

## 5.2 Prueba de matemáticas de los alumnos de secundaria (ronda 5)

### 5.2.1 Resultados con el punto de corte de 20 minutos

De la misma manera, para el caso de los alumnos de secundaria en el presente estudio, la tabla 9 indica el balanceo del emparejamiento realizado al mostrar la diferencia en los valores promedio de las covariables entre los alumnos con viajes largos y el grupo control de los alumnos con viajes cortos en la ronda 5 con el punto de corte de 20 minutos. Los t-tests de medias señalan que ningún promedio de las covariables entre los grupos de alumnos es significativamente distinto, siendo el porcentaje medio del sesgo entre las covariables de 2.7%. Añadiendo al análisis, el gráfico 44 señala cómo el emparejamiento ha causado que el grupo control de alumnos con viajes cortos sea comparable al grupo de alumnos con viajes largos, ya que el porcentaje del sesgo estandarizado al comparar los valores promedio de las covariables entre los grupos sin emparejar es de 22.8%.

Tabla 9

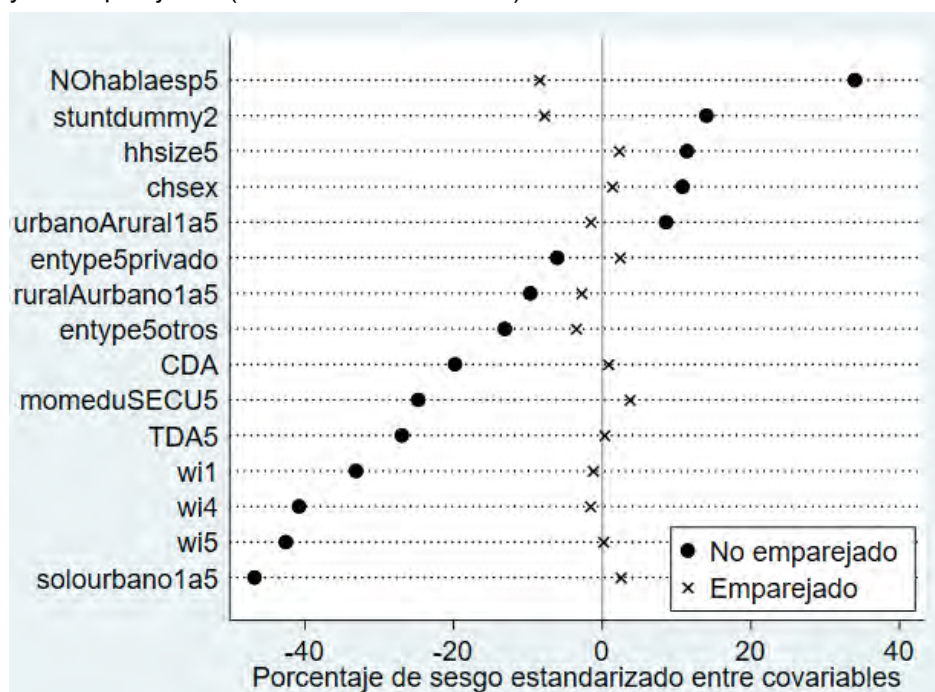
Test de medias de covariables entre alumnos con viajes largos y cortos (control) en la ronda 5 con el punto de corte de 20 minutos (alumnos de secundaria)

Variable	Promedio			t-test	
	Viajes Largos	Viajes Cortos (Control)	%sesgo	t	p>t
chsex	0.523	0.516	1.300	0.200	0.843
CDA	8.251	8.234	0.800	0.120	0.908
NOhablaesp5	0.167	0.196	-8.500	-1.140	0.255
ruralAurbano1a5	0.066	0.074	-2.800	-0.420	0.672
urbanoArural1a5	0.037	0.039	-1.600	-0.200	0.843
solourbano1a5	0.562	0.550	2.500	0.350	0.727
entype5otros	0.016	0.021	-3.500	-0.600	0.550
entype5privado	0.139	0.131	2.400	0.360	0.721
stuntdummy2	0.342	0.379	-7.800	-1.120	0.261
hhsize5	5.338	5.295	2.300	0.330	0.741
momeduSECU5	0.352	0.334	3.700	0.560	0.576
wi1	0.394	0.397	-1.200	-0.180	0.857
wi4	0.563	0.566	-1.600	-0.240	0.814
wi5	0.600	0.600	0.100	0.020	0.983
TDA5	0.543	0.542	0.300	0.040	0.965

Fuente: Boyden (2022a, 2022b, 2022c) y Sanchez et al. (2022). Elaboración propia.

Gráfico 44

Porcentaje de sesgo estandarizado entre las covariables de la ronda 5 con el punto de corte de 20 minutos de los alumnos de viajes cortos al ser emparejados y no emparejados (alumnos de secundaria)



Fuente: Boyden (2022a, 2022b, 2022c) y Sanchez et al. (2022).  
Elaboración propia.

Respecto al soporte común, la tabla 10 muestra como este emparejamiento ha resultado en que 33 alumnos del grupo que realizan viajes largos se encuentren fuera de este, por lo que 438 fueron los que se consideraron en este procedimiento de PSM, conformando un total de 1485 alumnos dentro del soporte común. Adicionalmente, el gráfico 45 muestra que, igualmente en el caso de los alumnos de primaria en la ronda 3, la población de alumnos de secundaria con viajes largos fuera del soporte común en la ronda 5 posee los PS más altos, mientras los que tienen un PS bajo tienen una amplia cantidad de alumnos con viajes cortos con los cuales pueden ser emparejados.

Tabla 10

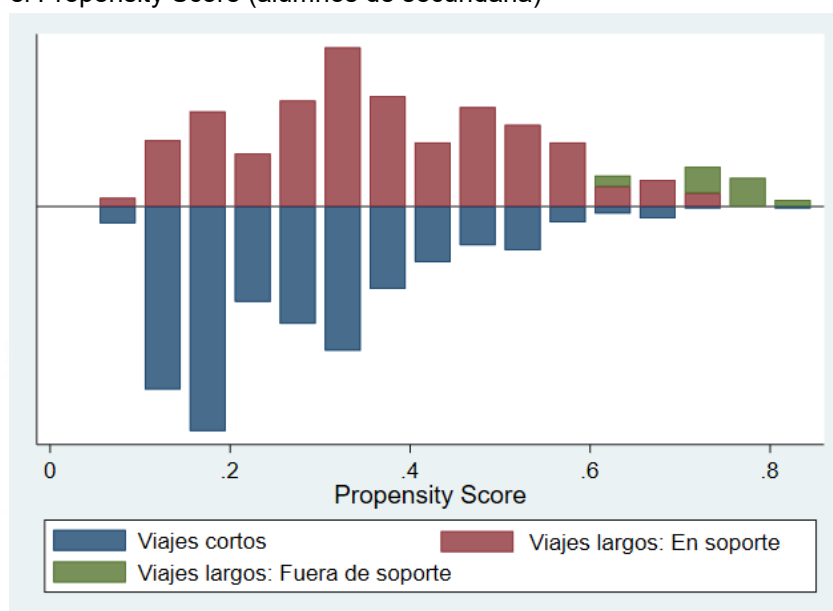
Soporte común en el emparejamiento en la ronda 5 con el punto de corte de 20 minutos (alumnos de secundaria)

Asignación	Soporte Común		Total
	Fuera	Adentro	
Viajes cortos (<20 minutos)	0	1047	1047
Viajes largos (>=20 minutos)	33	438	471
Total	33	1485	1518

Fuente: Boyden (2022a, 2022b, 2022c) y Sanchez et al. (2022). Elaboración propia.

Gráfico 45

Soporte común de la ronda 5 con el punto de corte de 20 minutos según el Propensity Score (alumnos de secundaria)



Fuente: Boyden (2022a, 2022b, 2022c) y Sanchez et al. (2022).  
Elaboración propia.

Finalmente, la tabla 11 muestra el ATT de los alumnos de secundaria en la ronda 5 con el punto de corte de 20 minutos. La variable “std\_math5” representa el puntaje obtenido estandarizado en el examen de matemáticas en esta ronda mencionada. De esta manera, el resultado de este efecto indica que el grupo de alumnos que realizan viajes largos posee un puntaje promedio menor en la prueba de matemáticas que quienes realizan viajes cortos en alrededor de 0.021 desviaciones estándar. No obstante, este impacto negativo en el rendimiento académico no es significativo al tener un T-stat de -0.32.

Tabla 11

Estimación del ATT en la ronda 5 con el punto de corte de 20 minutos, emparejamiento de 5 vecinos más cercanos y caliper de 0.005 (alumnos de secundaria)

Variable	Muestra	Viajes Largos	Viajes Cortos (Control)	Diferencia	Error std.	T-stat
std_math5	No emparejado	-0.177	0.074	-0.251	0.055	-4.560
	ATT	-0.145	-0.123	-0.021	0.068	-0.320

Fuente: Boyden (2022a, 2022b, 2022c) y Sanchez et al. (2022). Elaboración propia.

### 5.2.2 Resultados con los puntos de corte de 25 y 30 minutos

Así como en el caso de los alumnos cuando se encontraban en primaria, se utilizan adicionalmente los puntos de corte de 25 y 30 minutos para analizar el efecto de tiempos de desplazamiento más prolongados en el rendimiento académico de los alumnos en la ronda 5 cuando se encontraban en secundaria. En ambos cortes, como se evidencian en los gráficos y tablas del anexo C, todos los promedios de las covariables entre los grupos mencionados no son estadísticamente diferentes, evidenciando la calidad de los emparejamientos y la reducción del sesgo. Además, se confirma la presencia de un soporte común, perdiendo menos de 20 observaciones en el grupo de alumnos con viajes largos, siendo estos los que poseen los PS más altos.

De esta manera, las tablas 12 y 13 denotan el ATT para los puntos de corte de 25 y 30 minutos. Para el primer corte, el efecto es negativo e igual a -0.078 desviaciones estándar, mientras que para el corte de 30 minutos el efecto también es negativo e igual a -0.028 desviaciones estándar del puntaje en la prueba de matemáticas. Así, los resultados muestran que en ambos casos los alumnos de secundaria que realizan viajes largos poseen en promedio un puntaje ligeramente menor que quienes realizan viajes cortos, aunque con ambos puntos de corte este efecto no es significativo.

Tabla 12

Estimación del ATT en la ronda 5 con el punto de corte de 25 minutos, emparejamiento de 5 vecinos más cercanos y caliper de 0.005 (alumnos de secundaria)

Variable	Muestra	Viajes Largos	Viajes Cortos (Control)	Diferencia	Error std.	T-stat
std_math5	No emparejado	-0.237	0.054	-0.291	0.064	-4.550
	ATT	-0.241	-0.163	-0.078	0.076	-1.03

Fuente: Boyden (2022a, 2022b, 2022c) y Sanchez et al. (2022). Elaboración propia.

Tabla 13

Estimación del ATT en la ronda 5 con el punto de corte de 30 minutos, emparejamiento de 5 vecinos más cercanos y caliper de 0.005 (alumnos de secundaria)

Variable	Muestra	Viajes Largos	Viajes Cortos (Control)	Diferencia	Error std.	T-stat
std_math5	No emparejado	-0.262	0.055	-0.317	0.065	-4.850
	ATT	-0.254	-0.226	-0.028	0.083	-0.340

Fuente: Boyden (2022a, 2022b, 2022c) y Sanchez et al. (2022). Elaboración propia.

## 6. Interpretación de los resultados principales

En esta sección, se va a brindar las posibles explicaciones de los resultados hallados con la población total utilizada. Por tanto, se presentan las tablas 14 y 15 que muestran la dinámica del efecto del tiempo de desplazamiento en el rendimiento académico de los alumnos de primaria y secundaria a nivel nacional respectivamente de acuerdo a los puntos de corte de tiempo de 20, 25 y 30 minutos. Cabe reiterar que la medida del efecto se expresa con la desviación estándar del puntaje en el examen de matemáticas en la ronda rendida.

Tabla 14

Efecto del tiempo de desplazamiento en los alumnos de nivel primario de viajes largos (ATT) con los cortes de 20, 25 y 30 minutos en la ronda 3 (5 vecinos más cercanos y caliper 0.005)

Corte de tiempo	Viajes Largos	Viajes Cortos (Control)	Diferencia (ATT)	Error std.	T-stat
>= 20 minutos	-0.124	-0.110	-0.014	0.071	-0.190
>= 25 minutos	-0.129	-0.254	0.125	0.092	1.36
>= 30 minutos	-0.149	-0.337	0.188**	0.093	2.01

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

Fuente: Boyden (2022a, 2022b, 2022c) y Sanchez et al. (2022). Elaboración propia.

Tabla 15

Efecto del tiempo de desplazamiento en los alumnos de nivel secundario de viajes largos (ATT) con los cortes de 20, 25 y 30 minutos en la ronda 5 (5 vecinos más cercanos y caliper 0.005)

Corte de tiempo	Viajes Largos	Viajes Cortos (Control)	Diferencia (ATT)	Error std.	T-stat
>= 20 minutos	-0.145	-0.123	-0.021	0.068	-0.32
>= 25 minutos	-0.241	-0.163	-0.078	0.076	-1.03
>= 30 minutos	-0.254	-0.226	-0.028	0.083	-0.34

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

Fuente: Boyden (2022a, 2022b, 2022c) y Sanchez et al. (2022). Elaboración propia.

Para los alumnos de primaria, es evidente que mientras más alto sea el corte de tiempo, más grande y significativo es el efecto del tiempo de desplazamiento sobre el rendimiento académico, el cual es positivo. Este efecto consigue significancia al 5% en el último corte de 30 minutos o más. Por tanto, los alumnos de primaria que se demoran 30 minutos o más en su ida desde la residencia hacia su institución educativa poseen en promedio un mejor rendimiento académico. Este efecto se puede comprender por la mayor oferta educativa en este nivel educativo, donde es probable que una gran cantidad de los padres de familia prioriza la calidad educativa de la institución sobre la distancia y el tiempo de desplazamiento hacia esta (Thomas, 2016).

Así, los padres tienen una amplia variedad de instituciones que analizar y pueden elegir el que consideren apropiado de acuerdo a sus preferencias de educación para sus hijos si evalúan que la calidad brindada influye más en el rendimiento académico que factores negativos como el tiempo de viaje. Como comentan He & Giuliano (2017), diversas características de la demanda y oferta educativa como el nivel socioeconómico, la cercanía de la residencia a las instituciones y el tipo de gestión influyen en la decisión de matrícula de los alumnos. Por lo tanto, una oferta educativa abundante de instituciones educativas primarias puede compensar el efecto negativo del tiempo de desplazamiento en el rendimiento académico, pero solo si en estas se brindan una educación de calidad aprobada por los padres de familia de acuerdo a su perfil de características.

Por otro lado, los alumnos de secundaria no presentan un efecto conciso del tiempo de desplazamiento en su rendimiento académico. Si bien es negativo, este es bastante reducido y no significativo. Una razón importante detrás de este resultado podría yacer en la escasa oferta educativa secundaria comparada a la primaria. Los alumnos, especialmente en zonas rurales dado que la oferta educativa es limitada, acudirían a la misma institución educativa, por lo que recibirían la misma calidad educativa en su proceso de aprendizaje (Córdova & Ruiz, 2014; Chaman & Moali, 2020). Así, tomando en cuenta los alumnos que se demoran menos de dos horas en sus viajes de ida a la escuela utilizados en el presente estudio, el efecto de tiempo de desplazamiento no sería suficiente para engendrar una reducción significativa en el rendimiento académico cuando los alumnos reciben un semejante servicio de educación secundaria. Para comprobar la robustez de los resultados hallados de todos los cortes de tiempo en la ronda 3 y 5 para los alumnos de primaria y secundaria, se realizarán adicionales métodos de emparejamiento de PSM, así como un análisis desagregado de los alumnos de acuerdo a su zona de residencia y de la linealidad del efecto del tiempo de desplazamiento en los alumnos con viajes largos.

## 7. Análisis de sensibilidad

En esta sección se utilizan diversos emparejamientos para la técnica de PSM y diferentes divisiones de la población total utilizada previamente de acuerdo a las características principales de los alumnos de primaria y secundaria del estudio Niños del Milenio. De esta manera, se comprobará la consistencia de los resultados hallados con las principales estimaciones realizadas sobre el impacto del tiempo de desplazamiento en el rendimiento académico de los alumnos.

### 7.1 Uso de diferentes métodos de emparejamiento

En primer lugar, se realizó el emparejamiento con la técnica de PSM con regresiones logit para el cálculo del PS, la misma base de datos extraída del estudio Niños del Milenio, y las mismas covariables como en las estimaciones principales, pero con diferentes métodos de emparejamiento y valor de caliper para los alumnos de primaria y secundaria. Así, como se puede ver en las tablas 16, 17, 18 y 19, se realizó la estimación del ATT con emparejamientos de 2 a 5 vecinos más cercanos empleando valores de caliper de 0.005 y 0.01 desviaciones estándar del PS respectivo, así como emparejamientos de kernel (epanechnikov, uniforme y tricubo) para cada nivel educativo. La razón detrás del uso de emparejamiento por kernel se debe a que se desea comprobar la robustez de los resultados hallados dentro de la muestra total utilizada, ya que con este método se empareja la población de alumnos de viajes largos con el promedio ponderado de todas las observaciones dentro del grupo control de alumnos con viajes cortos dentro de una banda de acuerdo a sus respectivos PS. Asimismo, los pesos de las observaciones del grupo control son inversamente proporcionales a la distancia entre el PS del alumno emparejado con viajes largos y el suyo (Vélez et al., 2012; Weldegebriel, 2016).

