

MODELACIÓN CONJUNTA DE LA ELECCIÓN DE ESTABLECIMIENTO ESCOLAR Y LA LOCALIZACIÓN RESIDENCIAL EN SANTIAGO, CHILE

Felipe González - Espejo, Pontificia Universidad Católica de Chile, fjgonzalez5@uc.cl

Ricardo Hurtubia, Pontificia Universidad Católica de Chile, rhurtubia@uc.cl

Sebastian Astroza, Universidad de Concepción, sastroza@udec.cl

RESUMEN

En Santiago, sobre 900 mil escolares viajan diariamente del hogar al estudio. La literatura ha modelado la elección residencial con diferentes variables de demanda por transporte. Sin embargo, pocos estudios abordan el vínculo entre la elección residencial y de establecimiento escolar, cuando el viaje con propósito educación se realiza a diario. La investigación propone modelos Logit Mixtos, para comprender las preferencias de los hogares sobre la elección de localización residencial y de establecimiento escolar. Se consideran dos estructuras que emulan la secuencia de decisiones: hogares priorizadores de vivienda, y hogares priorizadores de establecimientos. Se incorporan factores no observados sobre el nivel educacional de los barrios de la residencia y escuela, afectando las preferencias y correlacionando las elecciones. Los resultados indican que los hogares del modelo priorizador de residencia tienen mayor disposición a pagar por estar cerca del establecimiento escolar, en comparación con una estructura que prioriza la escuela.

Palabras claves: Modelación conjunta, Elección escolar, Factores no observados

ABSTRACT

In Santiago, over 900 thousand students travel daily from homes to schools. Transport literature has vastly modelled the joint choice of residential location and transport demand variables. However, it has not explored the relationship between residential and school choice, despite being a daily trip for many. Understanding this relationship is vital. This research, models Mixed Logits to understand household's preferences for residential choice and school choice. Two choice structures are considered: dwelling seekers and school seekers. Unobserved factors linked to neighborhood's education level in both the residence and school choice are incorporated, affecting household's preferences and correlating both choices. Results show that the dwelling seeker structure is more sensitive to distance variations, willing to pay more on monthly school payments, in comparison to a school seeker structure.

Keywords: Joint choice modelling, School choice, Unobserved factors

1. INTRODUCCIÓN

La elección de establecimiento educacional en Chile ha sido relevante en las últimas décadas. Tras la implementación del sistema de *vouchers* que dio cabida a establecimientos particulares subvencionados, se ha evidenciado que la elección de establecimiento escolar por parte de los hogares se debe a sus preferencias sobre atributos académicos de los establecimientos y sobre la distancia al hogar (Chumacero et al., 2011; Gómez et al., 2012). Sin embargo, la evidencia también ha mostrado la importancia de aspectos socioeconómicos de los establecimientos en la elección, lo que a veces prima por sobre las características académicas (Bellei et al., 2020; Elacqua et al., 2006), exacerbando fenómenos de aglomeración y segregación socioespacial. Por otra parte, el actual sistema de admisión escolar (que asigna según preferencias a los alumnos) no prioriza la distancia, ni características de la localización residencial, como factores relevantes al momento de asignar el establecimiento al estudiante (Correa et al., 2019).

Entender las preferencias de los hogares por los atributos de los establecimientos educacionales puede ayudar a comprender los patrones de localización residencial y, en consecuencia, la distribución socioespacial y otros fenómenos relevantes. Por ejemplo, la demanda por transporte y el crecimiento urbano. La Encuesta Origen Destino 2012 en el Gran Santiago (SECTRA, 2014) estima que en el período punta mañana de un día laboral normal, hay sobre 900 mil viajes escolares realizados por estudiantes menores de 19 años. Entender ambas elecciones por separado puede ser un error, pues se dejan de percibir potenciales factores que vinculen a ambas decisiones, relevantes en el comportamiento en la ciudad. Modelarlas en conjunto permite comprender las preferencias de los hogares sobre cada elección, captando los factores que las determinan y correlacionan.

El objetivo de este estudio es modelar en conjunto las elecciones de localización residencial y de establecimiento escolar en Santiago, Chile. Se espera comprender cómo se vinculan ambas decisiones y qué estructuras de decisión son apropiadas para el contexto de elección. Se propone un modelo Logit Mixto que incorpora términos de error para captar factores no observados entre ambas elecciones para modelar la elección conjunta. Además, el estudio modela dos estructuras de decisión diferentes. La primera, donde los hogares escogen de manera no condicionada la localización residencial, y luego eligen establecimiento educacional condicionado a la elección residencial. La segunda, los hogares eligen primero el establecimiento, y luego la localización residencial dada la elección de establecimiento observada. Se utilizó información de la Encuesta Origen Destino 2012 (SECTRA, 2014), y se utilizaron técnicas de muestreo de alternativas según lo propuesto por Ben-Akiva y Lerman (1985), debido al amplio universo muestral. Se espera dilucidar las diferentes estructuras de decisión, la correlación entre elecciones, y analizar compensaciones presentes entre atributos de las alternativas escolares y de localización residencial.

La estructura del artículo es la siguiente. Primero, la Sección 2.1 describe la literatura de elección escolar en Chile, y la relevancia que tiene para entender las preferencias de los hogares. Luego, la Sección 2.2 analiza el estado del arte de la modelación conjunta de elecciones y el aporte del artículo en la literatura. Después, la Sección 3 describe la metodología, los modelos a utilizar y cómo se establecen correlaciones entre ambas elecciones. La Sección 4 describe los datos recopilados y procesados. A continuación, la Sección 5 presenta los resultados de los modelos utilizados y el análisis de las variables más relevantes. Además, expone resultados de tasas

marginales de sustitución entre diferentes parámetros del modelo. La Sección 6 presenta conclusiones de la investigación, al igual que sus alcances y trabajo futuro.

2. REVISIÓN DE LITERATURA

La investigación consta de dos aspectos relevantes a estudiar. Por un lado, la elección de establecimiento escolar en Chile y las preferencias de los hogares. Por el otro, el análisis de modelos de elección conjunta y cómo esta investigación propone ser un aporte al área.