Tabla 16

Estimación del ATT en la ronda 3 con emparejamientos de vecinos más cercanos, emparejamientos de kernel y caliper de 0.005 (alumnos de primaria)

ATT (caliper 0.005, ronda 3)							
Corte de tiempo	5 vecinos	4 vecinos	3 vecinos	2 vecinos	Kernel (epanechnikov)	Kernel (uniforme)	Kernel (tricubo)
>= 20 minutos	-0.014	-0.005	-0.020	-0.047	0.016	0.012	0.017
>= 25 minutos	0.125	0.124	0.139	0.076	0.114	0.106	0.118
>= 30 minutos	0.188**	0.178*	0.193**	0.225**	0.109	0.105	0.107

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

Fuente: Boyden (2022a, 2022b, 2022c) y Sanchez et al. (2022). Elaboración propia.

Tabla 17

Estimación del ATT en la ronda 3 con emparejamientos de vecinos más cercanos, emparejamientos de kernel y caliper de 0.01 (alumnos de primaria)

ATT (caliper 0.01, ronda 3)							
Corte de tiempo	5 vecinos	4 vecinos	3 vecinos	2 vecinos	Kernel (epanechnikov)	Kernel (uniforme)	Kernel (tricubo)
>= 20 minutos	-0.028	-0.018	-0.035	-0.059	0.016	0.012	0.017
>= 25 minutos	0.117	0.119	0.135	0.069	0.114	0.106	0.118
>= 30 minutos	0.183**	0.175*	0.187*	0.224**	0.109	0.105	0.107

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

Fuente: Boyden (2022a, 2022b, 2022c) y Sanchez et al. (2022). Elaboración propia.

Tabla 18

Estimación del ATT en la ronda 5 con emparejamientos de vecinos más cercanos, emparejamientos de kernel y caliper de 0.005 (alumnos de secundaria)

Corte de tiempo	ATT (caliper 0.005, ronda 5)						
	5 vecinos	4 vecinos	3 vecinos	2 vecinos	Kernel (epanechnikov)	Kernel (uniforme)	Kernel (tricubo)
>= 20 minutos	-0.021	-0.008	-0.006	-0.023	-0.056	-0.074	-0.044
>= 25 minutos	-0.078	-0.089	-0.089	-0.106	-0.038	-0.039	-0.036
>= 30 minutos	-0.028	-0.010	-0.006	-0.015	-0.035	-0.039	-0.032

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

Fuente: Boyden (2022a, 2022b, 2022c) y Sanchez et al. (2022). Elaboración propia.

Tabla 19

Estimación del ATT en la ronda 5 con emparejamientos de vecinos más cercanos, emparejamientos de kernel y caliper de 0.01 (alumnos de secundaria)

Corte de tiempo	ATT (caliper 0.01, ronda 5)						
	5 vecinos	4 vecinos	3 vecinos	2 vecinos	Kernel (epanechnikov)	Kernel (uniforme)	Kernel (tricubo)
>= 20 minutos	-0.017	-0.000	0.000	-0.016	-0.056	-0.074	-0.044
>= 25 minutos	-0.048	-0.060	-0.059	-0.082	-0.038	-0.039	-0.036
>= 30 minutos	-0.051	-0.029	-0.022	-0.036	-0.035	-0.039	-0.032

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

Fuente: Boyden (2022a, 2022b, 2022c) y Sanchez et al. (2022). Elaboración propia.

Las tablas anteriores indican que los resultados con los emparejamientos de vecinos más cercanos siguen la misma tendencia, donde el único punto de corte que es significativo es el de 30 minutos o más para los alumnos de primaria en la ronda 3, ya sea con el caliper de 0.005 o 0.01 desviaciones estándar del PS. Para los resultados con los emparejamientos con el método de kernel, ningún punto de corte para los alumnos de primaria y secundaria es significativo, usualmente generando valores más cercanos a cero comparados a los emparejamientos con vecinos más cercanos<sup>2</sup>. Por lo tanto, al utilizar todos los alumnos con viajes cortos dentro del grupo control para emparejar con los alumnos de viajes largos, los resultados pierden significancia.

## 7.2 Estimación de resultados por zona de residencia y análisis de linealidad

Diversos autores mencionados en la revisión de literatura solo utilizan muestras de alumnos en una determinada zona de residencia, ya sea urbana o rural, para obtener una población más uniforme en sus estimaciones y conseguir resultados más entendibles (Contreras et al., 2018; Córdoba & Ruiz, 2014; Villanueva, 2023). Este enfoque se ve respaldado por las posturas de Córdoba & Ruiz (2014) y Chaman & Moali (2020), quienes argumentan que las características de la oferta educativa del perfil socioeconómico de la demanda educativa inciden en la elección de la institución educativa, al priorizar atributos de estas como el tiempo de viaje a la institución o su calidad educativa, los cuales pueden ser diferentes de acuerdo a la zona de residencia del alumno. Esto se confirmó presentando los datos de la oferta educativa a nivel nacional brindado por ESCALE (2023) previamente, donde aproximadamente dos de cada tres instituciones educativas primarias se encontraban en zonas rurales, mientras que para las secundarias más de la mitad se hallan en áreas urbanas. Por otra parte, diferentes características del perfil de las familias varían de acuerdo si residen en zonas rurales o urbanas, siendo la diferencia principal un mayor nivel socioeconómico de las que residen en las ciudades comparadas a las que viven en áreas rurales al poseer estas últimas una mayor parte de su población en pobreza monetaria (Dirección Nacional de Prospectiva y Estudios Estratégicos, 2023).

---

<sup>2</sup> Si bien no se presentan los resultados de la calidad de todos los emparejamientos, esta es semejante al mostrado con 5 vecinos más cercanos con caliper 0.005 para los alumnos de primaria y secundaria en las tablas 4,9, 45, 47, 49 y 51. De igual manera, los datos se encuentran a disposición de cualquier persona que lo solicite.

Por ello, se realizará la técnica de PSM con una regresión logit para calcular los PS, separando a los alumnos que residen en zonas rurales y urbanas para analizar los resultados y compararlos con las estimaciones principales que usan la muestra total utilizada para los estudiantes de primaria y secundaria. Así, se utilizan los mismos cortes de tiempo y las mismas covariables que se implementaron en las estimaciones con la población total, descartando solo las que caracterizaban a los alumnos de acuerdo a su zona de residencia por la división de la población según este atributo. Para los alumnos de primaria en la ronda 3, 1114 residen en zonas rurales, mientras que 410 se encuentran en zonas urbanas. Adicionalmente, se han descartado las covariables de tipo de gestión de institución educativa para los puntos de corte de 25 y 30 minutos en la población de alumnos rurales debido a que quienes se demoran más que estos intervalos (los alumnos con viajes largos) solo asisten a escuelas públicas. En el caso del corte de 20 minutos se utilizan las covariables de gestión menos la de tipo alterna al descartar la única observación que había de la muestra rural al no existir alumnos en el grupo de control con viajes cortos con quienes compararse.

Mientras tanto, en el caso de los alumnos de la ronda 5 en secundaria, 1143 de ellos residen en zonas urbanas y 375 en áreas rurales sin eliminar ninguna covariable en las dos poblaciones. Debido a la menor cantidad de observaciones por la división de la población según la zona de residencia, se realizará la técnica de bootstrapping con 100 repeticiones para corroborar la validez de los resultados al generar muestras a partir de las ya existentes. Para comprobar la validez de esta técnica, se ha realizado el cálculo del ATT de los alumnos de primaria y secundaria como el estimado con la población total en la sección de interpretación de los resultados de las estimaciones principales, pero ahora utilizando bootstrapping con 100 repeticiones. Como se puede ver en la tabla 20, las estimaciones son idénticas a las mostradas previamente en las tablas 14 y 15.

Tabla 20

Estimación del ATT empleando bootstrapping en la ronda 3 (alumnos de primaria) y 5 (alumnos de secundaria) con emparejamiento de 5 vecinos más cercanos y caliper de 0.005

Corte de tiempo	ATT (5 vecinos, caliper 0.005)	
	Ronda 3	Ronda 5
>= 20 minutos	-0.014	-0.021
>= 25 minutos	0.125	-0.078
>= 30 minutos	0.188**	-0.028

\*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$ . Se utilizó bootstrapping para los errores estándar con 100 réplicas.

Fuente: Boyden (2022a, 2022b, 2022c) y Sanchez et al. (2022).  
Elaboración propia.

De esta manera, en las tablas 21, 22, 23 y 24 se muestra la calidad de los emparejamientos de 5 vecinos más cercanos y kernel para el corte de 20 minutos en los alumnos rurales y urbanos de nivel primario, la cual es adecuada en todos los casos al ser no significativas las diferencias entre los promedios de las covariables de los grupos de alumnos con viajes largos y cortos. Además, los t-tests de medias señalan que los contrastes entre los promedios de las covariables en los grupos de alumnos no son significativos. Para los demás emparejamientos, los respectivos balanceos se mantienen correctos, los cuales se pueden apreciar en el anexo D. Por consiguiente, en las tablas 25 y 26 se pueden apreciar los resultados de las estimaciones del ATT con los cortes de 20, 25 y 30 minutos con la población dividida en rural y urbana para los alumnos de primaria mediante los emparejamientos de 5 vecinos más cercanos utilizando un caliper de 0.005 y kernel (epanechnikov) respectivamente.

Tabla 21

Test de medias de covariables entre alumnos con viajes largos y cortos (control) en la ronda 3 con el punto de corte de 20 minutos, emparejamiento de 5 vecinos más cercanos y caliper de 0.005 (alumnos de primaria rurales)

Variable	Promedio			t-test	
	Viajes Largos	Viajes Cortos (Control)	%sesgo	t	p>t
chsex	0.474	0.443	6.400	0.530	0.598
CDA	7.664	7.647	0.800	0.060	0.949
NOhablaesp3	0.460	0.453	1.400	0.110	0.909
entype3privado	0.007	0.004	3.400	0.320	0.752
stuntdummy2	0.599	0.600	-0.400	-0.030	0.974
hhsiz3	6.022	6.057	-1.900	-0.160	0.870
momeduSECU3	0.058	0.064	-1.900	-0.180	0.860
wi1	0.220	0.234	-11.600	-0.970	0.332
wi2	0.248	0.255	-5.300	-0.460	0.647
wi3	0.333	0.339	-5.100	-0.450	0.652
TDA3	0.964	0.948	6.900	0.630	0.530

Fuente: Boyden (2022a, 2022b, 2022c) y Sanchez et al. (2022). Elaboración propia.

Tabla 22

Test de medias de covariables entre alumnos con viajes largos y cortos (control) en la ronda 3 con el punto de corte de 20 minutos, emparejamiento de 5 vecinos más cercanos y caliper de 0.005 (alumnos de primaria urbanos)

Variable	Promedio			t-test	
	Viajes Largos	Viajes Cortos (Control)	%sesgo	t	p>t
chsex	0.495	0.490	0.900	0.100	0.924
CDA	8.806	8.766	2.100	0.210	0.834
NOhablaesp3	0.044	0.036	4.000	0.390	0.694
entype3otros	0.019	0.019	0.000	0.000	1.000
entype3privado	0.155	0.152	0.800	0.080	0.935
stuntdummy2	0.233	0.236	-0.700	-0.070	0.943
hhsiz3	4.932	4.909	1.300	0.140	0.890
momeduSECU3	0.476	0.464	2.400	0.240	0.809
wi1	0.503	0.496	3.400	0.340	0.735
wi2	0.553	0.554	-0.300	-0.030	0.977
wi3	0.616	0.612	2.400	0.250	0.806
TDA3	0.529	0.535	-1.300	-0.120	0.903

Fuente: Boyden (2022a, 2022b, 2022c) y Sanchez et al. (2022). Elaboración propia.

Tabla 23

Test de medias de covariables entre alumnos con viajes largos y cortos (control) en la ronda 3 con el punto de corte de 20 minutos y emparejamiento de kernel epanechnikov (alumnos de primaria rurales)

Variable	Promedio			t-test	
	Viajes Largos	Viajes Cortos (Control)	%sesgo	t	p>t
chsex	0.469	0.458	2.200	0.190	0.853
CDA	7.572	7.673	-4.500	-0.390	0.699
NOhablaesp3	0.483	0.479	0.700	0.060	0.951
entype3privado	0.007	0.007	-0.200	-0.010	0.988
stuntdummy2	0.607	0.600	1.300	0.110	0.912
hysize3	6.035	5.986	2.600	0.220	0.822
momeduSECU3	0.062	0.065	-1.000	-0.090	0.928
wi1	0.217	0.223	-4.500	-0.400	0.689
wi2	0.250	0.248	2.000	0.180	0.860
wi3	0.326	0.326	-0.100	-0.010	0.994
TDA3	0.938	0.946	-3.600	-0.300	0.767

Fuente: Boyden (2022a, 2022b, 2022c) y Sanchez et al. (2022). Elaboración propia.

Tabla 24

Test de medias de covariables entre alumnos con viajes largos y cortos (control) en la ronda 3 con el punto de corte de 20 minutos y emparejamiento de kernel epanechnikov (alumnos de primaria urbanos)

Variable	Promedio			t-test	
	Viajes Largos	Viajes Cortos (Control)	%sesgo	t	p>t
chsex	0.500	0.506	-1.200	-0.130	0.899
CDA	8.781	8.768	0.600	0.070	0.947
NOhablaesp3	0.048	0.042	2.800	0.260	0.793
entype3otros	0.019	0.017	1.000	0.130	0.896
entype3privado	0.152	0.149	1.000	0.110	0.916
stuntdummy2	0.233	0.235	-0.400	-0.040	0.965
hysize3	4.910	4.938	-1.600	-0.170	0.861
momeduSECU3	0.471	0.461	2.100	0.210	0.832
wi1	0.499	0.496	1.500	0.150	0.880
wi2	0.551	0.550	0.600	0.060	0.951
wi3	0.610	0.611	-0.500	-0.050	0.957
TDA3	0.524	0.547	-4.900	-0.480	0.632

Fuente: Boyden (2022a, 2022b, 2022c) y Sanchez et al. (2022). Elaboración propia.

Tabla 25

Estimación del ATT empleando bootstrapping en la ronda 3 con emparejamiento de 5 vecinos más cercanos y caliper de 0.005 (alumnos de primaria rurales y urbanos)

ATT (5 vecinos, caliper 0.005, ronda 3)		
Corte de tiempo	Rural	Urbano
>= 20 minutos	0.138	0.036
>= 25 minutos	0.305**	0.110
>= 30 minutos	0.352**	-0.028

\*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$ . Se utilizó bootstrapping para los errores estándar con 100 réplicas.

Fuente: Boyden (2022a, 2022b, 2022c) y Sanchez et al. (2022).  
Elaboración propia.

Tabla 26

Estimación del ATT empleando bootstrapping en la ronda 3 con emparejamiento de kernel epanechnikov (alumnos de primaria rurales y urbanos)

ATT (Kernel epanechnikov, ronda 3)		
Corte de tiempo	Rural	Urbano
>= 20 minutos	0.102	0.001
>= 25 minutos	0.252**	0.074
>= 30 minutos	0.318***	-0.000

\*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$ . Se utilizó bootstrapping para los errores estándar con 100 réplicas.

Fuente: Boyden (2022a, 2022b, 2022c) y Sanchez et al. (2022).  
Elaboración propia.

Los resultados muestran que los alumnos rurales de primaria en la ronda 3 que se demoran 25 minutos o más y 30 minutos o más en sus viajes poseen en promedio un mayor rendimiento académico que sus correspondientes grupos de control de alumnos con viajes cortos al usar el emparejamiento de 5 vecinos más cercanos. Las diferencias son significativas al 5% y son iguales a 0.305 y 0.352 desviaciones estándar del puntaje en la prueba de matemáticas para cada corte respectivamente. Con kernel, los mismos cortes son significativos para los alumnos rurales, pero ahora el de 30 minutos es al 1%, siendo el ATT igual a 0.318 desviaciones estándar para este y 0.252 para el corte de 25 minutos. Estas estimaciones tienen una explicación similar a las que se realizaron con la muestra total, pues la mayor oferta educativa primaria en zonas rurales permitiría a las familias de los alumnos dejar de priorizar la distancia de la institución educativa al hogar, y en su lugar enfocarse en la calidad educativa que se ofrece. Por tanto, el efecto positivo de la calidad educativa

compensaría el impacto negativo del tiempo de desplazamiento en el rendimiento académico de acuerdo a las preferencias y nivel socioeconómico de las familias (Thomas, 2016). Por otro lado, el ATT de los alumnos urbanos resulta no significativo en todos los cortes, indicando un rendimiento académico similar entre los alumnos que demoran más y menos de los intervalos propuestos, potencialmente debido a una menor oferta educativa comparada a la zona rural que limita las alternativas de elección de las familias para decidir una institución según sus prioridades.

Análogamente, en las tablas 27, 28, 29 y 30 se presenta la calidad de los emparejamientos de 5 vecinos más cercanos y kernel para el corte de 20 minutos en los alumnos rurales y urbanos de nivel secundario, en las que también se demuestra el balanceo correcto de las covariables al evidenciarse que los t-tests de medias indican que los promedios de las covariables entre los grupos de alumnos con viajes largos y cortos no poseen diferencias significativas. De igual manera, los test de medias de las covariables de los demás emparejamientos para los alumnos de secundaria se encuentran en el anexo E. Así, en las tablas 31 y 32 se enseñan las estimaciones de los ATT con los cortes de 20, 25 y 30 minutos con la población separada según la zona de residencia para los alumnos de secundaria mediante los emparejamientos de 5 vecinos más cercanos implementando un caliper de 0.005 y el emparejamiento de kernel (epanechnikov) respectivamente.

Tabla 27

Test de medias de covariables entre alumnos con viajes largos y cortos (control) en la ronda 5 con el punto de corte de 20 minutos, emparejamiento de 5 vecinos más cercanos y caliper de 0.005 (alumnos de secundaria rurales)

Variable	Promedio			t-test	
	Viajes Largos	Viajes Cortos (Control)	%sesgo	t	p>t
chsex	0.518	0.517	0.200	0.020	0.987
CDA	7.573	7.489	3.800	0.350	0.725
NOhablaesp5	0.396	0.447	-10.500	-0.930	0.353
entype5otros	0.018	0.015	3.000	0.210	0.830
entype5privado	0.018	0.019	-0.200	-0.020	0.984
stuntdummy2	0.573	0.578	-1.000	-0.090	0.931
hhsiz5	5.494	5.440	2.800	0.250	0.800
momeduSECU5	0.110	0.109	0.300	0.030	0.979
wi1	0.225	0.212	9.800	0.930	0.354
wi4	0.399	0.397	1.400	0.130	0.893
wi5	0.436	0.450	-11.100	-1.090	0.277
TDA5	0.762	0.743	4.400	0.390	0.694

Fuente: Boyden (2022a, 2022b, 2022c) y Sanchez et al. (2022). Elaboración propia.

Tabla 28

Test de medias de covariables entre alumnos con viajes largos y cortos (control) en la ronda 5 con el punto de corte de 20 minutos, emparejamiento de 5 vecinos más cercanos y caliper de 0.005 (alumnos de secundaria urbanos)

Variable	Promedio			t-test	
	Viajes Largos	Viajes Cortos (Control)	%sesgo	t	p>t
chsex	0.533	0.523	2.100	0.240	0.810
CDA	8.722	8.764	-2.200	-0.260	0.795
NOhablaesp5	0.037	0.028	4.600	0.580	0.562
entype5otros	0.015	0.020	-3.200	-0.460	0.646
entype5privado	0.211	0.201	2.500	0.290	0.774
stuntdummy2	0.207	0.188	4.700	0.560	0.573
hhsiz5	5.152	5.121	1.600	0.190	0.846
momeduSECU5	0.493	0.530	-7.500	-0.870	0.386
wi1	0.503	0.512	-4.000	-0.460	0.646
wi4	0.667	0.679	-7.700	-0.930	0.355
wi5	0.703	0.709	-4.100	-0.490	0.627
TDA5	0.404	0.379	5.000	0.580	0.561

Fuente: Boyden (2022a, 2022b, 2022c) y Sanchez et al. (2022). Elaboración propia.

Tabla 29

Test de medias de covariables entre alumnos con viajes largos y cortos (control) en la ronda 5 con el punto de corte de 20 minutos y emparejamiento de kernel epanechnikov (alumnos de secundaria rurales)

Variable	Promedio			t-test	
	Viajes Largos	Viajes Cortos (Control)	%sesgo	t	p>t
chsex	0.533	0.531	0.600	0.060	0.956
CDA	7.595	7.672	-3.500	-0.360	0.719
NOhablaesp5	0.451	0.402	10.200	0.980	0.326
entype5otros	0.015	0.008	7.300	0.680	0.497
entype5privado	0.015	0.016	-0.200	-0.030	0.978
stuntdummy2	0.579	0.620	-8.200	-0.820	0.414
hhsiz5	5.610	5.435	9.100	0.910	0.365
momeduSECU5	0.108	0.114	-1.800	-0.190	0.853
wi1	0.223	0.208	10.600	1.100	0.272
wi4	0.389	0.392	-2.100	-0.220	0.828
wi5	0.424	0.431	-5.300	-0.550	0.584
TDA5	0.677	0.655	5.100	0.450	0.652

Fuente: Boyden (2022a, 2022b, 2022c) y Sanchez et al. (2022). Elaboración propia.

Tabla 30

Test de medias de covariables entre alumnos con viajes largos y cortos (control) en la ronda 5 con el punto de corte de 20 minutos y emparejamiento de kernel epanechnikov (alumnos de secundaria urbanos)

Variable	Promedio			t-test	
	Viajes Largos	Viajes Cortos (Control)	%sesgo	t	p>t
chsex	0.538	0.528	2.000	0.230	0.816
CDA	8.665	8.775	-5.700	-0.670	0.505
NOhablaesp5	0.036	0.034	1.000	0.130	0.900
entype5otros	0.015	0.019	-2.500	-0.380	0.702
entype5privado	0.215	0.216	-0.300	-0.030	0.974
stuntdummy2	0.204	0.205	-0.300	-0.040	0.970
hhsiz5	5.218	5.146	3.900	0.450	0.654
momeduSECU5	0.487	0.514	-5.400	-0.630	0.530
wi1	0.500	0.507	-3.500	-0.410	0.684
wi4	0.664	0.674	-6.500	-0.770	0.443
wi5	0.701	0.707	-4.200	-0.490	0.622
TDA5	0.396	0.382	2.800	0.330	0.739

Fuente: Boyden (2022a, 2022b, 2022c) y Sanchez et al. (2022). Elaboración propia.

Tabla 31

Estimación del ATT empleando bootstrapping en la ronda 5 con emparejamiento de 5 vecinos más cercanos y caliper de 0.005 (alumnos de secundaria rurales y urbanos)

ATT (5 vecinos, caliper 0.005, ronda 5)		
Corte de tiempo	Rural	Urbano
>= 20 minutos	0.008	-0.135*
>= 25 minutos	0.002	-0.169
>= 30 minutos	0.039	-0.022

\*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$ . Se utilizó bootstrapping para los errores estándar con 100 réplicas.

Fuente: Boyden (2022a, 2022b, 2022c) y Sanchez et al. (2022).  
Elaboración propia.

Tabla 32

Estimación del ATT empleando bootstrapping en la ronda 5 con emparejamiento de kernel epanechnikov (alumnos de secundaria rurales y urbanos)

ATT (Kernel epanechnikov, ronda 5)		
Corte de tiempo	Rural	Urbano
>= 20 minutos	0.039	-0.149**
>= 25 minutos	0.084	-0.141
>= 30 minutos	0.068	-0.128

\*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$ . Se utilizó bootstrapping para los errores estándar con 100 réplicas.

Fuente: Boyden (2022a, 2022b, 2022c) y Sanchez et al. (2022).  
Elaboración propia.

Para los alumnos de secundaria en la ronda 5, se puede observar que en los estudiantes urbanos el tiempo de desplazamiento afecta negativamente a su rendimiento académico, mientras que para los rurales este impacto es positivo y reducido. Sin embargo, este solo es significativo al 10% en el corte de 20 minutos para los alumnos urbanos, en el que el ATT es igual a -0.135 desviaciones estándar del puntaje en la prueba de matemáticas con el emparejamiento de 5 vecinos más cercanos. Para el emparejamiento con kernel, también solo es significativo la estimación del efecto con el corte de 20 minutos para los alumnos urbanos; sin embargo, la significancia ahora es al 5% y el ATT es igual a -0.149 desviaciones estándar. La significancia en este caso resulta atípica comparada al resto de los resultados tanto de los alumnos de secundaria urbanos y rurales, potencialmente atribuida a una muestra reducida al separar la población según zona de residencia.

Aparte de este resultado, las demás estimaciones para los alumnos rurales y urbanos de secundaria señalan que ellos no ven afectado significativamente su rendimiento académico en matemática por el tiempo de desplazamiento que realizan, lo cual podría estar atribuido a la menor oferta educativa comparado al nivel primario, por lo que la calidad educativa recibida sería similar para quienes residen en áreas aledañas al acceder a las mismas instituciones educativas. Adicionalmente, se destaca la similitud de las estimaciones utilizando el emparejamiento de vecinos más cercanos y kernel en las muestras de alumnos rurales y urbanos tanto para primaria y secundaria. Esto podría estar atribuido a la división de la población en grupos con características más homogéneas comparado con la muestra total al poseer insumos similares de la oferta y demanda educativa (Hanushek, 1992). Cabe añadir que los resultados de todas estas estimaciones del efecto del tiempo de desplazamiento en el rendimiento académico de los alumnos de primaria y secundaria según la zona de residencia sin implementar bootstrapping se encuentran en el anexo F.

Por otro lado, se realizó un análisis sobre la linealidad del efecto del tiempo de desplazamiento sobre el rendimiento académico en los alumnos utilizando la muestra total. La razón de este análisis yace en la postura Córdova y Ruiz (2014), quienes indican que un mayor tiempo de desplazamiento genera que los alumnos aprovechen en menor medida los insumos de la demanda y oferta educativa, mermando su proceso de aprendizaje. Por ello, se dividió la población con viajes largos en dos, los que se demoran menos de 40 minutos y los que tardan igual o más de 40 minutos en sus viajes de ida a la institución educativa para el nivel educativo primario y secundario con el motivo de analizar si dentro de esta población el efecto del tiempo de desplazamiento sobre el rendimiento académico es heterogéneo. De esta forma, la población de primaria con viajes largos mayores o iguales a 40 minutos es conformada por 61 alumnos de los 1524 totales, mientras que la de secundaria se constituye de 121 de los 1518.

Para el cálculo del ATT en este ejercicio se utilizaron las mismas especificaciones que en la sección de los resultados de las estimaciones principales (emparejamiento de 5 vecinos más cercanos, caliper de 0.005 y cortes de 20, 25 y 30 minutos). Además, para este análisis también se utilizó bootstrapping con 100 repeticiones por el reducido tamaño de las nuevas muestras. Así como en el análisis del efecto del tiempo de desplazamiento sobre el rendimiento académico según la

zona de residencia, se presentan en las tablas 33, 34, 35 y 36 el balanceo de las covariables con el corte de 20 minutos. De estas se puede observar una adecuada calidad de emparejamiento de las dos poblaciones de alumnos con viajes largos con el grupo control de estudiantes con viajes cortos para ambos niveles educativos, manteniéndose esta efectividad del emparejamiento en los demás cortes como se puede observar en el anexo G. Sin embargo, como se puede ver en las tablas 37<sup>3</sup> y 38, los resultados encontrados no señalan una diferencia significativa entre estas dos poblaciones de viajes largos en los alumnos ya sean de nivel primario o secundario respectivamente al utilizar los tres puntos de corte de 20, 25 y 30 minutos. Por último, cabe comentar que las estimaciones de este análisis para los alumnos de primaria y secundaria sin utilizar bootstrapping se encuentran en el anexo F.

Tabla 33

Test de medias de covariables entre alumnos con viajes largos menores a 40 minutos y viajes cortos (control) en la ronda 3 con el punto de corte de 20 minutos,

Variable	Promedio			t-test	
	Viajes Largos <40 min	Viajes Cortos (Control)	%sesgo	t	p>t
chsex	0.484	0.467	3.400	0.410	0.680
CDA	8.418	8.483	-3.100	-0.370	0.709
NOhablaesp3	0.178	0.161	4.900	0.540	0.589
ruralAurbano1a3	0.077	0.070	2.500	0.290	0.774
urbanoArural1a3	0.028	0.025	1.600	0.180	0.856
solourbano1a3	0.589	0.611	-4.600	-0.530	0.593
entype3otros	0.010	0.011	-0.500	-0.080	0.936
entype3privado	0.105	0.115	-3.200	-0.400	0.689
stuntdummy2	0.366	0.367	-0.200	-0.020	0.980
hhsiz3	5.369	5.362	0.400	0.050	0.962
momeduSECU3	0.324	0.336	-2.400	-0.300	0.767
wi1	0.402	0.408	-3.000	-0.360	0.717
wi2	0.449	0.462	-5.900	-0.710	0.478
wi3	0.517	0.529	-6.000	-0.720	0.474
TDA3	0.679	0.646	7.500	0.860	0.391

Fuente: Boyden (2022a, 2022b, 2022c) y Sanchez et al. (2022). Elaboración propia.

<sup>3</sup> Para los resultados de los alumnos de primaria en los cortes de 25 y 30 minutos, no se consideraron a los individuos (40 observaciones) que asistían a instituciones de gestión alterna en la estimación del ATT en la población de alumnos con viajes largos menores a 40 minutos debido a la ausencia de alumnos matriculados en instituciones educativas alternas en este último grupo.

Tabla 34

Test de medias de covariables entre alumnos con viajes largos mayores o iguales a 40 minutos y viajes cortos (control) en la ronda 3 con el punto de corte de 20 minutos, emparejamiento de 5 vecinos más cercanos y caliper de 0.005 (alumnos de primaria)

Variable	Promedio			t-test	
	Viajes Largos >= 40 min	Viajes Cortos (Control)	%sesgo	t	p>t
chsex	0.475	0.464	2.100	0.110	0.910
CDA	7.797	7.731	3.000	0.150	0.881
NOhablaesp3	0.390	0.369	5.200	0.240	0.814
ruralAurbano1a3	0.034	0.034	0.000	0.000	1.000
urbanoArural1a3	0.068	0.069	-0.400	-0.020	0.986
solourbano1a3	0.254	0.258	-0.800	-0.040	0.967
entype3otros	0.017	0.024	-4.300	-0.260	0.796
entype3privado	0.051	0.047	1.100	0.080	0.933
stuntdummy2	0.475	0.476	-0.200	-0.010	0.990
hhsiz3	5.593	5.580	0.700	0.040	0.967
momeduSECU3	0.237	0.258	-4.400	-0.250	0.800
wi1	0.323	0.316	2.900	0.160	0.873
wi2	0.342	0.342	-0.400	-0.020	0.982
wi3	0.418	0.426	-3.900	-0.210	0.831
TDA3	0.814	0.775	9.400	0.510	0.612

Fuente: Boyden (2022a, 2022b, 2022c) y Sanchez et al. (2022). Elaboración propia.

Tabla 35

Test de medias de covariables entre alumnos con viajes largos menores a 40 minutos y viajes cortos (control) en la ronda 5 con el punto de corte de 20 minutos, emparejamiento de 5 vecinos más cercanos y caliper de 0.005 (alumnos de secundaria)

Variable	Promedio			t-test	
	Viajes Largos <40 min	Viajes Cortos (Control)	%sesgo	t	p>t
chsex	0.512	0.535	-4.600	-0.590	0.553
CDA	8.261	8.261	0.000	0.000	1.000
NOhablaesp5	0.136	0.153	-5.000	-0.620	0.538
ruralAurbano1a5	0.064	0.068	-1.900	-0.250	0.802
urbanoArural1a5	0.033	0.025	5.000	0.610	0.544
solourbano1a5	0.633	0.629	0.900	0.110	0.911
entype5otros	0.012	0.018	-3.600	-0.580	0.563
entype5privado	0.167	0.153	3.800	0.480	0.629
stuntdummy2	0.315	0.323	-1.700	-0.220	0.830
hhsiz5	5.291	5.232	3.200	0.400	0.686
momeduSECU5	0.379	0.364	3.000	0.390	0.698
wi1	0.418	0.417	0.500	0.060	0.952
wi4	0.585	0.582	1.700	0.210	0.832
wi5	0.617	0.615	0.700	0.090	0.925
TDA5	0.524	0.529	-0.900	-0.110	0.909

Fuente: Boyden (2022a, 2022b, 2022c) y Sanchez et al. (2022). Elaboración propia.

Tabla 36

Test de medias de covariables entre alumnos con viajes largos mayores o iguales a 40 minutos y viajes cortos (control) en la ronda 5 con el punto de corte de 20 minutos, emparejamiento de 5 vecinos más cercanos y caliper de 0.005 (alumnos de secundaria)

Variable	Promedio			t-test	
	Viajes Largos >= 40 min	Viajes Cortos (Control)	%sesgo	t	p>t
chsex	0.537	0.495	8.400	0.610	0.539
CDA	8.009	8.319	-14.300	-1.010	0.314
NOhablaesp5	0.241	0.270	-7.600	-0.480	0.629
ruralAurbano1a5	0.074	0.056	7.000	0.550	0.583
urbanoArural1a5	0.046	0.029	9.900	0.660	0.511
solourbano1a5	0.343	0.400	-12.700	-0.870	0.385
entype5otros	0.028	0.017	6.500	0.550	0.582
entype5privado	0.056	0.041	4.800	0.490	0.625
stuntdummy2	0.417	0.387	6.200	0.440	0.659
hhsiz5	5.509	5.557	-2.500	-0.180	0.854
momeduSECU5	0.278	0.318	-8.500	-0.640	0.524
wi1	0.324	0.347	-10.600	-0.790	0.433
wi4	0.495	0.516	-11.000	-0.810	0.422
wi5	0.546	0.564	-10.000	-0.700	0.483
TDA5	0.574	0.502	14.700	1.060	0.291

Fuente: Boyden (2022a, 2022b, 2022c) y Sanchez et al. (2022). Elaboración propia.

Tabla 37

Estimación del ATT empleando bootstrapping en la ronda 3 con emparejamiento de 5 vecinos más cercanos, caliper de 0.005 y dividiendo los alumnos con viajes largos a partir del intervalo de 40 minutos (alumnos de primaria)

Corte de tiempo	ATT (5 vecinos, caliper 0.005, ronda 3)	
	Alumnos con viajes largos < 40 min	Alumnos con viajes largos >= 40 min
>= 20 minutos	-0.016	0.265
>= 25 minutos	0.194*	0.197
>= 30 minutos	0.152	0.164

\*\*\* p<0.01, \*\* p<0.05, \* p<0.1. Se utilizó bootstrapping para los errores estándar con 100 réplicas.

Fuente: Boyden (2022a, 2022b, 2022c) y Sanchez et al. (2022). Elaboración propia.

Tabla 38

Estimación del ATT empleando bootstrapping en la ronda 5 con emparejamiento de 5 vecinos más cercanos, caliper de 0.005 y dividiendo los alumnos con viajes largos a partir del intervalo de 40 minutos (alumnos de secundaria)

ATT (5 vecinos, caliper 0.005, ronda 5)		
Corte de tiempo	Alumnos con viajes largos < 40 min	Alumnos con viajes largos $\geq$ 40 min
$\geq$ 20 minutos	-0.029	-0.014
$\geq$ 25 minutos	-0.007	0.052
$\geq$ 30 minutos	-0.026	0.013

\*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$ . Se utilizó bootstrapping para los errores estándar con 100 réplicas.

Fuente: Boyden (2022a), Boyden (2022b) y Sanchez et al. (2022). Elaboración propia.