2.1 Elección escolar: el caso de Chile

La elección de establecimiento escolar en Chile ha sido ampliamente estudiada. El caso es singular, pues desde la implementación del sistema de *vouchers* en la década de 1980 se ha evidenciado una tendencia a incentivar la segregación social en la educación (Correa et al., 2019; Elacqua, 2012). Esto se ha debido a diferentes razones, dentro de las cuales está el sistema de selección escolar que permite a los establecimientos elegir entre una familia u otra, según sus características socioeconómicas y afinidad con el establecimiento (Gayo et al., 2019).

Sin embargo, los hogares también juegan un rol en esta dinámica social. Elacqua et al. (2006) evidencia cómo los padres y madres de cada hogar tienen una alta valoración por el nivel socioeconómico del cuerpo estudiantil. A pesar de declarar transversalmente el interés por atributos académicos y valóricos de los establecimientos, al momento de escoger las alternativas posibles tiende a primar la percepción del cuerpo estudiantil. Bellei et al. (2020) profundiza en el tema, comentando la relevancia para los hogares de nivel socioeconómico alto en Chile, del sentido de pertenencia al establecimiento y su comunidad. En ese sentido, los hogares en Chile tienden a compensar atributos de rendimiento académico con estatus social para elegir la mejor alternativa. Este comportamiento tiene sentido en el contexto chileno, donde se ha demostrado también la correlación entre el colegio de procedencia con el éxito profesional y universitario (Zimmerman, 2019). Por lo tanto, es razonable que un hogar valore la relevancia que la demografía del alumnado del colegio tenga para el éxito de sus hijos e hijas. En la misma línea, Gómez et al. (2012) discute cómo padres y madres declaran valorar el nivel académico, medida a través de pruebas estandarizadas, a pesar de no considerarlas luego en la elección (que puede ser un problema de información al momento de escoger). Esto es un reflejo de las preferencias de los hogares sobre los atributos de los establecimientos educacionales.

En la actualidad, el sistema de admisión escolar recibe las preferencias de los hogares que desean acceder a establecimientos municipales o subvencionados, y asigna a los estudiantes según diferentes criterios de prioridad. Los principales criterios son tener personas del núcleo familiar en el establecimiento, ser ex – alumnos (sin haber sido expulsados), necesidades especiales, rendimiento académico, entre otros. No obstante, la distancia no aparece como un criterio prioritario en la asignación. Es más, esta se considera exclusivamente cuando el estudiante no ha sido seleccionado en ningún establecimiento, y por ende se le asigna el de mayor cercanía al hogar. Esto evidencia la poca relevancia entregada al atributo. Sin embargo, Chumacero et al. (2011) demuestra que la distancia desde el hogar es tan relevante como otros atributos académicos en la

elección de establecimientos escolares en Chile. Esta investigación busca enfatizar en la relevancia de este atributo, sumado a las compensaciones que los hogares están dispuestos a realizar entre diferentes atributos académicos y sociales del establecimiento. Además, pretende incorporar en el análisis aspectos del barrio del establecimiento y de la localización residencial que determinen las preferencias de los hogares en la elección.

2.2 Modelos conjuntos de elección residencial y otras elecciones

La modelación conjunta de elecciones ha sido largamente estudiada en la literatura de modelación de transporte y uso de suelo. Por ejemplo, Ben-Akiva (1973) propone un modelo de elección conjunta entre la elección de destino y modo de viaje, a través de una estructura jerárquica (HL) con errores que distribuyen Generalized Extreme Value (GEV). En esa misma línea, han surgido investigaciones que tratan, en particular, la elección de localización residencial con otras decisiones. Lerman (1976) elabora un modelo Logit Multinomial (MNL) que combina decisiones de elección residencial, posesión vehicular y modo de viaje. Mientras que Waddell (1993) planteó un modelo de igual estructura, pero vinculando las elecciones de residencia y trabajo. Estas estructuras multinomiales demuestran el vínculo estrecho entre elecciones mediante atributos asociados a la combinación de alternativas. No obstante, no logran establecer causalidad de una decisión sobre la otra, pues trabajan la elección en simultáneo sin componentes de error que generen jerarquías. El modelo HL de Ben-Akiva es atractivo en este aspecto, pues estima correlaciones entre elecciones que determinan jerarquías de decisión.

Por otro lado, han existido otras propuestas para establecer estructuras que determinan causalidad en las decisiones. Waddell et al. (2007) estima las elecciones de localización residencial y de trabajo a través de un modelo de clases latentes, pero trabaja con una elección exógena que condiciona a la otra. Por ejemplo, modela la elección residencial dado que ya se conoce el lugar de trabajo, sin modelar ambas elecciones como variables endógenas. Se han propuesto también modelos con estructuras de error menos restringidas, como modelos Probit con términos de error compartidos entre las utilidades asociadas a distintas elecciones (Astroza et al., 2019; Bhat, 2015; Paleti et al., 2013). Estos modelos establecen causalidades entre las diferentes elecciones a través de atributos condicionados a las elecciones previas, y la componente normal de los errores permite estimar factores no observados entre las diferentes elecciones. Además, mediante enfoques de clases latentes (Astroza et al., 2019) se puede identificar diferentes tipos de tomadores de decisión según sus estilos de vida, opiniones y valores que determinen las estructuras de sus preferencias. No obstante, y aunque la brecha computacional ha disminuido gracias a nuevas técnicas de estimación (Bhat, 2011), estos modelos pueden ser desafiantes para universos muestrales de gran tamaño, como la elección de localización residencial o de establecimiento escolar. En ese sentido, lo propuesto por Pinjari et al. (2011) es interesante para los alcances del estudio. Este modela variadas elecciones multinomiales y ordinales, y las vincula a través de componentes de error compartidos en sus atributos. De esta manera busca captar la correlación entre elecciones, apoyándose de estructuras tratables como el Logit.

A pesar de la relevancia diaria que estas elecciones tienen, no se conoce (desde la perspectiva de los autores) una investigación que vincule las elecciones escolar y residencial a través de causalidades y estructuras de correlación. El trabajo de Barrow (2002) modela la elección

residencial condicionada a atributos escolares en EE.UU. Sin embargo, como en Estados Unidos el establecimiento depende del barrio residencial, los atributos se pueden entender como una medida de accesibilidad a educación del barrio. Además, propone una estructura de decisión donde es el establecimiento escolar el que determina la elección de localización residencial, y no en el sentido contrario. Por otra parte, Chumacero et al. (2011) modela la elección de establecimiento escolar y utiliza la distancia como atributo vinculante entre el establecimiento y la localización residencial. No obstante, no considera otros atributos ni modela la elección de esta última. Así, esta investigación pretende vincular ambas elecciones y comprender las estructuras de decisión que los hogares pueden tener sobre estas.

3. METODOLOGÍA

La metodología propuesta buscar modelar la elección de los hogares en torno a localización residencial y establecimiento educacional para las hijas e hijos de este. Sea $q = \{1, \dots, Q\}$, el hogar que debe elegir de entre las alternativas de localización residencial $r = \{1, \dots, R\}$ y de establecimiento educacional $c = \{1, \dots, C\}$. Ambos conjuntos de alternativas son no etiquetados.

Se asume que existen dos tipos de hogares con diferentes estructuras de elección. Estas estrategias definen una condicionalidad o prioridad de una elección sobre la otra:

- i) Hogares que seleccionan de manera incondicional su localización residencial, y luego eligen un establecimiento educacional, condicional a la elección residencial ya realizada (denominados priorizadores de localización residencial, PLR de ahora en adelante)
- ii) Hogares que seleccionan establecimiento educacional de manera no condicionada, para luego elegir una localización residencial, condicional al establecimiento seleccionado (denominados priorizadores de establecimiento educacional, PEE de ahora en adelante).

Cabe mencionar que la metodología propone dos modelos distintos, y en cada uno se asume que los todos los hogares se comportan o como hogares PLR o PEE de manera excluyente. Para una modelación que asume hogares PLR, la función de utilidad asociada a una alternativa de localización residencial viene dada por:

$$U_{qr} = \sum_i (\theta_i + \theta_{iq} \cdot x_q) \cdot z_{iqr} + \sum_j (\theta_j + \theta_{jq} \cdot x_q + \gamma_j) \cdot z_{jqr} + \varepsilon_{qr} \quad (1)$$

Donde los parámetros θ_i corresponden a los coeficientes que acompañan al i -ésimo atributo de la alternativa de localización r para el hogar q (z_{iqr}). Los parámetros θ_{iq} corresponden a los coeficientes que acompañan variaciones sistemáticas de gustos, con respecto al atributo z_{iqr} , según las características x_q del hogar q . Los parámetros θ_j y θ_{jq} son similares a los parámetros previos, pero para el j -ésimo atributo de la alternativa (z_{jqr}), que corresponde a un atributo que está también presente en las alternativas de otra decisión (en este caso la elección de establecimiento educacional) y que introduce una correlación entre las elecciones. El parámetro γ_j corresponde a un término de error aleatorio de distribución Normal, que busca captar el impacto de factores no observados del atributo z_{jqr} sobre la elección. Por último, ε_{qr} corresponde a un término de error que sigue una distribución Gumbel IID $(0, \mu)$.

Luego, la función de utilidad que determina la elección de establecimiento educacional, dada la localización residencial observada, r^* , está dada por:

$$U_{qc|r^*} = \sum_k (\beta_k + \beta_{kq} \cdot x_q) \cdot z_{kqc} + \sum_j (\beta_j + \beta_{jq} \cdot x_q + \gamma_j) \cdot z_{jqc} + \phi \cdot d_{qc|r^*} + \varepsilon_{qc} \quad (2)$$

Donde los parámetros β_k corresponden a los coeficientes que acompañan al k -ésimo atributo de la alternativa c para el hogar q (z_{kqc}). Los parámetros β_{kq} corresponden a los coeficientes que acompañan variaciones sistemáticas de gustos. Los parámetros β_j y β_{jq} acompañan a los atributos que coinciden con los j -ésimos atributos de las alternativas de localización residencial. El parámetro γ_j corresponde al mismo término de error de distribución Normal de la ecuación (1), que capta el impacto de factores no observados del atributo z_{jqc} . Es decir, este término de error se encuentra presente en ambos conjuntos de funciones de utilidad, y busca captar correlación entre ambas elecciones. Por lo tanto, z_{jqr} y z_{jqc} corresponden al mismo atributo, pero para alternativas de localización residencial y de elección de establecimiento escolar. Por ejemplo, el número de áreas verdes en el barrio cercano a la alternativa residencial/escolar, o la proporción de personas con nivel educacional alto en el barrio de la alternativa residencial/escolar.

Finalmente, el coeficiente ϕ acompaña al atributo de la distancia al establecimiento educacional, condicional a la elección residencial ya escogida. Este atributo es el que determina la secuencialidad en las elecciones. El término ε_{qc} corresponde a un término de error con distribución Gumbel IID $(0, \lambda)$.

La probabilidad para un hogar q de escoger una localización residencial r y un establecimiento escolar c está dada por:

$$P_q(r, c) = \int \frac{\exp(V_{qr})}{\sum_{l=1}^R \exp(V_{ql})} \cdot \frac{\exp(V_{qc|r^*})}{\sum_{n=1}^C \exp(V_{qn|r^*})} \cdot f(\gamma) \cdot d\gamma \quad (3)$$

Donde V_{qr} y V_{qc} corresponden a las utilidades sistemáticas de las U_{qr} y U_{qc} . $f(\gamma)$ corresponde a la función de distribución de probabilidad de los términos de error γ_j que explican los factores no observados que comparten ambas elecciones.