## Conclusiones

El presente trabajo analiza el impacto del tiempo de desplazamiento desde la residencia hasta la institución educativa en el rendimiento académico de los alumnos de primaria y secundaria en el Perú. Mediante la implementación de un modelo económico que señala que insumos pasados y presentes de la demanda educativa, oferta educativa, habilidades innatas e influencias externas inciden en el rendimiento académico de los alumnos, este documento propone que el tiempo de desplazamiento de ida desde su residencia hacia su institución educativa de los alumnos de primaria y secundaria del Perú irrumpe en esta relación mermando la efectividad de los insumos en el resultado del rendimiento académico. Este factor perjudica el proceso de aprendizaje de los alumnos al perturbar la asignación de su tiempo, siendo los principales canales el menor tiempo de estudio, disminución de la supervisión de los padres, cansancio y potenciales distracciones en el trayecto (Córdova & Ruiz, 2014). No obstante, la literatura previa encuentra impactos positivos, negativos y ambiguos del tiempo de desplazamiento sobre el rendimiento académico de los alumnos, por lo que se considera la existencia de efectos heterogéneos de este tiempo de acuerdo a diferentes características de la demanda y oferta educativa.

Por ello, se plantea que el tiempo de desplazamiento posee un efecto negativo en el rendimiento académico de los alumnos de primaria y secundaria del Perú, siendo este impacto más profundo en este último nivel por una menor oferta educativa comparada a la primaria. Así, con el uso de la base de datos de Niños del Milenio y la técnica de Propensity Score Matching utilizando cortes de tiempo de 20, 25 y 30 minutos para separar y emular el recibimiento de un tratamiento en la metodología de impacto, se procede a calcular el efecto del tiempo de desplazamiento en el rendimiento académico reflejado en las calificaciones obtenidas de los alumnos en las evaluaciones de matemática realizadas y registradas en la base mencionada.

De esta manera, los resultados no coinciden con la hipótesis planteada, indicando que el tiempo de desplazamiento beneficia el rendimiento académico de los alumnos de primaria de manera significativa usando el corte más largo de 30 minutos, mientras que para los de secundaria el impacto es negativo en la mayoría de los cortes, pero sin ser significativo. Asimismo, al analizar esta interacción desagregando la población según la zona de residencia, se halla el resultado que para los alumnos rurales de primaria el efecto es positivo y significativo con los cortes de 25 y 30 minutos.

En el caso de secundaria, el impacto es negativo y significativo al 10% para los alumnos urbanos con el corte de 20 minutos, siendo el resto de los casos no significativos. Adicionalmente, al analizar la linealidad del efecto del tiempo de desplazamiento, no se encuentra una tendencia consistente y significativa entre la población de alumnos con viajes largos al dividirlos a partir del intervalo de 40 minutos.

Por lo tanto, las estimaciones sugieren que un mayor tiempo de desplazamiento a la institución educativa puede representar un costo que las familias y alumnos deciden asumir con el objetivo de asistir a un colegio con mayor calidad educativa comparado a los que están más cercanos a su residencia, dada una oferta educativa ostentosa y un tiempo de viaje lo suficiente que refleje este sacrificio (Hanushek, 1992). De esta manera, la mejor calidad educativa tendría un efecto positivo en el rendimiento académico del alumno, y compensaría el potencial efecto negativo del tiempo de viaje, lo que se evidencia en los resultados de los alumnos de primaria, particularmente en los que residen en zonas rurales. Por último, cabe añadir el aporte de literatura del presente trabajo que consiste en utilizar tanto el nivel primario como secundario en el análisis del efecto del tiempo de desplazamiento sobre el rendimiento, así como su comparación.

### Limitaciones

Las limitaciones principales del trabajo recaen en la reducida muestra de alumnos en los niveles de primaria y secundaria, así como la dificultad de su desagregación en subgrupos, por lo que se procedió a implementar la técnica de bootstrapping. Adicionalmente, la escasa disponibilidad de variables que definan las instituciones educativas supone una limitación para controlar el impacto de la oferta educativa sobre el proceso de aprendizaje de los alumnos. Por ejemplo, si bien es una característica difícil de interpretar, la falta de una variable que refleje la calidad educativa de la institución educativa percibida por las familias o los alumnos restringe un análisis más exhaustivo del efecto del tiempo de desplazamiento sobre el rendimiento académico de estos últimos. No obstante, esta limitación se buscó corregir implementando las covariables de tipo de gestión de la institución educativa a la que asiste el alumno.



## Referencias Bibliográficas:

- Advisory Committee on Mathematics Education. (2011). *Mathematical Needs: Mathematics in The Workplace and in Higher Education*. The Royal Society. <https://royalsociety.org/news-resources/publications/2011/mathematical-needs-mathematics-in-the-workplace-and-in-higher-education/>
- Aguilar, D. (2017). *Una aproximación econométrica a los efectos del programa JUNTOS en el rendimiento escolar de los niños peruanos* [Tesis de Maestría, Pontificia Universidad Católica del Perú]. Repositorio PUCP. <http://hdl.handle.net/20.500.12404/9373>
- Boyden, J. (2022a). *Young Lives: an International Study of Childhood Poverty: Round 2, 2006*. [data collection]. 4th Edition. UK Data Service. SN: 6852, DOI: <http://doi.org/10.5255/UKDA-SN-6852-4>
- Boyden, J. (2022b). *Young Lives: an International Study of Childhood Poverty: Round 3, 2009*. [data collection]. 4th Edition. UK Data Service. SN: 6853, DOI: <http://doi.org/10.5255/UKDA-SN-6853-4>
- Boyden, J. (2022c). *Young Lives: an International Study of Childhood Poverty: Rounds 1-5 Constructed Files, 2002-2016*. [data collection]. 5th Edition. UK Data Service. SN: 7483 DOI: <http://doi.org/10.5255/UKDA-SN-7483-5>
- Cabuquin, J.C. & Abocejo, F.T. (2023). Mathematics learners' performance and academic achievement at a public high school institution in Leyte, Philippines. *Formatif: Jurnal Ilmiah Pendidikan MIPA*, 13(2), 123-136. <http://dx.doi.org/10.30998/formatif.v13i2.17235>
- Contreras, D., Hojman, D., Matas, M., Rodríguez, P., & Suárez, N. (2018). The impact of commuting time over educational achievement: a machine learning approach (Series Documentos de Trabajo No. 472). Universidad de Chile. <https://repositorio.uchile.cl/handle/2250/153416>
- Córdova, C., & Ruiz, F. (2014). Impacto del tiempo de desplazamiento a la institución educativa sobre el rendimiento académico en el ámbito rural. En *Economía aplicada: Ensayos de investigación económica 2013* (pp. 79-104). Universidad del Pacífico.

- Chaman, J. & Moali, J. (2020). *Lo público vs lo privado: Un análisis de los determinantes de la elección escolar en la escuela primaria en el Perú* [Tesis de bachillerato, Pontificia Universidad Católica del Perú]. Repositorio PUCP. <http://hdl.handle.net/20.500.12404/21578>
- Cueto, S., Escobal, J., Felipe, C., Pazos, N., Penny, M., Rojas, V. & Sánchez, A. (2018) *¿Qué hemos aprendido del estudio longitudinal Niños del Milenio? Síntesis de hallazgos*. Lima: Niños del Milenio; GRADE.
- Ding, P., & Feng, S. (2022). How School Travel Affects Children's Psychological Well-Being and Academic Achievement in China. *International journal of environmental research and public health*, 19(21), 13881
- Dirección Nacional de Prospectiva y Estudios Estratégicos. (2023). *Integración urbano-rural para el desarrollo sostenible en el Perú*. Centro Nacional de Planeamiento Estratégico. <https://www.gob.pe/institucion/ceplan/informes-publicaciones/4099589-integracion-urbano-rural-para-el-desarrollo-sostenible-en-el-peru>
- Duraisamy, P., Tan, J., James, E., & Lane, J. (1997). Is there a quantity-quality tradeoff as enrollments increase? Evidence from Tamil Nadu, India. *Policy Research Working Papers*.
- Espinoza, P. & Douglas, S. (2022). *Cognitive and Achievement Tests in the Young Lives Study*. Technical Note 53, Oxford: Young Lives.
- García-Hermoso, A., Saavedra, J. M., Olloquequi, J., & Ramírez-Vélez, R. (2017). Associations between the duration of active commuting to school and academic achievement in rural Chilean adolescents. *Environmental Health and Preventive Medicine*, 22(1).
- Girard C, Bastelica T, Léone J, Epinat-Duclos J, Longo L, Prado J (2021) The relation between home numeracy practices and a variety of math skills in elementary school children. *PLoS ONE*, 16(9). <https://doi.org/10.1371/journal.pone.0255400>
- Gutiérrez, J. (2022). *¿Mayor conectividad entre escuelas reduce la deserción escolar? Un análisis geoespacial en la transición educativa primaria – secundaria en el Perú* [Tesis de maestría, Universidad del Pacífico]. Repositorio Institucional UP. <https://hdl.handle.net/11354/3393>

- Guhl, P. (2019). The impact of early math and numeracy skills on academic achievement in elementary school. *Master's Theses & Capstone Projects Education, Northwestern College, Iowa*. [https://nwcommons.nwciowa.edu/education\\_masters/155/](https://nwcommons.nwciowa.edu/education_masters/155/)
- Hanushek, E. A. (1979). Conceptual and empirical issues in the estimation of educational production functions. *The Journal of Human Resources*, 14(3), 351.
- Hanushek, E. (1992). The Trade-off between Child Quantity and Quality. *The Journal of Political Economy*, 100(1), 84–117. <https://www.jstor.org/stable/2138807>
- He, S. Y., & Giuliano, G. (2017). School choice: Understanding the trade-off between travel distance and school quality. *Transportation*, 45(5), 1475-1498. <https://doi.org/10.1007/s11116-017-9773-3>
- Instituto Nacional de Estadística e Informática. (2024a). *Encuesta Nacional de Uso del Tiempo 2024*. Instituto Nacional de Estadística e Informática. [https://proyectos.inei.gob.pe/microdatos/Consulta\\_por\\_Encuesta.asp](https://proyectos.inei.gob.pe/microdatos/Consulta_por_Encuesta.asp)
- Instituto Nacional de Estadística e Informática. (2024a). *Encuesta Nacional de Uso del Tiempo 2024: Principales Resultados*. Instituto Nacional de Estadística e Informática. <https://cdn.www.gob.pe/uploads/document/file/7838198/6610655-encuesta-nacional-de-uso-del-tiempo-2024-principales-resultados.pdf?v=1743092246>
- Instituto Peruano de Economía. (2021). *Efectos del COVID-19 en la educación*. Informes. <https://www.ipe.org.pe/portal/efectos-del-covid-19-en-la-educacion/#:~:text=Esto%20significa%20que%20m%C3%A1s%20de,de%2093%25%20a%2081%25>
- Leite, W. (2017). *Practical propensity score methods using r*. SAGE Publications, Inc, <https://doi.org/10.4135/9781071802854>
- Li, L. & Haoming, L. (2009). *The Effect of Primary School Availability on Middle School Attainment in Rural China*.
- Levin, H. M. (1974). Measuring efficiency in educational production. *Public Finance Quarterly*, 2(1), 3-24.

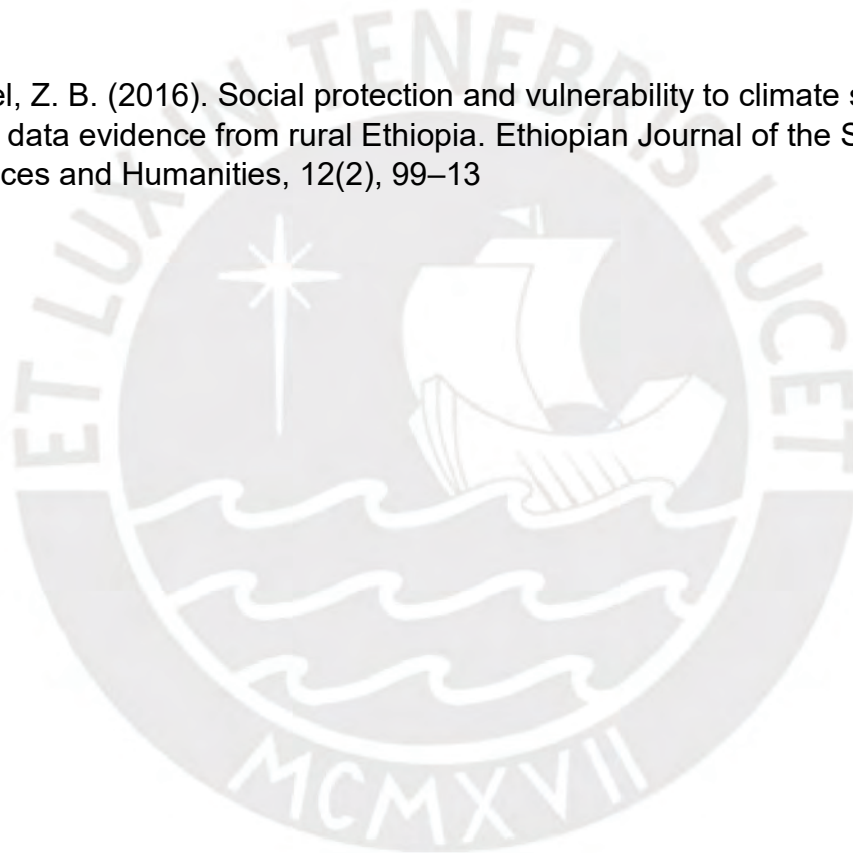
- Ministerio de Educación. (2023). *PISA 2022: El Perú mantiene sus resultados en las competencias de Lectura Y Ciencia*. <https://umc.minedu.gob.pe/pisa-2022-el-peru-mantiene-sus-resultados-en-las-competencias-de-lectura-y-ciencia/>
- Portocarrero, J. (2023). *Evolución del Gasto Público en Educación en el Perú 2015-2021*. Campaña Latinoamericana por el Derecho a la Educación.
- Ryan, J., & Brockmann, F. (2018). *A Practitioner's Introduction to Equating* (1ra edición revisada). Council of Chief State School Officers. <https://ccsso.org/resourcelibrary/practitioners-introduction-equating>
- Sanchez, A., Woldehanna, T., Duc, L. Thuc, Boyden, J., Penny, M., Galab, S. (2022). *Young Lives: an International Study of Childhood Poverty: Round 5, 2016*. [data collection]. 2nd Edition. UK Data Service. SN: 8357, DOI: <http://doi.org/10.5255/UKDA-SN-8357-2>
- Steinmayr, R., Meißner, A., Weidinger, A. F., & Wirthwein, L. (2014). Academic achievement. Oxford Bibliographies Online
- Thomas, C. (2016). *Going the Distance! How Distance to School Relates to Education Outcomes* [Tesis de doctorado, Universidad de California]. <https://escholarship.org/uc/item/97f0n3mq>
- Tigre, R. (2014). *The impact of commuting duration on youth's academic performance: could mobility compromise school achievement?* [Tesis de Maestría, Universidade Federal De Pernambuco]. Repositorio Digital da UFPE. <https://repositorio.ufpe.br/handle/123456789/14998>
- Tigre, R., Sampaio, B., & Menezes, T. (2017). The impact of commuting time on youth's school performance. *Journal of Regional Science*, 57(1), 28–47. <https://doi.org/10.1111/jors.12289>
- Unidad Estadística de Calidad Educativa. (2023). *Magnitudes de la educación en el Perú*. [https://escale.minedu.gob.pe/magnitudes-portlet/reporte/cuadro?anio=35&cuadro=647&forma=C&dpto=&dre=&tipo\\_ambito=ambito-ubigeo#urbana](https://escale.minedu.gob.pe/magnitudes-portlet/reporte/cuadro?anio=35&cuadro=647&forma=C&dpto=&dre=&tipo_ambito=ambito-ubigeo#urbana)
- Vélez, M., Egurrola, J., & Jaimes Barragán, F. (2012). Ronda clínica Y epidemiológica. USO de la puntuación de propension (propensity score) en estudios no experimentales. *Iatreia*, 26(1). <https://doi.org/10.17533/udea.iatreia.13647>

Villanueva, A. (2023). *El efecto del tiempo de desplazamiento de la vivienda al centro educativo sobre el rendimiento escolar en el área rural del Perú* [Tesis de licenciatura, Pontificia Universidad Católica del Perú]. Repositorio PUCP. <http://hdl.handle.net/20.500.12404/26926>

Vuri, D. (2007). *The effects of availability and distance from school on children's time allocation in Ghana and Guatemala*. Working Paper May 2007, Understanding Children Work (UCW) Project, University of Rome, Rome Tor Vergata.

Yeung, R., & Nguyen-Hoang, P. (2020). It's the journey, not the destination: the effect of school travel mode on student achievement. *Journal of Urbanism: International Research on Placemaking and Urban Sustainability*.

Weldegebriel, Z. B. (2016). Social protection and vulnerability to climate shocks: a panel data evidence from rural Ethiopia. *Ethiopian Journal of the Social Sciences and Humanities*, 12(2), 99–13



## Anexos

Anexo A: Regresiones logit para el cálculo del Propensity score con los puntos de corte de 20, 25 y 30 minutos en las rondas 3 y 5 (5 vecinos más cercanos y caliper 0.005)

Tabla 39

Regresión logit para el Propensity score con el corte de 20 minutos en la ronda 3 (alumnos de primaria)

Observaciones = 1,524  
LR chi2(15) = 118.66  
Prob > chi2 = 0.0000  
Pseudo R2 = 0.0715

Log likelihood = -770.27805

TREAT_20MIN3	Coefficiente	Error std.	z	P >  z	[Intervalo	Conf. 95%]
chsex	-0.097	0.127	-0.770	0.442	-0.346	0.151
CDA	-0.005	0.033	-0.140	0.890	-0.069	0.060
NOhablaesp3	0.662***	0.199	3.320	0.001	0.271	1.052
ruralAurbano1a3	0.011	0.282	0.040	0.969	-0.542	0.564
urbanoArural1a3	0.134	0.397	0.340	0.736	-0.644	0.912
solourbano1a3	-0.643***	0.205	-3.140	0.002	-1.044	-0.241
entype3otros	-0.821	0.547	-1.500	0.134	-1.893	0.252
entype3privado	-0.325	0.230	-1.420	0.156	-0.775	0.125
stuntdummy2	0.019	0.149	0.130	0.896	-0.273	0.312
hhsiz3	-0.060*	0.037	-1.650	0.099	-0.132	0.011
momeduSECU3	-0.026	0.169	-0.150	0.879	-0.356	0.305
wi1	0.033	0.528	0.060	0.950	-1.001	1.068
wi2	0.589	0.619	0.950	0.341	-0.623	1.802
wi3	-1.778***	0.617	-2.880	0.004	-2.987	-0.569
TDA3	-1.000***	0.162	-6.170	0.000	-1.317	-0.682
Constante	0.964**	0.426	2.260	0.024	0.128	1.800

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

Fuente: Boyden (2022a, 2022b, 2022c) y Sanchez et al. (2022). Elaboración propia.