Por el otro lado, para el modelo con una estructura PEE, las funciones de utilidad están dadas por:

$$U_{qc} = \sum_k (\beta_k + \beta_{kq} \cdot x_q) \cdot z_{kqc} + \sum_j (\beta_j + \beta_{jq} \cdot x_q + \gamma_j) \cdot z_{jqc} + \varepsilon_{qc} \quad (4)$$

$$U_{qr|c^*} = \sum_i (\theta_i + \theta_{iq} \cdot x_q) \cdot z_{iqr} + \sum_j (\theta_j + \theta_{jq} \cdot x_q + \gamma_j) \cdot z_{jqr} + \phi \cdot d_{qr|c^*} + \varepsilon_{qr} \quad (5)$$

Donde todos los parámetros poseen el mismo significado que los expuestos en las ecuaciones (1) y (2). La principal diferencia para este tipo de hogar es que el atributo de distancia y su coeficiente ϕ se encuentran en la función de utilidad de elección residencial. Por lo tanto, la distancia es condicional al establecimiento educacional observado, c^* . Así, la causalidad en las decisiones está determinada por el atributo de distancia y la condicionalidad que posea. De este modo, la

probabilidad del hogar q de escoger un establecimiento c y una localización residencial r está dada por:

$$P_q(c, r) = \int \frac{\exp(V_{qc})}{\sum_{n=1}^C \exp(V_{qn})} \cdot \frac{\exp(V_{qr|c^*})}{\sum_{l=1}^R \exp(V_{ql|c^*})} \cdot f(\gamma) \cdot d\gamma \quad (6)$$

Finalmente, para ambos modelos, la función de log-verosimilitud a maximizar corresponde a:

$$l(\theta, \beta, \gamma, \phi) = \sum_{q \in Q} \sum_{r \in R} \sum_{c \in C} g_{qr} \cdot g_{qc} \cdot \ln(P_q(r, c)) \quad (7)$$

Donde g_{qc} y g_{qr} toman valor 1 si el hogar q escoge la alternativa r y c , respectivamente. θ, β, γ y ϕ corresponden a vectores que contienen todos los parámetros a estimar según lo expuesto.

4. DATOS

Los datos provienen de la Encuesta Origen Destino 2012 para el Gran Santiago (SECTRA, 2014). De esta, se tomaron aquellos hogares, personas (estudiantes menores de 18 años) y viajes con propósito educación escolar. Cada observación posee información del hogar y sus características: composición y estructura, disponibilidad de vehículos, ingreso, entre otros, al igual que información con respecto a los estudiantes del hogar (sexo, nivel de estudios, edad, entre otros). Además, se posee información de los modos utilizados en el viaje escolar, el tiempo de viaje, al igual que las coordenadas de origen (localización del hogar) y destino (localización del establecimiento educacional). Se utilizaron Sistemas de Información Geográfica (SIG) para emparejar el destino del viaje y la localización de establecimientos en el gran Santiago, identificando de esta forma los establecimientos educacionales seleccionados.

Las alternativas de localización residencial son caracterizadas con datos poblacionales del Censo 2017 (INE, 2018) y del Catastro de Bienes Raíces del Servicio de Impuestos Internos (SII, 2014). Se posee información con respecto a los habitantes, su ingreso y nivel educacional, además de los m² de suelo, número de viviendas y calidad de la edificación. Información adicional, como el valor del suelo, se obtuvo a través del arriendo declarado por los encuestados en la EOD 2012 (SECTRA, 2014). Así, la unidad de análisis corresponde a hexágonos regulares de 72,5 m de lado con la información descrita. El universo muestral se compone de 48.724 hexágonos.

Los establecimientos educacionales son caracterizados a partir de información entregada por el Ministerio de Educación (MINEDUC, 2014). Se obtuvo información con respecto a los establecimientos, los niveles educacionales impartidos, número de docentes, matrículas, orientación religiosa, subvención, resultados de pruebas estandarizadas, al igual que información del barrio sobre el cuál se localiza el establecimiento. El universo muestral corresponde a 938 establecimientos educacionales. De ellos, todos imparten al menos enseñanza básica o media. Se realizaron técnicas de muestreo de alternativas explicadas en la Sección 4.1. Cabe mencionar que los valores monetarios corresponden a los vigentes en cada base de datos, por lo que no existen ajustes sobre estos.

Para los conjuntos de alternativas, se calculó la distancia euclidiana entre cada posible localización residencial y establecimiento. La **Tabla 1** describe a la muestra de hogares y estudiantes.

Tabla 1: Características de la muestra de hogares

| Hogares | Cantidad | % | Estudiantes | Cantidad | % |
|--------------------------------|----------|------|--------------------------------------|----------|-------|
| Observaciones | 1981 | 100 | Observaciones | 2582 | 100,0 |
| Integrantes | | | Sexo | | |
| 2 personas | 56 | 2,8 | Femenino | 1261 | 48,8 |
| 3 personas | 385 | 19,4 | Edad | | |
| 4 personas | 683 | 34,5 | 5 - 6 Años | 289 | 11,2 |
| 5 o + personas | 489 | 24,7 | 7 - 13 años | 1280 | 49,6 |
| 6 o + | 378 | 19,1 | 7 - 13 años | 1013 | 39,2 |
| Ingreso per cápita | | | Dependencia | | 0,0 |
| 0 - 100 mil CLP | 651 | 32,9 | Municipal | 603 | 23,4 |
| 100 - 400 mil CLP | 1201 | 60,6 | Particular Subvencionado | 1586 | 61,4 |
| > 400 mil CLP | 129 | 6,5 | Particular | 393 | 15,2 |
| Posee vehículo | 910 | 45,9 | Enseñanza Media | 839 | 32,5 |
| 1 | 711 | 35,9 | Científico - Humanista | 769 | 29,8 |
| 2 o + | 199 | 10,0 | Técnico Profesional | 70 | 2,7 |
| 2 o más escolares | 504 | 25,4 | Orientación religiosa escolar | | 0,0 |
| En el mismo establecimiento | 320 | 16,2 | Laica | 1681 | 65,1 |
| En establecimientos diferentes | 184 | 9,3 | Católica | 901 | 34,9 |
| Casa Propia | 1510 | 76,2 | Modo de viaje a la escuela | | 0,0 |
| Ubicación geográfica | | | Motorizado | 1917 | 74,2 |
| Norte | 205 | 10,3 | Auto | 516 | 20,0 |
| Poniente | 486 | 24,5 | Bus Transantiago | 702 | 27,2 |
| Oriente | 252 | 12,7 | Metro | 138 | 5,3 |
| Centro | 77 | 3,9 | Taxi | 56 | 2,2 |
| Sur | 468 | 23,6 | Bus (No Transantiago) | 10 | 0,4 |
| Sur - Oriente | 493 | 24,9 | Otro modo | 593 | 23,0 |
| | | | No Motorizado | 665 | 25,8 |
| | | | Caminata | 630 | 24,4 |
| | | | Bicicleta | 35 | 1,4 |

De los hogares se observa que la mediana posee 4 integrantes familiares, mientras que la mayoría tiene un ingreso per cápita entre los 100 mil y 400 mil CLP. Menos de la mitad de la muestra posee vehículo, mientras que sólo el 10% posee 2 o más automóviles. También, un 25,4% de los hogares tiene más de un estudiante escolar. A estos hogares se les aplicó un efecto panel para captar de mejor manera heterogeneidad en las preferencias. Esto significa que el término de error aleatorio γ_j se repite para observaciones del mismo hogar. De estos 504 hogares, 184 escogen diferentes establecimientos para los estudiantes del hogar. Por último, de las 7 macrozonas definidas por la EOD del Gran Santiago (por falta de información, se eliminaron las observaciones de las 2

restantes: Extensión Sur – Poniente y Exterior a RM. No obstante, corresponden a ciudades periféricas que no siempre realizan viajes hacia Santiago, por lo que su descarte no es determinante). El sector Sur y Sur – Oriente capta un 48,5% de la muestra, mientras que el centro solo un 3,9%. 3 de cada 4 hogares son propietarios de la localización donde residen.

Por parte de los estudiantes, se consideraron 2582 observaciones, repartidas equitativamente por sexo. De la muestra (menores de 18 años), un 11,2% declaró tener edad entre 5 y 6 años. Un 49,6% se encuentra entre los 7 y 13 años, y un 39,2% son mayores de 14 años. De los establecimientos escogidos, 1 cada 4 estudiantes atienden un establecimiento municipal, mientras que más del 60% atienden a establecimientos particulares subvencionados. Solo un 15% atienden colegios privados. Además, de los estudiantes que cursan educación media (un 32,5% de la muestra), la mayoría optó por una educación Científico - Humanista. Por último, 3 de cada 4 estudiantes viajan en un modo motorizado. Dentro de los modos que utilizan (uno o varios), se encuentran principalmente el auto, bus Transantiago u otros modos. Estos últimos, potencialmente podrían corresponder a transporte escolar (SECTRA, 2014). Un 24,4% declara caminar al colegio.

4.1 Muestreo de alternativas

Dado que los universos de alternativas para ambas elecciones son de gran tamaño, se realizó un muestreo de alternativas para cada observación. En la elección de localización residencial, se realizó un muestreo de 19 alternativas (20 considerando la escogida) según la localización de este. En función de la macrozona EOD (7 macrozonas posibles) en el que se localiza el hogar, se realizó un muestreo aleatorio para las primeras 9 alternativas. Luego, las otras 10 alternativas se obtenían de manera aleatoria de los otros seis sectores posibles. Mientras, en la elección de establecimiento escolar, se realizó un muestreo de 19 alternativas (20 considerando la escogida), según la dependencia de este. Si el establecimiento escogido era municipal, las primeras 9 alternativas se muestrearon aleatoriamente de las alternativas escolares municipales, luego se muestrearon 5 alternativas de establecimientos particulares, y 5 de establecimientos subvencionados.

A cada observación, y para cada conjunto de alternativas, se aplicó un factor de corrección asociado a la probabilidad de escoger el conjunto de alternativas del universo muestral, según la técnica de muestreo de alternativas explicada en profundidad por Ben-Akiva y Lerman (1985). Esta técnica es equivalente a la propuesta por McFadden (1978). Este factor se incorpora en las probabilidades de elección.

5. RESULTADOS Y DISCUSIÓN

Se estimaron los modelos de elección con estructuras PLR y PEE planteadas en la Sección 3. Los resultados se presentan en la **Tabla 2**. Con 2000 realizaciones de Montecarlo, se obtuvieron estimaciones de los coeficientes que acompañan atributos de la elección de localización residencial y establecimiento escolar. Se optó por mantener aquellos coeficientes poco significativos para comparar de igual manera ambas especificaciones, entendiendo para qué estructuras algunos atributos se hacen significativos y cuáles no. El único término que es diferente es el de

“Accesibilidad a educación”, ya que la segunda estructura toma una elección de localización residencial ya conociendo el establecimiento escolar escogido.

Ambos modelos exponen una alta significancia del parámetro de distancia entre residencia y establecimiento. Para el modelo PLR, esta distancia es condicionada a la localización residencial escogida, mientras que para el modelo PEE, la distancia es condicionada al establecimiento escogido. En ambos, se muestra una sensibilidad negativa relevante por parte de los individuos.

Tabla 2: Resultados de modelos con estructuras PLR y PEE.