Tabla 40

Regresión logit para el Propensity score con el corte de 25 minutos en la ronda 3 (alumnos de primaria)

Observaciones = 1,524  
 LR chi2(15) = 126.88  
 Prob > chi2 = 0.0000  
 Pseudo R2 = 0.1054

Log likelihood = -538.35809

TREAT_25MIN3	Coefficiente	Error std.	z	P >  z	[Intervalo	Conf. 95%]
chsex	-0.007	0.159	-0.040	0.965	-0.318	0.304
CDA	0.003	0.040	0.080	0.938	-0.076	0.082
NOhablaesp3	0.735***	0.220	3.350	0.001	0.305	1.166
ruralAurbano1a3	-0.084	0.325	-0.260	0.796	-0.721	0.553
urbanoArural1a3	-0.342	0.490	-0.700	0.485	-1.303	0.618
solourbano1a3	-1.013***	0.253	-4.000	0.000	-1.509	-0.516
entype3otros	-1.276	1.033	-1.230	0.217	-3.301	0.749
entype3privado	0.059	0.303	0.190	0.846	-0.536	0.654
stuntdummy2	-0.069	0.181	-0.380	0.702	-0.425	0.286
hysize3	-0.060	0.046	-1.310	0.190	-0.150	0.030
momeduSECU3	0.035	0.226	0.150	0.878	-0.409	0.478
wi1	0.115	0.687	0.170	0.867	-1.230	1.461
wi2	0.140	0.780	0.180	0.857	-1.389	1.669
wi3	-2.200***	0.768	-2.870	0.004	-3.704	-0.695
TDA3	-0.916***	0.215	-4.270	0.000	-1.336	-0.495
Constante	0.599	0.528	1.130	0.257	-0.435	1.633

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

Fuente: Boyden (2022a, 2022b, 2022c) y Sanchez et al. (2022). Elaboración propia.

Tabla 41

Regresión logit para el Propensity score con el corte de 30 minutos en la ronda 3 (alumnos de primaria)

Observaciones = 1,524  
 LR chi2(15) = 116.15  
 Prob > chi2 = 0.0000  
 Pseudo R2 = 0.1046

Log likelihood = -497.36059

TREAT_30MIN3	Coefficiente	Error std.	z	P >  z	[Intervalo	Conf. 95%]
chsex	0.026	0.167	0.150	0.877	-0.301	0.352
CDA	-0.000	0.042	-0.010	0.996	-0.083	0.082
NOhablaesp3	0.613***	0.226	2.710	0.007	0.170	1.057
ruralAurbano1a3	-0.288	0.343	-0.840	0.402	-0.961	0.385
urbanoArural1a3	-0.496	0.522	-0.950	0.342	-1.518	0.526
solourbano1a3	-1.239***	0.269	-4.610	0.000	-1.767	-0.712
entype3otros	-1.134	1.037	-1.090	0.274	-3.166	0.897
entype3privado	0.117	0.322	0.360	0.717	-0.514	0.748
stuntdummy2	0.015	0.189	0.080	0.937	-0.355	0.385
hysize3	-0.043	0.048	-0.900	0.370	-0.136	0.051
momeduSECU3	0.172	0.241	0.710	0.475	-0.300	0.644
wi1	0.153	0.734	0.210	0.835	-1.285	1.591
wi2	-0.090	0.828	-0.110	0.913	-1.712	1.531
wi3	-1.747**	0.808	-2.160	0.031	-3.330	-0.163
TDA3	-0.824***	0.229	-3.600	0.000	-1.272	-0.376
Constante	0.227	0.551	0.410	0.680	-0.852	1.307

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

Fuente: Boyden (2022a, 2022b, 2022c) y Sanchez et al. (2022). Elaboración propia.

Tabla 42

Regresión logit para el Propensity score con el corte de 20 minutos en la ronda 5 (alumnos de secundaria)

Observaciones = 1,518  
 LR chi2(15) = 177.67  
 Prob > chi2 = 0.0000  
 Pseudo R2 = 0.0945

Log likelihood = -851.29535

TREAT_20MIN5	Coficiente	Error std.	z	P >  z	[Intervalo	Conf. 95%]
chsex	0.211	0.118	1.780	0.075	-0.021	0.443
CDA	-0.033	0.031	-1.060	0.289	-0.093	0.028
NOhablaesp5	0.300	0.199	1.510	0.132	-0.090	0.690
ruralAurbano1a5	-1.113***	0.262	-4.250	0.000	-1.626	-0.600
urbanoArural1a5	-0.307	0.370	-0.830	0.407	-1.032	0.418
solourbano1a5	-1.032***	0.198	-5.220	0.000	-1.420	-0.645
entype5otros	-0.513	0.437	-1.180	0.240	-1.369	0.343
entype5privado	0.164	0.186	0.880	0.380	-0.201	0.529
stuntdummy2	-0.223	0.144	-1.550	0.122	-0.505	0.059
hhsiz5	0.020	0.033	0.610	0.542	-0.044	0.084
momeduSECU5	-0.119	0.152	-0.780	0.435	-0.416	0.179
wi1	0.032	0.416	0.080	0.939	-0.783	0.846
wi4	-0.773	0.556	-1.390	0.165	-1.862	0.317
wi5	-0.823	0.610	-1.350	0.177	-2.019	0.373
TDA5	-0.977***	0.131	-7.440	0.000	-1.235	-0.720
Constante	1.601***	0.421	3.800	0.000	0.776	2.426

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

Fuente: Boyden (2022a, 2022b, 2022c) y Sanchez et al. (2022). Elaboración propia.

Tabla 43

Regresión logit para el Propensity score con el corte de 25 minutos en la ronda 5 (alumnos de secundaria)

Observaciones = 1,518  
 LR chi2(15) = 211.66  
 Prob > chi2 = 0.0000  
 Pseudo R2 = 0.1400

Log likelihood = -650.16313

TREAT_25MIN5	Coficiente	Error std.	z	P >  z	[Intervalo	Conf. 95%]
chsex	0.041	0.140	0.290	0.771	-0.234	0.316
CDA	-0.018	0.036	-0.500	0.618	-0.088	0.052
NOhablaesp5	0.323	0.210	1.540	0.124	-0.089	0.735
ruralAurbano1a5	-1.352***	0.304	-4.450	0.000	-1.948	-0.756
urbanoArural1a5	-0.418	0.396	-1.060	0.291	-1.193	0.358
solourbano1a5	-1.535***	0.229	-6.710	0.000	-1.984	-1.087
entype5otros	-0.157	0.512	-0.310	0.759	-1.160	0.846
entype5privado	0.160	0.234	0.690	0.493	-0.298	0.618
stuntdummy2	-0.175	0.166	-1.050	0.293	-0.500	0.151
hhsiz5	0.006	0.038	0.160	0.874	-0.069	0.081
momeduSECU5	-0.269	0.188	-1.430	0.153	-0.637	0.100
wi1	0.302	0.507	0.600	0.551	-0.691	1.295
wi4	-1.328**	0.650	-2.040	0.041	-2.603	-0.053
wi5	-0.387	0.718	-0.540	0.590	-1.795	1.020
TDA5	-1.109***	0.159	-6.970	0.000	-1.422	-0.797
Constante	1.295***	0.489	2.650	0.008	0.337	2.254

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

Fuente: Boyden (2022a, 2022b, 2022c) y Sanchez et al. (2022). Elaboración propia.

Tabla 44

Regresión logit para el Propensity score con el corte de 30 minutos en la ronda 5 (alumnos de secundaria)

Observaciones = 1,518  
 LR chi2(15) = 227.79  
 Prob > chi2 = 0.0000  
 Pseudo R2 = 0.1563

Log likelihood = -614.79169

TREAT_30MIN5	Coficiente	Error std.	z	P >  z	[Intervalo	Conf. 95%]
chsex	0.104	0.145	0.720	0.471	-0.180	0.389
CDA	-0.031	0.037	-0.830	0.406	-0.103	0.042
NOhablaesp5	0.383*	0.211	1.810	0.070	-0.031	0.798
ruralAurbano1a5	-1.292***	0.305	-4.230	0.000	-1.891	-0.693
urbanoArural1a5	-0.312	0.397	-0.790	0.432	-1.091	0.467
solourbano1a5	-1.546***	0.234	-6.600	0.000	-2.005	-1.087
entype5otros	0.043	0.518	0.080	0.934	-0.972	1.058
entype5privado	0.268	0.245	1.090	0.274	-0.212	0.749
stuntdummy2	-0.185	0.170	-1.090	0.276	-0.518	0.148
hhsiz5	-0.008	0.040	-0.200	0.842	-0.086	0.070
momeduSECU5	-0.353	0.197	-1.790	0.073	-0.740	0.034
wi1	0.130	0.527	0.250	0.805	-0.903	1.163
wi4	-1.288*	0.667	-1.930	0.053	-2.594	0.019
wi5	-0.563	0.737	-0.760	0.445	-2.007	0.882
TDA5	-1.141***	0.166	-6.890	0.000	-1.466	-0.816
Constante	1.505***	0.504	2.980	0.003	0.516	2.493

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

Fuente: Boyden (2022a, 2022b, 2022c) y Sanchez et al. (2022). Elaboración propia.

Anexo B: Tablas y gráficos de test de medias, reducción de sesgo y soporte común para los alumnos de primaria en la ronda 3 con los cortes de 25 y 30 minutos (5 vecinos más cercanos y caliper 0.005)

Tabla 45

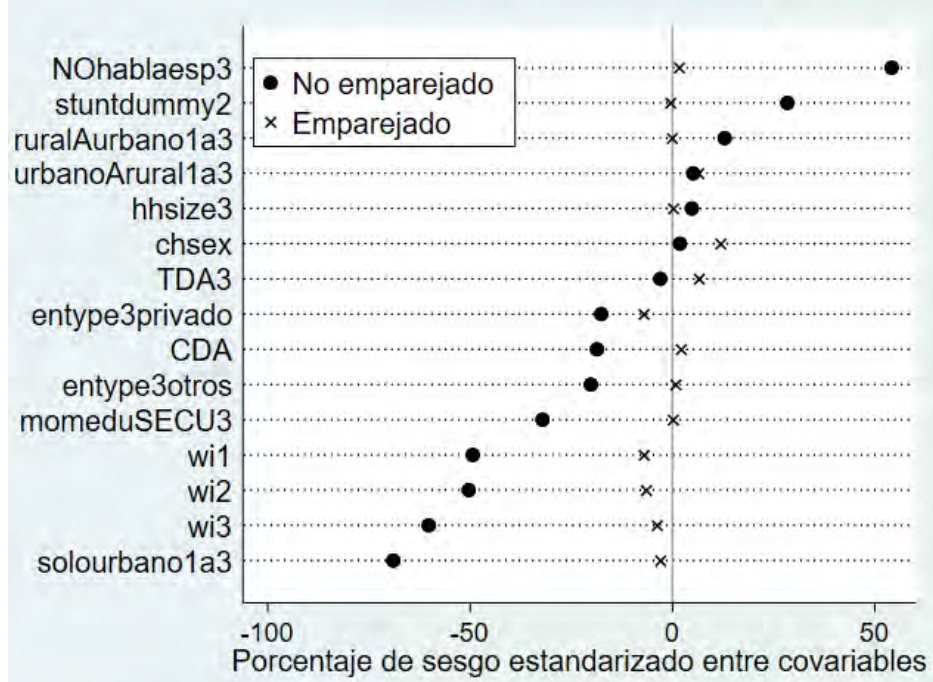
Test de medias de covariables entre alumnos con viajes largos y cortos (control) en la ronda 3 con el punto de corte de 25 minutos (alumnos de primaria)

Variable	Promedio			t-test	
	Viajes Largos	Viajes Cortos (Control)	%sesgo	t	p>t
chsex	0.505	0.446	11.900	1.180	0.238
CDA	8.252	8.208	2.100	0.210	0.834
NOhablaesp3	0.288	0.282	1.600	0.140	0.890
ruralAurbano1a3	0.081	0.081	-0.100	-0.010	0.990
urbanoArural1a3	0.030	0.020	6.400	0.640	0.523
solourbano1a3	0.409	0.423	-3.000	-0.280	0.776
entype3otros	0.005	0.004	0.800	0.150	0.882
entype3privado	0.091	0.114	-7.100	-0.740	0.457
stuntdummy2	0.439	0.442	-0.500	-0.050	0.960
hhsiz3	5.480	5.477	0.200	0.020	0.986
momeduSECU3	0.258	0.257	0.100	0.010	0.991
wi1	0.346	0.363	-7.100	-0.720	0.475
wi2	0.388	0.403	-6.500	-0.650	0.514
wi3	0.458	0.465	-3.900	-0.380	0.701
TDA3	0.742	0.714	6.500	0.640	0.525

Fuente: Boyden (2022a, 2022b, 2022c) y Sanchez et al. (2022). Elaboración propia.

Gráfico 46

Porcentaje de sesgo estandarizado entre las covariables de la ronda 3 con el punto de corte de 25 minutos de los alumnos de viajes cortos al ser emparejados y no emparejados (alumnos de primaria)



Fuente: Boyden (2022a, 2022b, 2022c) y Sanchez et al. (2022). Elaboración propia.

Tabla 46

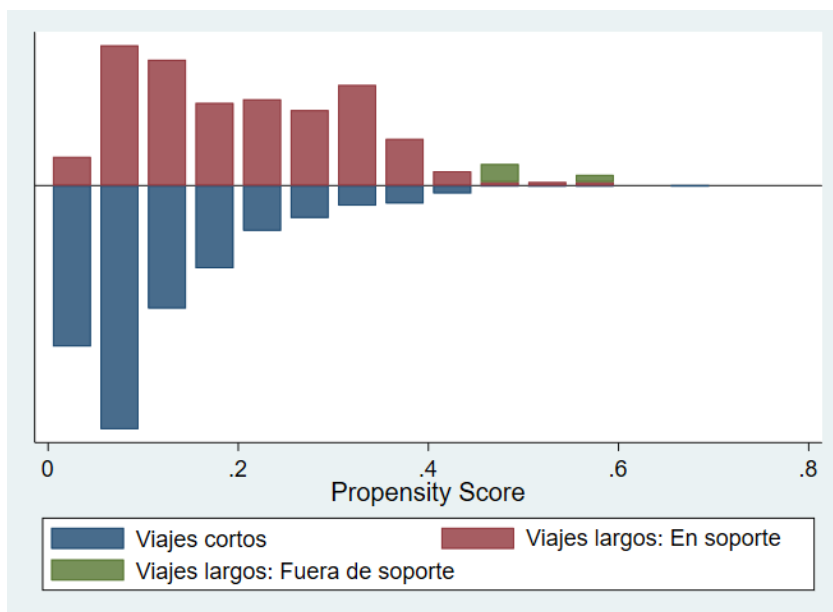
Soporte común en el emparejamiento en la ronda 3 con el punto de corte de 25 minutos (alumnos de primaria)

Asignación	Soporte Común		Total
	Fuera	Adentro	
Viajes cortos (<25 minutos)	0	1319	1319
Viajes largos (>=25 minutos)	7	198	205
Total	7	1517	1524

Fuente: Boyden (2022a, 2022b, 2022c) y Sanchez et al. (2022). Elaboración propia.

Gráfico 47

Soporte común de la ronda 3 con el punto de corte de 25 minutos según el Propensity Score (alumnos de primaria)



Fuente: Boyden (2022a, 2022b, 2022c) y Sanchez et al. (2022).  
Elaboración propia.