| Especificación | Estructura PLR | | Estructura PEE | |
|--|----------------|--------|----------------|--------|
| | Valor | t-test | Test-t | t-test |
| Atributo | | | | |
| <i>Distancia entre residencia y establecimiento (en km)</i> | -0,73 | -36,9 | -0,78 | -40,5 |
| <i>Interacción para hogares con vehículo</i> | 0,11 | 4,4 | 0,12 | 4,8 |
| Coeficientes de elección escolar | | | | |
| Establecimiento con educación Científico - Humanista | -0,4 | -3,5 | -0,28 | -3,4 |
| <i>Interacción para hogares con Ingreso per Cápita 100 - 400 mil CLP</i> | 0,24 | 1,6 | 0,13 | 1,2 |
| <i>Interacción para hogares con Ingreso per Cápita > 400 mil CLP</i> | 1,89 | 4,3 | 1,68 | 4,9 |
| Índice de Vulnerabilidad | | | | |
| <i>Interacción para hogares con Ingreso per Cápita \geq 100 mil CLP</i> | -0,92 | -3,3 | -1,47 | -7,6 |
| Mensualidad (en MM CLP) | -2,53 | -4,6 | -3,38 | -8,2 |
| Establecimiento no coeducacional | -1,61 | -16,5 | -0,99 | -14,9 |
| % de jefes de hogar en el barrio del establecimiento con educación superior aprobada (valor entre 0 - 1) | -0,11 | -0,3 | -2,55 | -9,3 |
| <i>Interacción para hogares con Ingreso per Cápita 100 - 400 mil CLP</i> | 2,02 | 4,0 | 2,04 | 6,5 |
| <i>Interacción para hogares con Ingreso per Cápita > 400 mil CLP</i> | 3,19 | 3,7 | 7,05 | 12,5 |
| <i>Desv. Estándar % jefes de hogar con Ed. Superior</i> | 3,55 | 19,0 | 2,98 | 17,7 |
| Coeficientes de elección residencial | | | | |
| Valor promedio de la renta mensual (en MM CLP) | -0,68 | -3,1 | -0,98 | -3,0 |
| Accesibilidad a comercio | 1,58 | 2,6 | -1,61 | -2,3 |
| Accesibilidad a educación | 3,51 | 11,4 | - | - |
| Accesibilidad a industria | -3,05 | -13,9 | -2,19 | -8,0 |
| Accesibilidad a oficinas | -1,24 | -2,4 | 0,63 | 1,0 |
| % de jefes de hogar en el barrio residencial con educación superior aprobada (valor entre 0 - 1) | -6,66 | -18,4 | -5,3 | -11,9 |
| <i>Interacción para hogares con Ingreso per Cápita 100 - 400 mil CLP</i> | 11,7 | 18,0 | 9,92 | 12,1 |
| <i>Interacción para hogares con Ingreso per Cápita > 400 mil CLP</i> | 4,21 | 11,6 | 3,35 | 7,0 |
| <i>Desv. Estándar % jefes de hogar con Ed. Superior</i> | 3,55 | 19,0 | 2,98 | 17,7 |
| Estadísticas generales | | | | |
| N° de observaciones | 2582 | | 2582 | |
| Log-Verosimilitud | -12769,5 | | -12368,5 | |

| Especificación | Estructura PLR | Estructura PEE |
|-----------------------|---------------------------|---------------------------|
| Atributo | Valor t-test | Test-t t-test |
| ρ^2 ajustado | 0,315 | 0,337 |
| AIC | 25578,6 | 24775,1 |

Con respecto a los atributos escolares, ambos modelos exponen una preferencia hacia establecimientos con enseñanza Científico Humanista a medida que aumentan los niveles de ingreso de los hogares. Para hogares de menor ingreso, el parámetro es negativo, pero esto puede deberse a problemas de endogeneidad en las elecciones. No se puede definir si estos hogares prefieren establecimientos técnicos profesionales, o hay variables no incorporadas en el modelo que se expliquen a través de este parámetro. En torno al Índice de Vulnerabilidad (IVE), que mide qué porcentaje del alumnado proviene de contextos vulnerables, se demostró en ambos modelos que hogares de ingresos por sobre los 100 mil CLP per cápita tienden a percibir de manera negativa mayores niveles de vulnerabilidad que los hogares con ingresos menores, quienes son indiferentes. Se demostró también una sensibilidad negativa con respecto a la mensualidad a pagar en los establecimientos, lo cual es razonable pues implica un costo para las familias. Además, existe una preferencia por establecimientos coeducacionales por sobre los no coeducacionales en ambos modelos, para los hogares de la muestra. No se encontró significancia en variables de rendimiento académico como pruebas SIMCE o PSU, lo que va en línea con la literatura (Gómez et al., 2012; Elacqua et al., 2006).

Por otra parte, se estimó el impacto que tiene el nivel educacional del barrio del establecimiento escolar, medido en la proporción de jefes de hogar del barrio que poseen educación superior. Los resultados muestran que para los hogares de ingresos medios y altos (sobre los 100 mil CLP), se prefieren establecimientos en barrios de mayor nivel educacional en comparación con hogares de menor ingreso. También, se probó que en ambos modelos de manera significativa existen factores no observados sobre el nivel educacional del barrio que determinan la preferencia de esta elección.

En los coeficientes de los atributos de la elección de localización residencial existen los siguientes resultados. Primero, parámetros como los asociados al valor mensual de la renta a pagar, accesibilidad a industria y educación tienen comportamientos razonables. Los hogares tienden a preferir pagar menos por una localización, al igual que ubicarse cerca de escuelas y lejos de actividad industrial. No obstante, es interesante el comportamiento opuesto que se presenta en atributos de accesibilidad a comercio y oficinas. Los resultados para el modelo con estructura PLR indican una preferencia por parte de los hogares hacia localizaciones con buena accesibilidad a comercio, mientras que no prefiere localizaciones que tengan actividades de trabajo cerca. En cambio, los resultados para el modelo de estructura PEE propone lo opuesto, buscando estar cerca de oportunidades de trabajo, pero lejos del comercio. Esto podría deberse a diferentes factores no observados de cada hogar, en torno a sus preferencias, opiniones, valores o estilos de vida.

Además, la variable asociada al nivel educacional del barrio, la variable seleccionada para detectar correlación entre las elecciones tuvo un comportamiento similar a su símil en la elección escolar. Es decir, a medida que aumenta el nivel de ingreso del hogar, existe una preferencia por escoger alternativas residenciales cuyo entorno tenga más jefes de hogar con educación superior. Esta variable es un *proxy* del nivel socioeconómico del barrio, tanto de la alternativa residencial como

escolar. La endogeneidad para el caso base puede deberse a que hogares de ingresos menores, no tienden a poder acceder a barrios de mayor nivel socioeconómico. También, pueden existir atributos omitidos, observables y no observables que afecten esta variable. Además, se comprueba que existen factores no observados que determinan la sensibilidad sobre esta elección. Este término está presente tanto en la elección residencial como escolar, para ambos modelos. Entonces, factores no observados con respecto al barrio tanto de la localización residencial como del establecimiento escolar, afectan de la misma manera la preferencia del hogar sobre ambas elecciones. Estos factores podrían estar asociados a la valoración que los hogares tienen sobre el nivel socioeconómico de los establecimientos, comentado en la Sección 2.1.