Tabla 47

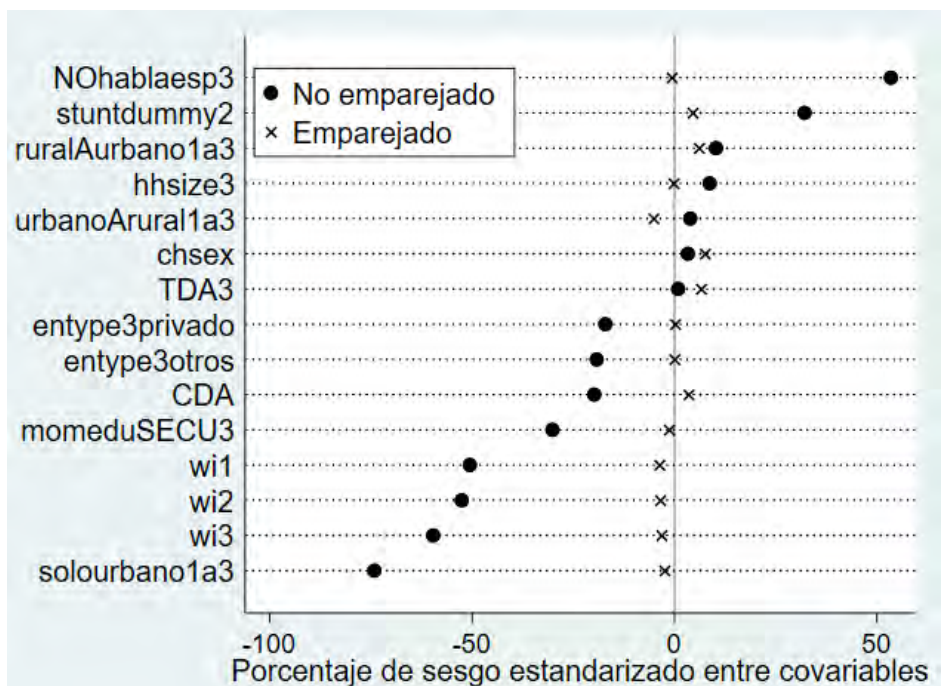
Test de medias de covariables entre alumnos con viajes largos y cortos (control) en la ronda 3 con el punto de corte de 30 minutos (alumnos de primaria)

Variable	Promedio			t-test	
	Viajes Largos	Viajes Cortos (Control)	%sesgo	t	p>t
chsex	0.511	0.474	7.500	0.700	0.484
CDA	8.195	8.121	3.500	0.320	0.750
NOhablaesp3	0.287	0.290	-0.700	-0.050	0.958
ruralAurbano1a3	0.075	0.060	6.100	0.560	0.579
urbanoArural1a3	0.029	0.037	-5.200	-0.420	0.674
solourbano1a3	0.385	0.397	-2.500	-0.220	0.827
entype3otros	0.006	0.006	0.000	0.000	1.000
entype3privado	0.092	0.091	0.200	0.020	0.985
stuntdummy2	0.460	0.438	4.500	0.400	0.687
hhsiz3	5.563	5.566	-0.200	-0.020	0.987
momeduSECU3	0.259	0.265	-1.400	-0.130	0.894
wi1	0.342	0.351	-3.700	-0.360	0.719
wi2	0.381	0.389	-3.500	-0.330	0.740
wi3	0.456	0.462	-3.200	-0.300	0.764
TDA3	0.759	0.730	6.600	0.610	0.544

Fuente: Boyden (2022a, 2022b, 2022c) y Sanchez et al. (2022). Elaboración propia.

Gráfico 48

Porcentaje de sesgo estandarizado entre las covariables de la ronda 3 con el punto de corte de 30 minutos de los alumnos de viajes cortos al ser emparejados y no emparejados (alumnos de primaria)



Fuente: Boyden (2022a, 2022b, 2022c) y Sanchez et al. (2022).  
Elaboración propia.

Tabla 48

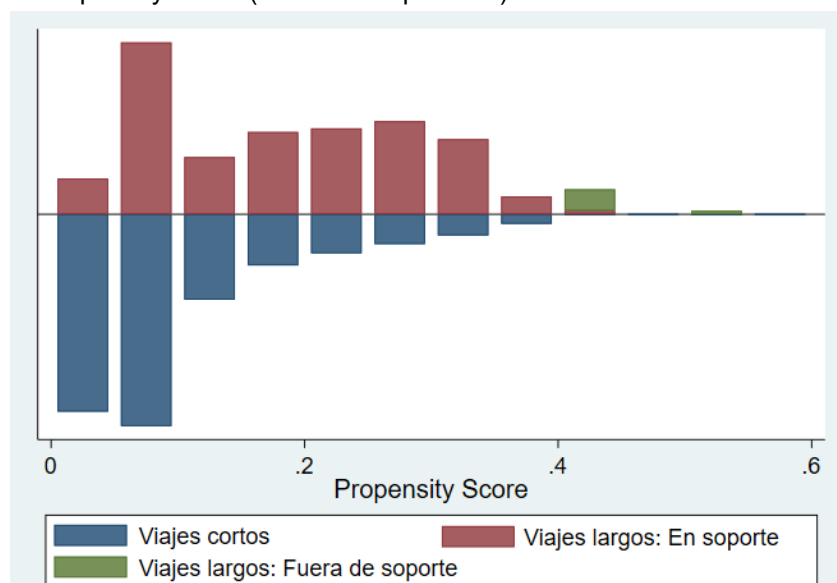
Soporte común en el emparejamiento en la ronda 3 con el punto de corte de 30 minutos (alumnos de primaria)

Asignación	Soporte Común		Total
	Fuera	Adentro	
Viajes cortos (<30 minutos)	0	1343	1343
Viajes largos (>=30 minutos)	7	174	181
Total	7	1517	1524

Fuente: Boyden (2022a, 2022b, 2022c) y Sanchez et al. (2022). Elaboración propia.

Gráfico 49

Soporte común de la ronda 3 con el punto de corte de 30 minutos según el Propensity Score (alumnos de primaria)



Fuente: Boyden (2022a, 2022b, 2022c) y Sanchez et al. (2022).  
Elaboración propia.



Anexo C: Tablas y gráficos de test de medias, reducción de sesgo y soporte común para los alumnos de secundaria en la ronda 5 con los cortes de 25 y 30 minutos (5 vecinos más cercanos y caliper 0.005)

Tabla 49

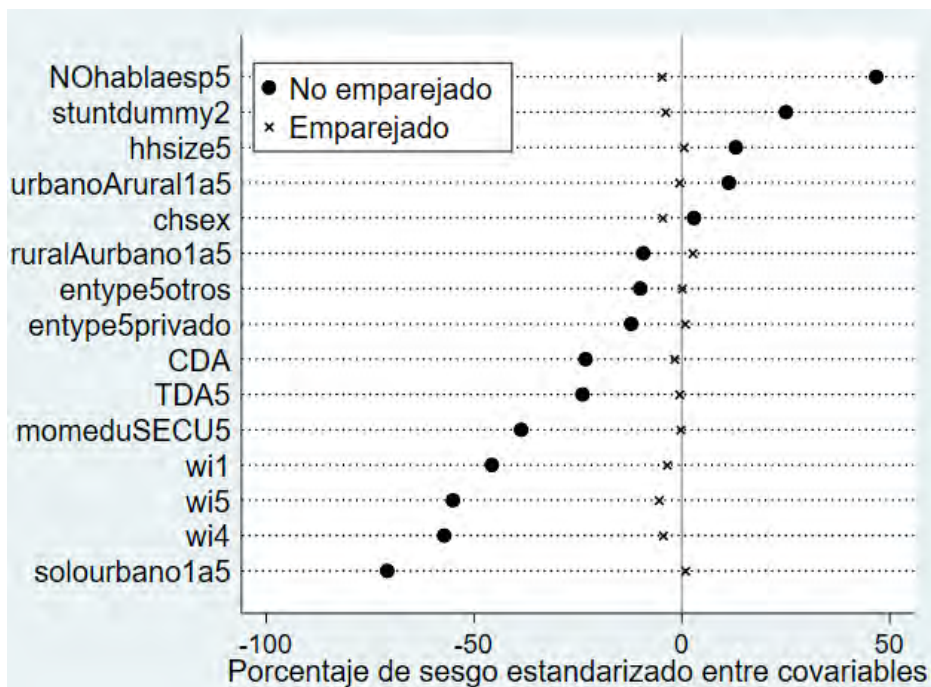
Test de medias de covariables entre alumnos con viajes largos y cortos (control) en la ronda 5 con el punto de corte de 25 minutos (alumnos de secundaria)

Variable	Promedio			t-test	
	Viajes Largos	Viajes Cortos (Control)	%sesgo	t	p>t
chsex	0.503	0.527	-4.700	-0.560	0.578
CDA	8.115	8.153	-1.800	-0.210	0.836
NOhablaesp5	0.238	0.256	-4.800	-0.500	0.617
ruralAurbano1a5	0.063	0.056	2.600	0.330	0.738
urbanoArural1a5	0.042	0.043	-0.600	-0.060	0.951
solourbano1a5	0.430	0.426	0.800	0.090	0.926
entype5otros	0.017	0.017	0.000	0.000	1.000
entype5privado	0.119	0.116	0.800	0.100	0.923
stuntdummy2	0.399	0.417	-4.000	-0.460	0.648
hhsiz5	5.392	5.383	0.400	0.060	0.956
momeduSECU5	0.280	0.281	-0.300	-0.040	0.967
wi1	0.358	0.366	-3.600	-0.430	0.669
wi4	0.521	0.530	-4.600	-0.520	0.601
wi5	0.565	0.576	-5.500	-0.630	0.527
TDA5	0.524	0.528	-0.600	-0.070	0.941

Fuente: Boyden (2022a, 2022b, 2022c) y Sanchez et al. (2022). Elaboración propia.

Gráfico 50

Porcentaje de sesgo estandarizado entre las covariables de la ronda 5 con el punto de corte de 25 minutos de los alumnos de viajes cortos al ser emparejados y no emparejados (alumnos de secundaria)



Fuente: Boyden (2022a, 2022b, 2022c) y Sanchez et al. (2022).  
Elaboración propia.

Tabla 50

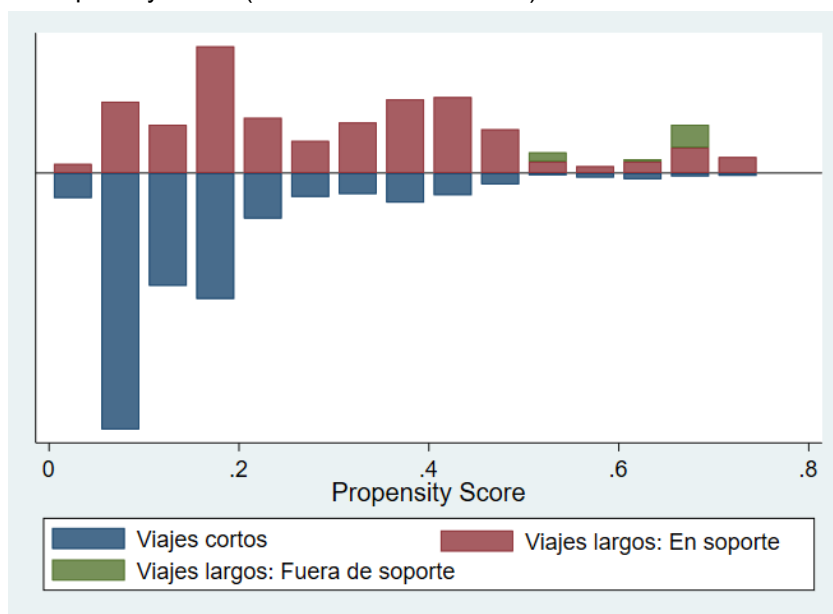
Soporte común en el emparejamiento en la ronda 5 con el punto de corte de 25 minutos (alumnos de secundaria)

Asignación	Soporte Común		Total
	Fuera	Adentro	
Viajes cortos (<25 minutos)	0	1217	1217
Viajes largos (>=25 minutos)	15	286	301
Total	15	1503	1518

Fuente: Boyden (2022a, 2022b, 2022c) y Sanchez et al. (2022). Elaboración propia.

Gráfico 51

Soporte común de la ronda 5 con el punto de corte de 25 minutos según el Propensity Score (alumnos de secundaria)



Fuente: Boyden (2022a, 2022b, 2022c) y Sanchez et al. (2022).  
Elaboración propia.

Tabla 51

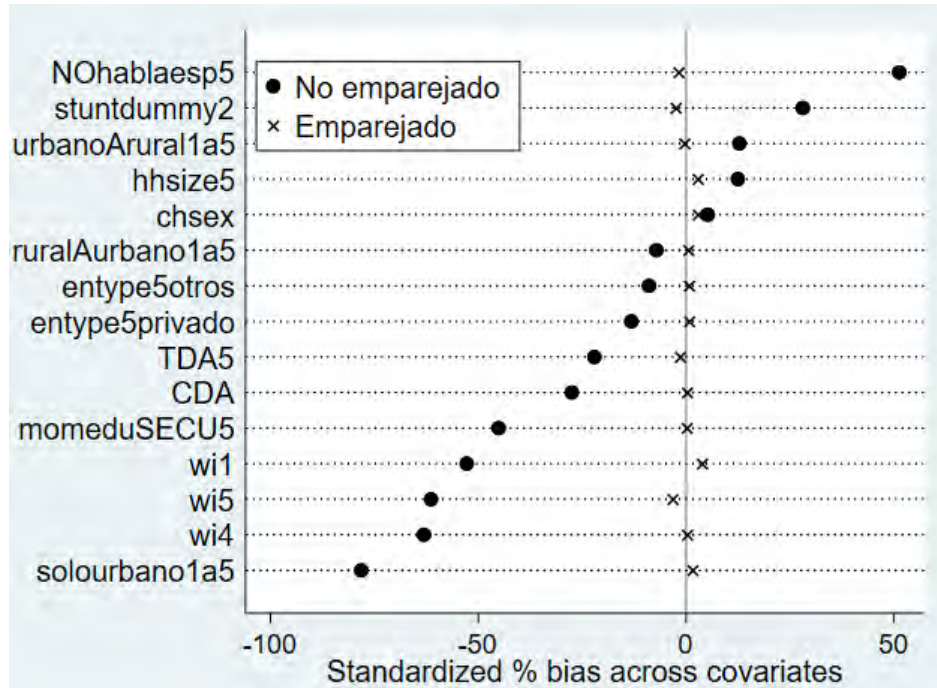
Test de medias de covariables entre alumnos con viajes largos y cortos (control) en la ronda 5 con el punto de corte de 30 minutos (alumnos de secundaria)

Variable	Promedio			t-test	
	Viajes Largos	Viajes Cortos (Control)	%sesgo	t	p>t
chsex	0.524	0.509	2.900	0.340	0.738
CDA	8.044	8.038	0.200	0.030	0.977
NOhablaesp5	0.276	0.283	-1.800	-0.180	0.861
ruralAurbano1a5	0.065	0.064	0.600	0.070	0.945
urbanoArural1a5	0.044	0.044	-0.400	-0.040	0.970
solourbano1a5	0.385	0.378	1.600	0.180	0.861
entype5otros	0.018	0.017	0.700	0.100	0.923
entype5privado	0.113	0.110	0.800	0.090	0.925
stuntdummy2	0.415	0.426	-2.500	-0.280	0.780
hhsiz5	5.400	5.347	2.800	0.330	0.741
momeduSECU5	0.244	0.243	0.200	0.020	0.980
wi1	0.340	0.331	3.900	0.470	0.642
wi4	0.507	0.507	0.300	0.030	0.976
wi5	0.549	0.555	-3.200	-0.370	0.712
TDA5	0.524	0.531	-1.400	-0.160	0.870

Fuente: Boyden (2022a, 2022b, 2022c) y Sanchez et al. (2022). Elaboración propia.

Gráfico 52

Porcentaje de sesgo estandarizado entre las covariables de la ronda 5 con el punto de corte de 30 minutos de los alumnos de viajes cortos al ser emparejados y no emparejados (alumnos de secundaria)



Fuente: Boyden (2022a, 2022b, 2022c) y Sanchez et al. (2022).  
Elaboración propia.

Tabla 52

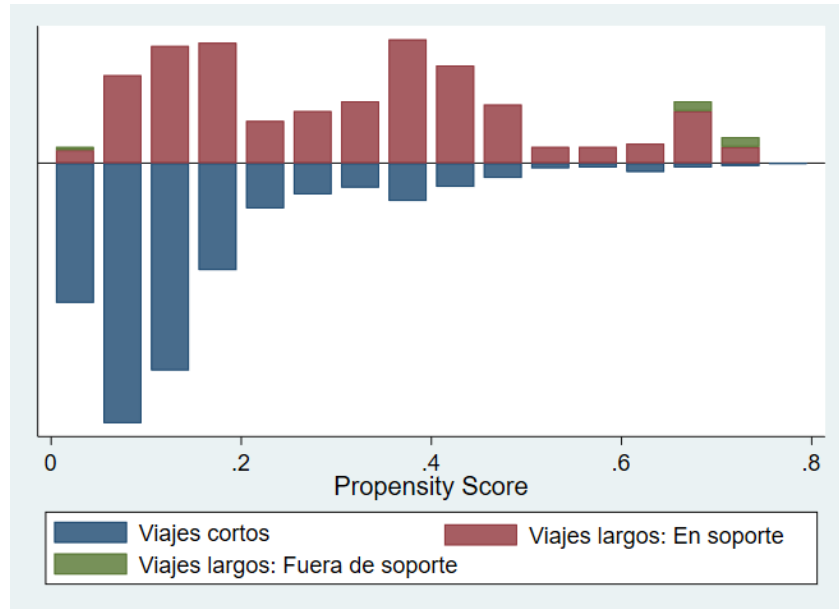
Soporte común en el emparejamiento en la ronda 5 con el punto de corte de 30 minutos (alumnos de secundaria)

Asignación	Soporte Común		Total
	Fuera	Adentro	
Viajes cortos (<30 minutos)	0	1236	1236
Viajes largos (>=30 minutos)	7	275	282
Total	7	1511	1518

Fuente: Boyden (2022a, 2022b, 2022c) y Sanchez et al. (2022). Elaboración propia.

Gráfico 53

Soporte común de la ronda 5 con el punto de corte de 30 minutos según el Propensity Score (alumnos de secundaria)



Fuente: Boyden (2022a, 2022b, 2022c) y Sanchez et al. (2022).  
Elaboración propia.



Anexo D: Tablas de test de medias de covariables entre alumnos con viajes largos y cortos (control) en la ronda 3 con los cortes de 25 y 30 minutos utilizando emparejamientos de 5 vecinos más cercanos (caliper de 0.005) y kernel epanechnikov (alumnos de primaria rurales y urbanos)

Tabla 53

Test de medias de covariables entre alumnos con viajes largos y cortos (control) en la ronda 3 con el punto de corte de 25 minutos, emparejamiento de 5 vecinos más cercanos y caliper de 0.005 (alumnos de primaria rurales)

Variable	Promedio			t-test	
	Viajes Largos	Viajes Cortos (Control)	%sesgo	t	p>t
chsex	0.476	0.462	2.700	0.190	0.848
CDA	7.660	7.371	13.000	0.890	0.374
NOhablaesp3	0.495	0.484	2.200	0.150	0.877
stuntdummy2	0.602	0.629	-5.400	-0.390	0.694
hysize3	5.922	5.957	-1.900	-0.140	0.891
momeduSECU3	0.058	0.047	4.400	0.370	0.709
wi1	0.222	0.218	3.300	0.260	0.797
wi2	0.248	0.252	-2.700	-0.210	0.835
wi3	0.336	0.339	-2.900	-0.230	0.820
TDA3	0.951	0.944	3.400	0.250	0.804

Fuente: Boyden (2022a, 2022b, 2022c) y Sanchez et al. (2022). Elaboración propia.

Tabla 54

Test de medias de covariables entre alumnos con viajes largos y cortos (control) en la ronda 3 con el punto de corte de 25 minutos, emparejamiento de 5 vecinos más cercanos y caliper de 0.005 (alumnos de primaria urbanos)

Variable	Promedio			t-test	
	Viajes Largos	Viajes Cortos (Control)	%sesgo	t	p>t
chsex	0.527	0.536	-1.900	-0.130	0.899
CDA	8.796	8.886	-4.600	-0.310	0.758
NOhablaesp3	0.054	0.033	9.400	0.690	0.489
entype3otros	0.011	0.011	0.000	0.000	1.000
entype3privado	0.183	0.181	0.600	0.040	0.970
stuntdummy2	0.247	0.259	-2.700	-0.180	0.858
hysize3	5.032	5.084	-3.100	-0.210	0.832
momeduSECU3	0.473	0.410	12.500	0.860	0.392
wi1	0.481	0.475	2.600	0.180	0.858
wi2	0.539	0.539	0.000	0.000	0.999
wi3	0.591	0.580	5.600	0.370	0.709
TDA3	0.516	0.528	-2.400	-0.160	0.876

Fuente: Boyden (2022a, 2022b, 2022c) y Sanchez et al. (2022). Elaboración propia.

Tabla 55

Test de medias de covariables entre alumnos con viajes largos y cortos (control) en la ronda 3 con el punto de corte de 30 minutos, emparejamiento de 5 vecinos más cercanos y caliper de 0.005 (alumnos de primaria rurales)

Variable	Promedio			t-test	
	Viajes Largos	Viajes Cortos (Control)	%sesgo	t	p>t
chsex	0.458	0.447	2.300	0.160	0.876
CDA	7.563	7.780	-9.900	-0.690	0.489
NOhablaesp3	0.490	0.507	-3.600	-0.250	0.806
stuntdummy2	0.604	0.583	4.200	0.290	0.770
hhsiz3	5.938	6.000	-3.400	-0.250	0.805
momeduSECU3	0.063	0.053	3.600	0.280	0.782
wi1	0.224	0.226	-1.900	-0.140	0.886
wi2	0.250	0.252	-1.500	-0.120	0.907
wi3	0.337	0.339	-1.000	-0.080	0.937
TDA3	0.958	0.956	0.900	0.070	0.948

Fuente: Boyden (2022a, 2022b, 2022c) y Sanchez et al. (2022). Elaboración propia.

Tabla 56

Test de medias de covariables entre alumnos con viajes largos y cortos (control) en la ronda 3 con el punto de corte de 30 minutos, emparejamiento de 5 vecinos más cercanos y caliper de 0.005 (alumnos de primaria urbanos)

Variable	Promedio			t-test	
	Viajes Largos	Viajes Cortos (Control)	%sesgo	t	p>t
chsex	0.582	0.511	14.400	0.900	0.370
CDA	8.949	9.027	-3.900	-0.250	0.801
NOhablaesp3	0.051	0.018	15.700	1.110	0.269
entype3otros	0.013	0.008	3.200	0.320	0.752
entype3privado	0.203	0.220	-4.500	-0.270	0.787
stuntdummy2	0.278	0.233	10.600	0.650	0.515
hhsiz3	5.152	5.094	3.400	0.200	0.841
momeduSECU3	0.506	0.513	-1.300	-0.080	0.937
wi1	0.488	0.484	1.400	0.090	0.926
wi2	0.540	0.552	-5.700	-0.360	0.719
wi3	0.599	0.595	1.800	0.120	0.908
TDA3	0.506	0.476	6.300	0.380	0.705

Fuente: Boyden (2022a, 2022b, 2022c) y Sanchez et al. (2022). Elaboración propia.

Tabla 57

Test de medias de covariables entre alumnos con viajes largos y cortos (control) en la ronda 3 con el punto de corte de 25 minutos y emparejamiento de kernel epanechnikov (alumnos de primaria rurales)

Variable	Promedio			t-test	
	Viajes Largos	Viajes Cortos (Control)	%sesgo	t	p>t
chsex	0.467	0.460	1.500	0.110	0.911
CDA	7.495	7.550	-2.500	-0.180	0.859
NOhablaesp3	0.514	0.492	4.500	0.320	0.750
stuntdummy2	0.598	0.613	-3.000	-0.220	0.825
hhsiz3	5.897	5.914	-0.900	-0.070	0.947
momeduSECU3	0.056	0.055	0.300	0.020	0.982
wi1	0.219	0.218	1.400	0.110	0.916
wi2	0.247	0.245	1.500	0.120	0.905
wi3	0.331	0.333	-1.100	-0.090	0.931
TDA3	0.944	0.944	0.000	0.000	0.997

Fuente: Boyden (2022a, 2022b, 2022c) y Sanchez et al. (2022). Elaboración propia.

Tabla 58

Test de medias de covariables entre alumnos con viajes largos y cortos (control) en la ronda 3 con el punto de corte de 25 minutos y emparejamiento de kernel epanechnikov (alumnos de primaria urbanos)

Variable	Promedio			t-test	
	Viajes Largos	Viajes Cortos (Control)	%sesgo	t	p>t
chsex	0.546	0.541	1.100	0.080	0.937
CDA	8.876	8.873	0.200	0.010	0.990
NOhablaesp3	0.072	0.071	0.500	0.030	0.978
entype3otros	0.010	0.015	-2.700	-0.270	0.789
entype3privado	0.186	0.174	3.000	0.210	0.834
stuntdummy2	0.247	0.245	0.500	0.030	0.975
hhsiz3	5.010	5.064	-3.200	-0.230	0.819
momeduSECU3	0.464	0.457	1.400	0.100	0.921
wi1	0.474	0.474	-0.100	-0.010	0.992
wi2	0.530	0.532	-1.000	-0.070	0.945
wi3	0.578	0.584	-3.200	-0.220	0.830
TDA3	0.515	0.583	-14.000	-0.940	0.348

Fuente: Boyden (2022a, 2022b, 2022c) y Sanchez et al. (2022). Elaboración propia.

Tabla 59

Test de medias de covariables entre alumnos con viajes largos y cortos (control) en la ronda 3 con el punto de corte de 30 minutos y emparejamiento de kernel epanechnikov (alumnos de primaria rurales)

Variable	Promedio			t-test	
	Viajes Largos	Viajes Cortos (Control)	%sesgo	t	p>t
chsex	0.455	0.455	0.000	0.000	0.999
CDA	7.515	7.563	-2.200	-0.150	0.879
NOhablaesp3	0.505	0.482	4.600	0.320	0.751
stuntdummy2	0.596	0.599	-0.700	-0.050	0.961
hhsiz3	5.889	5.902	-0.700	-0.050	0.959
momeduSECU3	0.061	0.061	-0.300	-0.030	0.980
wi1	0.220	0.217	1.800	0.140	0.889
wi2	0.248	0.246	1.900	0.140	0.886
wi3	0.331	0.331	-0.200	-0.010	0.990
TDA3	0.960	0.962	-1.000	-0.080	0.938

Fuente: Boyden (2022a, 2022b, 2022c) y Sanchez et al. (2022). Elaboración propia.

Tabla 60

Test de medias de covariables entre alumnos con viajes largos y cortos (control) en la ronda 3 con el punto de corte de 30 minutos y emparejamiento de kernel epanechnikov (alumnos de primaria urbanos)

Variable	Promedio			t-test	
	Viajes Largos	Viajes Cortos (Control)	%sesgo	t	p>t
chsex	0.588	0.557	6.100	0.390	0.700
CDA	8.925	8.866	3.000	0.190	0.852
NOhablaesp3	0.050	0.049	0.300	0.020	0.986
entype3otros	0.013	0.018	-3.800	-0.300	0.761
entype3privado	0.200	0.190	2.600	0.160	0.873
stuntdummy2	0.275	0.253	5.100	0.310	0.756
hhsiz3	5.125	5.133	-0.500	-0.030	0.977
momeduSECU3	0.500	0.484	3.200	0.200	0.839
wi1	0.482	0.488	-2.600	-0.160	0.870
wi2	0.534	0.543	-4.000	-0.250	0.805
wi3	0.593	0.603	-5.200	-0.320	0.749
TDA3	0.500	0.572	-14.800	-0.910	0.366

Fuente: Boyden (2022a, 2022b, 2022c) y Sanchez et al. (2022). Elaboración propia.

Anexo E: Tablas de test de medias de covariables entre alumnos con viajes largos y cortos (control) en la ronda 5 con los cortes de 25 y 30 minutos utilizando emparejamientos de 5 vecinos más cercanos (caliper de 0.005) y kernel epanechnikov (alumnos de secundaria rurales y urbanos)

Tabla 61

Test de medias de covariables entre alumnos con viajes largos y cortos (control) en la ronda 5 con el punto de corte de 25 minutos, emparejamiento de 5 vecinos más cercanos y caliper de 0.005 (alumnos de secundaria rurales)

Variable	Promedio			t-test	
	Viajes Largos	Viajes Cortos (Control)	%sesgo	t	p>t
chsex	0.524	0.535	-2.200	-0.190	0.853
CDA	7.490	7.700	-9.600	-0.810	0.418
NOhablaesp5	0.434	0.483	-9.900	-0.820	0.413
entype5otros	0.014	0.022	-7.500	-0.500	0.615
entype5privado	0.014	0.021	-5.000	-0.490	0.625
stuntdummy2	0.600	0.583	3.400	0.290	0.771
hhsiz5	5.710	5.497	11.100	0.960	0.340
momeduSECU5	0.083	0.059	7.600	0.800	0.422
wi1	0.214	0.214	-0.100	-0.010	0.990
wi4	0.381	0.391	-7.300	-0.690	0.490
wi5	0.420	0.425	-3.900	-0.350	0.726
TDA5	0.738	0.731	1.500	0.130	0.900

Fuente: Boyden (2022a, 2022b, 2022c) y Sanchez et al. (2022). Elaboración propia.

Tabla 62

Test de medias de covariables entre alumnos con viajes largos y cortos (control) en la ronda 5 con el punto de corte de 25 minutos, emparejamiento de 5 vecinos más cercanos y caliper de 0.005 (alumnos de secundaria urbanos)

Variable	Promedio			t-test	
	Viajes Largos	Viajes Cortos (Control)	%sesgo	t	p>t
chsex	0.482	0.445	7.500	0.620	0.533
CDA	8.741	8.720	1.000	0.090	0.930
NOhablaesp5	0.050	0.045	2.700	0.220	0.822
entype5otros	0.022	0.016	3.400	0.350	0.724
entype5privado	0.223	0.248	-6.300	-0.490	0.622
stuntdummy2	0.209	0.205	0.900	0.080	0.939
hhsiz5	5.115	5.231	-6.200	-0.510	0.613
momeduSECU5	0.468	0.488	-4.000	-0.340	0.738
wi1	0.506	0.499	3.100	0.260	0.798
wi4	0.669	0.664	3.000	0.260	0.798
wi5	0.716	0.719	-1.900	-0.170	0.867
TDA5	0.302	0.309	-1.500	-0.130	0.897

Fuente: Boyden (2022a, 2022b, 2022c) y Sanchez et al. (2022). Elaboración propia.

Tabla 63

Test de medias de covariables entre alumnos con viajes largos y cortos (control) en la ronda 5 con el punto de corte de 30 minutos, emparejamiento de 5 vecinos más cercanos y caliper de 0.005 (alumnos de secundaria rurales)

Variable	Promedio			t-test	
	Viajes Largos	Viajes Cortos (Control)	%sesgo	t	p>t
chsex	0.511	0.511	-0.100	-0.010	0.994
CDA	7.482	7.750	-12.200	-1.020	0.308
NOhablaesp5	0.440	0.441	-0.200	-0.020	0.984
entype5otros	0.014	0.011	3.300	0.260	0.796
entype5privado	0.014	0.017	-1.600	-0.170	0.866
stuntdummy2	0.596	0.570	5.200	0.440	0.664
hhsiz5	5.745	5.639	5.500	0.450	0.654
momeduSECU5	0.085	0.100	-4.600	-0.420	0.672
wi1	0.217	0.210	5.500	0.490	0.624
wi4	0.376	0.393	-12.400	-1.140	0.255
wi5	0.417	0.424	-4.900	-0.470	0.642
TDA5	0.752	0.701	11.500	0.950	0.343

Fuente: Boyden (2022a, 2022b, 2022c) y Sanchez et al. (2022). Elaboración propia.

Tabla 64

Test de medias de covariables entre alumnos con viajes largos y cortos (control) en la ronda 5 con el punto de corte de 30 minutos, emparejamiento de 5 vecinos más cercanos y caliper de 0.005 (alumnos de secundaria urbanos)

Variable	Promedio			t-test	
	Viajes Largos	Viajes Cortos (Control)	%sesgo	t	p>t
chsex	0.500	0.530	-6.000	-0.460	0.644
CDA	8.608	8.682	-3.700	-0.290	0.775
NOhablaesp5	0.042	0.038	1.500	0.130	0.896
entype5otros	0.025	0.027	-1.000	-0.080	0.935
entype5privado	0.225	0.255	-7.500	-0.540	0.588
stuntdummy2	0.200	0.198	0.400	0.030	0.974
hhsiz5	5.092	5.165	-3.900	-0.290	0.772
momeduSECU5	0.442	0.420	4.400	0.340	0.736
wi1	0.497	0.485	5.300	0.400	0.689
wi4	0.669	0.664	3.100	0.250	0.803
wi5	0.713	0.708	3.400	0.270	0.790
TDA5	0.292	0.297	-1.100	-0.080	0.933

Fuente: Boyden (2022a, 2022b, 2022c) y Sanchez et al. (2022). Elaboración propia.

Tabla 65

Test de medias de covariables entre alumnos con viajes largos y cortos (control) en la ronda 5 con el punto de corte de 25 minutos y emparejamiento de kernel epanechnikov (alumnos de secundaria rurales)

Variable	Promedio			t-test	
	Viajes Largos	Viajes Cortos (Control)	%sesgo	t	p>t
chsex	0.535	0.551	-3.400	-0.300	0.765
CDA	7.572	7.591	-0.900	-0.080	0.937
NOhablaesp5	0.453	0.442	2.200	0.190	0.848
entype5otros	0.013	0.008	4.100	0.370	0.713
entype5privado	0.019	0.018	0.800	0.080	0.938
stuntdummy2	0.579	0.597	-3.600	-0.320	0.746
hhsiz5	5.686	5.659	1.400	0.120	0.903
momeduSECU5	0.094	0.097	-0.700	-0.070	0.947
wi1	0.222	0.213	6.500	0.610	0.545
wi4	0.383	0.386	-2.600	-0.240	0.812
wi5	0.420	0.421	-0.700	-0.060	0.950
TDA5	0.698	0.700	-0.400	-0.030	0.973

Fuente: Boyden (2022a, 2022b, 2022c) y Sanchez et al. (2022). Elaboración propia.

Tabla 66

Test de medias de covariables entre alumnos con viajes largos y cortos (control) en la ronda 5 con el punto de corte de 25 minutos y emparejamiento de kernel epanechnikov (alumnos de secundaria urbanos)

Variable	Promedio			t-test	
	Viajes Largos	Viajes Cortos (Control)	%sesgo	t	p>t
chsex	0.482	0.493	-2.200	-0.180	0.854
CDA	8.723	8.805	-4.100	-0.340	0.733
NOhablaesp5	0.057	0.049	3.900	0.310	0.760
entype5otros	0.021	0.020	0.800	0.080	0.934
entype5privado	0.220	0.224	-1.000	-0.080	0.935
stuntdummy2	0.220	0.213	1.700	0.140	0.887
hhsiz5	5.135	5.131	0.200	0.020	0.988
momeduSECU5	0.461	0.488	-5.400	-0.450	0.650
wi1	0.502	0.510	-3.400	-0.280	0.781
wi4	0.665	0.673	-5.200	-0.430	0.665
wi5	0.715	0.717	-1.400	-0.120	0.902
TDA5	0.298	0.305	-1.600	-0.140	0.891

Fuente: Boyden (2022a, 2022b, 2022c) y Sanchez et al. (2022). Elaboración propia.

Tabla 67

Test de medias de covariables entre alumnos con viajes largos y cortos (control) en la ronda 5 con el punto de corte de 30 minutos y emparejamiento de kernel epanechnikov (alumnos de secundaria rurales)

Variable	Promedio			t-test	
	Viajes Largos	Viajes Cortos (Control)	%sesgo	t	p>t
chsex	0.538	0.554	-3.200	-0.280	0.538
CDA	7.571	7.585	-0.600	-0.060	7.571
NOhablaesp5	0.462	0.464	-0.500	-0.040	0.462
entype5otros	0.013	0.008	4.500	0.400	0.013
entype5privado	0.019	0.017	1.200	0.110	0.019
stuntdummy2	0.577	0.597	-4.100	-0.360	0.577
hhsiz5	5.667	5.717	-2.600	-0.230	5.667
momeduSECU5	0.096	0.097	-0.300	-0.030	0.096
wi1	0.223	0.215	5.900	0.540	0.223
wi4	0.381	0.385	-3.200	-0.300	0.381
wi5	0.418	0.419	-0.700	-0.060	0.418
TDA5	0.705	0.703	0.600	0.050	0.705

Fuente: Boyden (2022a, 2022b, 2022c) y Sanchez et al. (2022). Elaboración propia.