5.1 Tasas marginales de sustitución

Los modelos Logit poseen problemas de identificabilidad, pues los parámetros de escala acompañan a todos los coeficientes estimados. Entonces, para poder comparar las magnitudes de las preferencias entre modelos, es necesario comparar las compensaciones que los hogares hacen entre atributos, para entender el efecto de los parámetros sin considerar el factor de escala. Esto se logra mediante tasas marginales de sustitución, determinadas por la razón entre los coeficientes estimados. Estas representan la disposición a compensar de un atributo por otro. La **Tabla 3** expone las razones de variables de interés. Cabe mencionar que el modelo no refleja necesariamente los valores subjetivos reales de la población, pero sirve para ilustrar diferencias entre estructuras de decisión.

Tabla 3: Tasas marginales de sustitución

| Tasa de Utilidades Marginales | Unidad | PLR | PEE |
|-------------------------------------|---------------------------|--------|--------|
| $UMg_{Distancia}/UMg_{Mensualidad}$ | $\frac{mil\ CLP}{km}$ | | |
| Hogares sin vehículo | | 288,54 | 230,18 |
| Hogares con vehículo | | 243,87 | 195,56 |
| $UMg_{IVE}/UMg_{Distancia}$ | $\frac{km}{\% IVE}$ | | |
| Hogares sin vehículo | | 0,013 | 0,019 |
| Hogares con vehículo | | 0,015 | 0,022 |
| $UMg_{IVE}/UMg_{Mensualidad}$ | $\frac{mil\ CLP}{\% IVE}$ | | |
| Ingreso per cápita > \$100mil CLP | | 3,64 | 4,35 |
| $UMg_{Distancia}/UMg_{Renta}$ | $\frac{MM\ CLP}{km}$ | | |
| Hogares sin vehículo | | 1,07 | 0,79 |
| Hogares con vehículo | | 0,9 | 0,67 |
| $UMg_{Renta}/UMg_{Distancia}$ | $\frac{CLP}{CLP}$ | | |
| | | 0,27 | 0,29 |

Para la estructura de elección PLR, los hogares tienen una mayor sensibilidad a la distancia del establecimiento escolar con respecto a la mensualidad. Es decir, 1 kilómetro más de distancia entre el colegio y la residencia equivale a una mayor mensualidad que para el modelo PEE. Además, la estructura con hogares PLR tiene una mayor sensibilidad de la distancia con respecto a la renta.

Esto significa que alejarse más del colegio es equivalente a un mayor aumento del precio a pagar por la residencia, en comparación con la estructura PEE.

Por otra parte, para el modelo que asume hogares PEE, hay más sensibilidad al nivel socioeconómico del establecimiento. Un aumento porcentual en la vulnerabilidad de este equivale a vivir más lejos, o pagar más mensualidad, en comparación con el modelo PLR. También, para la estructura PEE, existe una valoración mayor de la mensualidad sobre la renta. Un peso más de renta equivale a 0,29 pesos más de mensualidad. Esto reflejaría una disposición a pagar más mensualidad escolar que pagar más por la localización, en comparación con la estructura PLR. Esto puede evidenciar que cada estructura valora una elección por sobre la otra.

En general, estas tasas pueden reflejar que la estructura PLR se adapta a hogares que les dan más relevancia a variables como la distancia, lo que podría estar determinado por diferentes opiniones, valores o estilos de vida, por ejemplo, usar modos sustentables. Mientras, la estructura PEE podría adecuarse a hogares con menor sensibilidad a la distancia, pero mayor interés por el establecimiento educacional y sus atributos. Cabe notar que los modelos imponen sobre los hogares una u otra estructura de decisión, a pesar de que estos puedan no elegir de esa forma. Sin embargo, el hecho que existan diferencias entre modelos da luces de que puede existir heterogeneidad en las preferencias y existen diferentes clases de hogares con estructuras de decisión diferentes.

6. CONCLUSIONES

La relación entre las decisiones de localización residencial y establecimiento educacional ha sido pocas veces planteada en la literatura de modelos de elección discreta. No obstante, y dado que el viaje por educación se realiza diariamente por miles de estudiantes, entender esta relación es de vital importancia para comprender el comportamiento de los hogares en la ciudad. Particularmente, en un contexto de heterogeneidad en las estructuras de decisión.

Este artículo, mediante información de las personas, localizaciones residenciales y establecimientos educacionales de Santiago, modela en conjunto la elección de localización residencial y establecimiento escolar. A través de un modelo Logit Mixto, estima factores no observados que vinculan ambas decisiones, y establece causalidad a través de atributos condicionados a la elección previa. Para captar diferentes estructuras de decisión, se estimaron dos modelos. El primero, con una estructura de decisión PLR y el segundo una estructura PEE.

De los resultados se puede concluir que existen factores no observados, asociados al nivel educacional del entorno de la residencia y del colegio, que afectan la utilidad de una alternativa escolar (o residencial) sobre la otra. Esto permite establecer correlación entre la elección de localización residencial y la elección escolar. Además, factores asociados al nivel del barrio del establecimiento demuestran la relevancia que los hogares entregan a atributos socioeconómicos de este, corroborando lo expuesto por la literatura de elección escolar. También, se comprueba que independiente de la estructura de decisión, la distancia es un factor relevante que los hogares consideran al elegir. Esto da indicios que debería ser un aspecto por considerar al momento de asignar a estudiantes a los establecimientos educacionales.

También, mediante el cálculo de tasas marginales de sustitución, se pueden comparar los *trade-offs* presentes según las estructuras de decisión determinadas. Para el modelo PLR hay una mayor sensibilidad con respecto a la distancia al colegio en comparación con aspectos como mensualidad, valor de la renta o índice de vulnerabilidad escolar. Mientras, para la estructura PEE, hay mayor sensibilidad hacia aspectos vinculados con el nivel socioeconómico del colegio en comparación con la distancia o la mensualidad. A pesar de que cada modelo asume que todos los hogares se comportan de una u otra forma, revela que existe heterogeneidad en las preferencias, que se podría captar según diferentes tipos de hogares, según factores no observados asociados a opiniones, valores o estilos de vida de estos. Todo esto, en miras de comprender de mejor forma las preferencias de los hogares sobre decisiones de localización y de actividades que moldean la estructura de viajes diaria de la ciudad.

REFERENCIAS

- Astroza, S., Garikapati, V. M., Pendyala, R. M., Bhat, C. R., y Mokhtarian, P. L. (2019). Representing heterogeneity in structural relationships among multiple choice variables using a latent segmentation approach. **Transportation**, 46(5), 1755–1784. <https://doi.org/10.1007/s11116-018-9882-7>
- Barrow, L. (2002). School choice through relocation: Evidence from the Washington, D.C. area. **Journal of Public Economics**, 86(2), 155–189. [https://doi.org/10.1016/S0047-2727\(01\)00141-4](https://doi.org/10.1016/S0047-2727(01)00141-4)
- Bellei Carvacho, C., Orellana Calderón, V., y Canales Cerón, M. (2020). School choice within the upper class. Community, identity and social closure. **Education Policy Analysis Archives**, 28. <https://doi.org/10.14507/epaa.28.3884>
- Ben-Akiva, M. (1973). **Structure of Passenger Travel Demand Models**. PhD Thesis, Massachusetts Institute of Technology. <https://dspace.mit.edu/handle/1721.1/14790>
- Ben-Akiva, M., y Stephen R. Lerman. (1985). **Discrete Choice Analysis: Theory and Application to Travel Demand (3rd ed.)**. The MIT Press.
- Bhat, C. R. (2011). The maximum approximate composite marginal likelihood (MACML) estimation of multinomial probit-based unordered response choice models. **Transportation Research Part B: Methodological**, 45(7), 923–939. <https://doi.org/10.1016/j.trb.2011.04.005>
- Bhat, C. R. (2015). A new generalized heterogeneous data model (GHDM) to jointly model mixed types of dependent variables. **Transportation Research Part B: Methodological**, 79, 50–77. <https://doi.org/10.1016/j.trb.2015.05.017>
- Cal, M. G., Cabrol, G. O., y Méndez, M. L. (2019). School choice and the selection of families: Upper middle class reproduction in Santiago de Chile. **Revista Internacional de Sociología**, 77(1), 1–16. <https://doi.org/10.3989/ris.2019.77.1.17.310>

- Chumacero, R. A., Gómez, D., y Paredes, R. D. (2011). I would walk 500 miles (if it paid): Vouchers and school choice in Chile. **Economics of Education Review**, 30(5), 1103–1114. <https://doi.org/10.1016/j.econedurev.2011.05.015>
- Correa, J., Epstein, R., Escobar, J., Rios, I., Bahamondes, B., Bonet, C., Epstein, N., Aramayo, N., Castillo, M., Cristi, A., y Epstein, B. (2019). School choice in Chile. **ACM EC 2019 - Proceedings of the 2019 ACM Conference on Economics and Computation**, 325–343. <https://doi.org/10.1145/3328526.3329580>
- Elacqua, G. (2012). The impact of school choice and public policy on segregation: Evidence from Chile. **International Journal of Educational Development**, 32(3), 444–453. <https://doi.org/10.1016/j.ijedudev.2011.08.003>
- Elacqua, G., Schneider, M., y Buckley, J. (2006). School choice in Chile: Is it class or the classroom? **Journal of Policy Analysis and Management**, 25(3), 577–601. <https://doi.org/10.1002/pam.20192>
- Gómez, D., Chumacero, R. A., y Paredes, R. D. (2012). School choice and information. **Estudios de economía**, 39(2), 143–157. <https://doi.org/10.4067/s0718-52862012000200003>
- INE. (2018). *Resultados Censo 2017*.
- McFadden. (1978). *Modelling the choice of residential location*. Cowles Foundation for Research in Economics. Yale University.
- MINEDUC. (2014). *Directorio Establecimientos Educativos 2014*. <http://datos.mineduc.cl/dataviews/235565/vista-directorio-de-establecimientos-2014/>
- Paleti, R., Bhat, C., y Pendyala, R. (2013). Integrated model of residential location, work location, vehicle ownership, and commute tour characteristics. **Transportation Research Record**, 2382, 162–172. <https://doi.org/10.3141/2382-18>
- Pinjari, A. R., Pendyala, R. M., Bhat, C. R., y Waddell, P. A. (2011). Modeling the choice continuum: An integrated model of residential location, auto ownership, bicycle ownership, and commute tour mode choice decisions. **Transportation**, 38(6), 933–958. <https://doi.org/10.1007/s11116-011-9360-y>
- SECTRA. (2014). *Actualización y recolección de información del sistema de transporte urbano, Etapa IX: Encuesta Origen Destino Santiago 2012*.
- SII. (2014). *Catastro de Bienes Raíces del Servicio de impuestos Internos*. https://zeus.sii.cl/cvc_cgi/dfmun/dfmun_repGobierno.cgi
- Stephen R. Lerman. (1976). Location, Housing, Automobile Ownership, and Mode To Work: a Joint Choice Model. **Transportation Research Record**, 610, 6–11. <http://pubsindex.trb.org/view.aspx?id=53665>
- Waddell, P. (1993). Models : Is the Assumption Valid ? **Geographical**, 25(1).
- Waddell, P., Bhat, C., Eluru, N., Wang, L., y Pendyala, R. M. (2007). Modeling interdependence in household residence and workplace choices. **Transportation Research Record**, 2003, 84–92. <https://doi.org/10.3141/2003-11>
- Zimmerman, S. D. (2019). Elite colleges and upward mobility to top jobs and top incomes. **American Economic Review**, 109(1), 1–47. <https://doi.org/10.1257/aer.20171019>