Tabla 68

Test de medias de covariables entre alumnos con viajes largos y cortos (control) en la ronda 5 con el punto de corte de 30 minutos y emparejamiento de kernel epanechnikov (alumnos de secundaria urbanos)

Variable	Promedio			t-test	
	Viajes Largos	Viajes Cortos (Control)	%sesgo	t	p>t
chsex	0.500	0.506	-1.200	-0.090	0.925
CDA	8.637	8.702	-3.200	-0.250	0.800
NOhablaesp5	0.056	0.052	2.100	0.160	0.876
entype5otros	0.024	0.026	-0.900	-0.070	0.941
entype5privado	0.226	0.224	0.500	0.040	0.969
stuntdummy2	0.226	0.228	-0.500	-0.040	0.969
hhsiz5	5.121	5.124	-0.100	-0.010	0.991
momeduSECU5	0.427	0.443	-3.200	-0.250	0.801
wi1	0.489	0.493	-1.800	-0.140	0.889
wi4	0.660	0.668	-5.100	-0.400	0.691
wi5	0.712	0.713	-1.000	-0.080	0.934
TDA5	0.282	0.305	-4.700	-0.380	0.701

Fuente: Boyden (2022a, 2022b, 2022c) y Sanchez et al. (2022). Elaboración propia.

Anexo F: Tablas de estimaciones del ATT sin bootstrapping en la ronda 3 y 5 con emparejamientos de 5 vecinos más cercanos (caliper de 0.005) y kernel epanechnikov (alumnos de primaria y secundaria según zona de residencia)

Tabla 69

Estimación del ATT sin bootstrapping en la ronda 3 con emparejamiento de 5 vecinos más cercanos y caliper de 0.005 (alumnos de primaria rurales y urbanos)

ATT (5 vecinos, caliper 0.005, ronda 3)		
Corte de tiempo	Rural	Urbano
>= 20 minutos	0.138	0.036
>= 25 minutos	0.305**	0.110
>= 30 minutos	0.352***	-0.028

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

Fuente: Boyden (2022a, 2022b, 2022c) y Sanchez et al. (2022).  
Elaboración propia.

Tabla 70

Estimación del ATT sin bootstrapping en la ronda 3 con emparejamiento de kernel epanechnikov (alumnos de primaria rurales y urbanos)

ATT (Kernel epanechnikov, ronda 3)		
Corte de tiempo	Rural	Urbano
>= 20 minutos	0.102	0.001
>= 25 minutos	0.252**	0.074
>= 30 minutos	0.318***	-0.000

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

Fuente: Boyden (2022a, 2022b, 2022c) y Sanchez et al. (2022).  
Elaboración propia.

Tabla 71

Estimación del ATT sin bootstrapping en la ronda 5 con emparejamiento de 5 vecinos más cercanos y caliper de 0.005 (alumnos de secundaria rurales y urbanos)

ATT (5 vecinos, caliper 0.005, ronda 5)		
Corte de tiempo	Rural	Urbano
>= 20 minutos	0.008	-0.135*
>= 25 minutos	0.002	-0.169*
>= 30 minutos	0.039	-0.022

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

Fuente: Boyden (2022a, 2022b, 2022c) y Sanchez et al. (2022).  
Elaboración propia.

Tabla 72

Estimación del ATT sin bootstrapping en la ronda 5 con emparejamiento de kernel epanechnikov (alumnos de secundaria rurales y urbanos)

ATT (Kernel epanechnikov, ronda 5)		
Corte de tiempo	Rural	Urbano
>= 20 minutos	0.039	-0.149**
>= 25 minutos	0.084	-0.141
>= 30 minutos	0.068	-0.128

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

Fuente: Boyden (2022a, 2022b, 2022c) y Sanchez et al. (2022).  
Elaboración propia.

Anexo G: Tablas de test de medias de covariables entre alumnos con viajes largos y cortos (control) en las rondas 3 y 5 con los puntos de corte de 25 y 30 minutos, emparejamiento de 5 vecinos más cercanos (caliper de 0.005) y dividiendo los alumnos con viajes largos a partir del intervalo de 40 minutos (alumnos de primaria y secundaria)

Tabla 73

Test de medias de covariables entre alumnos con viajes largos menores a 40 minutos y viajes cortos (control) en la ronda 3 con el punto de corte de 25 minutos, emparejamiento de 5 vecinos más cercanos y caliper de 0.005 (alumnos de primaria)

Variable	Promedio			t-test	
	Viajes Largos <40 min	Viajes Cortos (Control)	%sesgo	t	p>t
chsex	0.504	0.459	8.800	0.730	0.463
CDA	8.338	8.346	-0.400	-0.030	0.976
NOhablaesp3	0.245	0.290	-11.900	-0.850	0.397
ruralAurbano1a3	0.094	0.104	-4.000	-0.300	0.764
urbanoArural1a3	0.014	0.017	-2.200	-0.190	0.848
solourbano01a3	0.475	0.469	1.200	0.100	0.924
entype3privado	0.108	0.109	-0.400	-0.040	0.969
stuntdummy2	0.417	0.451	-7.000	-0.560	0.574
hhsiz3	5.482	5.591	-6.000	-0.510	0.614
momeduSECU3	0.266	0.259	1.600	0.140	0.892
wi1	0.355	0.352	1.600	0.130	0.893
wi2	0.407	0.405	0.700	0.060	0.955
wi3	0.471	0.476	-2.300	-0.190	0.846
TDA3	0.719	0.734	-3.300	-0.280	0.782

Fuente: Boyden (2022a, 2022b, 2022c) y Sanchez et al. (2022). Elaboración propia.

Tabla 74

Test de medias de covariables entre alumnos con viajes largos mayores o iguales a 40 minutos y viajes cortos (control) en la ronda 3 con el punto de corte de 25 minutos, emparejamiento de 5 vecinos más cercanos y caliper de 0.005 (alumnos de primaria)

Variable	Promedio			t-test	
	Viajes Largos >= 40 min	Viajes Cortos (Control)	%sesgo	t	p>t
chsex	0.483	0.510	-5.400	-0.290	0.770
CDA	7.800	7.706	4.300	0.210	0.830
NOhablaesp3	0.400	0.384	4.000	0.180	0.856
ruralAurbano1a3	0.033	0.040	-3.300	-0.190	0.848
urbanoArural1a3	0.067	0.046	9.900	0.480	0.634
solourbano1a3	0.250	0.260	-2.300	-0.120	0.901
entype3otros	0.017	0.027	-6.500	-0.370	0.710
entype3privado	0.050	0.063	-4.600	-0.310	0.755
stuntdummy2	0.467	0.449	3.800	0.200	0.844
hhsiz3	5.550	5.520	1.700	0.090	0.927
momeduSECU3	0.233	0.270	-8.000	-0.460	0.647
wi1	0.319	0.336	-7.400	-0.410	0.681
wi2	0.338	0.346	-3.500	-0.190	0.846
wi3	0.413	0.436	-10.900	-0.610	0.543
TDA3	0.817	0.770	11.400	0.630	0.532

Fuente: Boyden (2022a, 2022b, 2022c) y Sanchez et al. (2022). Elaboración propia.

Tabla 75

Test de medias de covariables entre alumnos con viajes largos menores a 40 minutos y viajes cortos (control) en la ronda 3 con el punto de corte de 30 minutos, emparejamiento de 5 vecinos más cercanos y caliper de 0.005 (alumnos de primaria)

Variable	Promedio			t-test	
	Viajes Largos < 40 min	Viajes Cortos (Control)	%sesgo	t	p>t
chsex	0.534	0.561	-5.400	-0.420	0.677
CDA	8.322	8.347	-1.200	-0.090	0.929
NOhablaesp3	0.254	0.275	-5.400	-0.350	0.725
ruralAurbano1a3	0.093	0.063	11.300	0.850	0.394
urbanoArural1a3	0.008	0.005	2.800	0.320	0.752
solourbano1a3	0.441	0.463	-4.600	-0.340	0.735
entype3privado	0.110	0.098	3.700	0.310	0.756
stuntdummy2	0.441	0.424	3.500	0.260	0.794
hhsiz3	5.559	5.578	-1.000	-0.080	0.939
momeduSECU3	0.263	0.263	-0.100	-0.010	0.992
wi1	0.347	0.359	-5.700	-0.450	0.655
wi2	0.394	0.408	-6.100	-0.470	0.637
wi3	0.467	0.483	-7.600	-0.580	0.565
TDA3	0.729	0.683	10.300	0.770	0.443

Fuente: Boyden (2022a, 2022b, 2022c) y Sanchez et al. (2022). Elaboración propia.

Tabla 76

Test de medias de covariables entre alumnos con viajes largos mayores o iguales a 40 minutos y viajes cortos (control) en la ronda 3 con el punto de corte de 30 minutos, emparejamiento de 5 vecinos más cercanos y caliper de 0.005 (alumnos de primaria)

Variable	Promedio			t-test	
	Viajes Largos >= 40 min	Viajes Cortos (Control)	%sesgo	t	p>t
chsex	0.483	0.496	-2.400	-0.130	0.895
CDA	7.800	7.901	-4.600	-0.240	0.808
NOhablaesp3	0.400	0.390	2.400	0.110	0.912
ruralAurbano1a3	0.033	0.030	1.600	0.100	0.918
urbanoArural1a3	0.067	0.053	6.500	0.310	0.761
solourbano1a3	0.250	0.273	-5.300	-0.290	0.774
entype3otros	0.017	0.023	-4.400	-0.260	0.796
entype3privado	0.050	0.050	0.000	0.000	1.000
stuntdummy2	0.467	0.473	-1.300	-0.070	0.945
hysize3	5.550	5.497	3.000	0.160	0.871
momeduSECU3	0.233	0.253	-4.400	-0.250	0.801
wi1	0.319	0.326	-3.000	-0.170	0.863
wi2	0.338	0.340	-1.200	-0.070	0.946
wi3	0.413	0.426	-6.000	-0.330	0.740
TDA3	0.817	0.738	19.200	1.030	0.303

Fuente: Boyden (2022a, 2022b, 2022c) y Sanchez et al. (2022). Elaboración propia.

Tabla 77

Test de medias de covariables entre alumnos con viajes largos menores a 40 minutos y viajes cortos (control) en la ronda 5 con el punto de corte de 25 minutos, emparejamiento de 5 vecinos más cercanos y caliper de 0.005 (alumnos de secundaria)

Variable	Promedio			t-test	
	Viajes Largos < 40 min	Viajes Cortos (Control)	%sesgo	t	p>t
chsex	0.483	0.449	6.800	0.640	0.524
CDA	8.171	8.160	0.500	0.050	0.963
NOhablaesp3	0.233	0.277	-11.900	-0.940	0.350
ruralAurbano1a3	0.057	0.067	-4.000	-0.400	0.692
urbanoArural1a3	0.040	0.028	6.700	0.590	0.558
solourbano1a3	0.489	0.444	9.400	0.830	0.406
entype3otros	0.011	0.014	-1.600	-0.190	0.848
entype3privado	0.159	0.133	7.100	0.680	0.494
stuntdummy2	0.392	0.418	-5.500	-0.490	0.622
hysize3	5.318	5.259	3.100	0.300	0.761
momeduSECU3	0.284	0.252	6.700	0.670	0.502
wi1	0.385	0.359	10.900	1.030	0.304
wi2	0.545	0.539	2.900	0.260	0.792
wi3	0.578	0.568	5.600	0.500	0.619
TDA3	0.483	0.545	-12.600	-1.160	0.246

Fuente: Boyden (2022a, 2022b, 2022c) y Sanchez et al. (2022). Elaboración propia.

Tabla 78

Test de medias de covariables entre alumnos con viajes largos mayores o iguales a 40 minutos y viajes cortos (control) en la ronda 5 con el punto de corte de 25 minutos, emparejamiento de 5 vecinos más cercanos y caliper de 0.005 (alumnos de secundaria)

Variable	Promedio			t-test	
	Viajes Largos ≥ 40 min	Viajes Cortos (Control)	%sesgo	t	p>t
chsex	0.553	0.548	1.000	0.080	0.938
CDA	8.105	8.199	-4.400	-0.320	0.748
NOhablaesp3	0.263	0.251	3.100	0.210	0.837
ruralAurbano1a3	0.070	0.072	-0.700	-0.050	0.959
urbanoArural1a3	0.044	0.053	-5.500	-0.330	0.740
solourbano1a3	0.325	0.337	-2.700	-0.200	0.845
entype3otros	0.026	0.023	2.100	0.170	0.865
entype3privado	0.053	0.038	5.000	0.540	0.590
stuntdummy2	0.412	0.415	-0.500	-0.040	0.972
hhsiz3	5.500	5.469	1.600	0.120	0.902
momeduSECU3	0.263	0.258	1.100	0.090	0.928
wi1	0.319	0.327	-3.600	-0.280	0.776
wi2	0.483	0.492	-4.400	-0.330	0.740
wi3	0.538	0.543	-3.100	-0.230	0.820
TDA3	0.579	0.541	7.700	0.570	0.567

Fuente: Boyden (2022a, 2022b, 2022c) y Sanchez et al. (2022). Elaboración propia.

Tabla 79

Test de medias de covariables entre alumnos con viajes largos menores a 40 minutos y viajes cortos (control) en la ronda 5 con el punto de corte de 30 minutos, emparejamiento de 5 vecinos más cercanos y caliper de 0.005 (alumnos de secundaria)

Variable	Promedio			t-test	
	Viajes Largos < 40 min	Viajes Cortos (Control)	%sesgo	t	p>t
chsex	0.490	0.479	2.400	0.210	0.834
CDA	7.943	7.881	2.900	0.250	0.807
NOhablaesp3	0.261	0.223	10.200	0.790	0.428
ruralAurbano1a3	0.064	0.056	2.900	0.280	0.776
urbanoArural1a3	0.045	0.049	-2.500	-0.190	0.852
solourbano1a3	0.446	0.423	4.900	0.410	0.683
entype3otros	0.013	0.024	-7.800	-0.750	0.452
entype3privado	0.153	0.132	5.900	0.540	0.592
stuntdummy2	0.414	0.421	-1.400	-0.120	0.903
hhsiz3	5.325	5.426	-5.500	-0.470	0.641
momeduSECU3	0.242	0.231	2.300	0.220	0.826
wi1	0.362	0.357	2.200	0.190	0.848
wi2	0.529	0.524	2.400	0.210	0.836
wi3	0.561	0.559	1.400	0.110	0.909
TDA3	0.503	0.512	-1.800	-0.160	0.875

Fuente: Boyden (2022a, 2022b, 2022c) y Sanchez et al. (2022). Elaboración propia.

Tabla 80

Test de medias de covariables entre alumnos con viajes largos mayores o iguales a 40 minutos y viajes cortos (control) en la ronda 5 con el punto de corte de 30 minutos, emparejamiento de 5 vecinos más cercanos y caliper de 0.005 (alumnos de secundaria)

Variable	Promedio			t-test	
	Viajes Largos ≥ 40 min	Viajes Cortos (Control)	%sesgo	t	p>t
chsex	0.559	0.524	7.000	0.540	0.591
CDA	8.110	8.210	-4.700	-0.350	0.723
NOhablaesp3	0.288	0.275	3.500	0.230	0.818
ruralAurbano1a3	0.068	0.071	-1.200	-0.100	0.922
urbanoArural1a3	0.042	0.047	-2.500	-0.160	0.875
solourbano1a3	0.314	0.329	-3.400	-0.250	0.803
entype3otros	0.025	0.030	-2.700	-0.210	0.833
entype3privado	0.051	0.045	2.000	0.210	0.832
stuntdummy2	0.415	0.418	-0.600	-0.040	0.965
hhsiz3	5.542	5.632	-4.700	-0.340	0.735
momeduSECU3	0.254	0.257	-0.500	-0.040	0.966
wi1	0.316	0.334	-8.100	-0.640	0.521
wi2	0.482	0.495	-6.700	-0.520	0.605
wi3	0.539	0.542	-1.800	-0.140	0.890
TDA3	0.559	0.502	11.600	0.870	0.383

Fuente: Boyden (2022a, 2022b, 2022c) y Sanchez et al. (2022). Elaboración propia.

Anexo H: Tablas de estimaciones del ATT sin bootstrapping en las rondas 3 y 5 con emparejamiento de 5 vecinos más cercanos (caliper de 0.005) dividiendo los alumnos con viajes largos a partir del intervalo de 40 minutos (alumnos de primaria y secundaria)

Tabla 81

Estimación del ATT sin bootstrapping en la ronda 3 con emparejamiento de 5 vecinos más cercanos, caliper de 0.005 y dividiendo los alumnos con viajes largos a partir del intervalo de 40 minutos (alumnos de primaria)

ATT (5 vecinos, caliper 0.005, ronda 3)		
Corte de tiempo	Alumnos con viajes largos < 40 min	Alumnos con viajes largos $\geq$ 40 min
$\geq$ 20 minutos	-0.016	0.265*
$\geq$ 25 minutos	0.194*	0.197
$\geq$ 30 minutos	0.152	0.164

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

Fuente: Boyden (2022a, 2022b, 2022c) y Sanchez et al. (2022).  
Elaboración propia.

Tabla 82

Estimación del ATT empleando bootstrapping en la ronda 5 con emparejamiento de 5 vecinos más cercanos, caliper de 0.005 y dividiendo los alumnos con viajes largos a partir del intervalo de 40 minutos (alumnos de secundaria)

ATT (5 vecinos, caliper 0.005, ronda 5)		
Corte de tiempo	Alumnos con viajes largos < 40 min	Alumnos con viajes largos $\geq$ 40 min
$\geq$ 20 minutos	-0.029	-0.014
$\geq$ 25 minutos	-0.007	0.052
$\geq$ 30 minutos	-0.026	0.013

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

Fuente: Boyden (2022a, 2022b, 2022c) y Sanchez et al. (2022).  
Elaboración propia.