



UNIVERSIDAD DE CHILE
FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS
DEPARTAMENTO DE INGENIERÍA INDUSTRIAL
DEPARTAMENTO DE INGENIERÍA MATEMÁTICA

PATRONES DE MOVILIDAD ESCOLAR BAJO EL NUEVO SISTEMA DE ADMISIÓN

TESIS PARA OPTAR AL GRADO DE MAGÍSTER EN ECONOMÍA APLICADA
MEMORIA PARA OPTAR AL TÍTULO DE INGENIERO CIVIL MATEMÁTICO

JOAQUÍN ESTEBAN FARÍÑA VALDEBENITO

PROFESOR GUÍA:
ALEJANDRA MIZALA SALCES

MIEMBROS DE LA COMISIÓN:
JORGE AMAYA ARRIAGADA
PATRICIO RODRÍGUEZ VALDÉS

Este trabajo ha sido parcialmente financiado por el Centro de Investigación Avanzada en Educación (CIAE) y el proyecto FB0003 del Programa de Investigación Asociativa de CONICYT

SANTIAGO DE CHILE
2020

RESUMEN DE LA MEMORIA PARA OPTAR
AL TÍTULO DE: Ingeniero Civil Matemático y grado de
Magíster en Economía Aplicada
POR: Joaquín Esteban Fariña Valdebenito
FECHA: 29/01/2020
PROFESORA GUÍA: Alejandra Mizala Salces

PATRONES DE MOVILIDAD ESCOLAR BAJO EL NUEVO SISTEMA DE ADMISIÓN

Chile es uno de los países con mayor segregación escolar de la OCDE (OCDE (2018)). El nuevo Sistema de Admisión Escolar (SAE), que se ha implementado gradualmente desde 2015, busca orientar el sistema hacia uno más inclusivo y equitativo (Biblioteca del Congreso Nacional de Chile, 2015). Sin embargo, múltiples polémicas se han instalado en la opinión pública sobre su funcionamiento en términos de justicia y eficiencia (Carrasco y Honey, 2019). Por ende, surge la interrogante sobre los efectos que este nuevo sistema está teniendo sobre la movilidad escolar.

Esta tesis busca caracterizar la elección escolar en términos de variables de establecimientos educacionales y estudiantes, y analizar cómo el escenario de cambio de escuelas se ha modificado producto de la implementación del SAE. Para lo cual, se indaga en los patrones de movilidad escolar entre los años 2013 y 2019 en las aglomeraciones urbanas de Iquique - Alto Hospicio, Coquimbo - La Serena, Rancagua - Machalí, Osorno - Puerto Montt, y Punta Arenas, en los niveles de sexto y octavo básico. Más aún, se estudian los casos en los que el cambio es forzado (establecimiento no ofrece siguiente nivel) y aquellos en que no lo es (el cambio ocurre aunque la escuela ofrezca el próximo nivel).

Específicamente, se trabaja con una nueva metodología, que a partir de las características los establecimientos de origen y destino de los estudiantes que cambian de escuela, identifica patrones, o *clusters*, significativos. Para identificar los efectos del SAE, estos *clusters* se analizan en tres escenarios: (i) antes del SAE, (ii) post SAE considerando las preferencias declaradas de las familias, y (iii) post SAE con respecto a la asignación que realiza el sistema.

Los principales resultados evidencian que la dependencia administrativa es el atributo más relevante en la elección de escuela. Además, se observa que hay una demanda por EE con mejores resultados SIMCE que excede la capacidad del sistema.

En general, cuando los cambios son No Forzados, se observa un aumento de demanda y asignación a establecimientos particulares subvencionados.

Además, hay una alta correlación entre variables socioeconómicas de los estudiantes y el tipo de colegio que escogen: estudiantes con mayor Ingreso del Hogar, mayor Educación de los Padres, no Beneficiarios SEP, con mejores puntajes SIMCE, son los que componen y prefieren establecimientos con mayor rendimiento SIMCE, menos vulnerables (IVE) y de dependencia particular subvencionada.

Finalmente, se deduce que la sociedad se encuentra en un estado de adaptación al nuevo sistema de admisión. Si bien el SAE provee un escenario más justo en lo que respecta a *school choice*, los resultados muestran que la demanda en los escenarios SAE por ciertos tipos de establecimientos excede la oferta que el sistema escolar provee actualmente.

A mis padres, que me han regalado la libertad de soñar.

Agradecimientos

Se agradece al Centro de Investigación Avanzada en Educación (CIAE) y al proyecto FB0003 del Programa de Investigación Asociativa de CONICYT por el apoyo otorgado durante el desarrollo de esta tesis. En particular, quisiera agradecer a Alejandra Mizala por su guía, comprensión y apoyo durante todo este proceso, sin duda su sabiduría me ha convertido en un mejor profesional.

También agradezco a Patricio Rodríguez y Manuel Matas por sus valiosas contribuciones en la elaboración de las bases de datos necesarias para este estudio, así como en la elección y desarrollo de la metodología. Agradezco además a Jorge Amaya por sus valiosos comentarios e ideas, y muy especialmente por su comprensión y paciencia.

Agradezco a mi padres, Patricia y Rodolfo, por su entrega incondicional y compromiso con mi formación y desarrollo (y por insistirme que termine la tesis); a mis hermanos, a mi abuela y a todos los miembros de mi familia, los que siempre han demostrado un cariño y apoyo sincero en los procesos que he vivido.

También agradezco a Javiera, mi compañera incondicional durante este largo periodo de tesis, gracias por apoyarme en tantos momentos difíciles y por permitirme ser parte de tu vida.

Además, quisiera agradecer también a mis amigos, colegas, jefes, y a todos los que han compartido conmigo parte de su sabiduría a lo largo de mis años universitarios, especialmente a Leo, Joaquín, Rodrigo, Esteban, Chomi, Perro, Danilo, Robert, Sam, Saki, Simón, Ricardo, Alejandro, Vicente, Favio, Olivia, Seba G., Seba R., Marce, Juanpa, Toti, JF, Daniela, Teresa, Gina, Alex, Belén, Jose, Gabo, Ángelo, Roberto y Michelle.

Finalmente, agradezco profundamente a todos mis estudiantes del Colegio Claudio Matte de Viña del Mar y del Liceo Bicentenario Italia de Santiago. Gracias por darme un espacio en sus vidas, por permitirme acompañarlos en sus caminos y sobre todo por enseñarme que es posible, que el talento está en todas partes y que como sociedad tenemos una tremenda deuda con ustedes: que todos los estudiantes tengan las mismas oportunidades.

Tabla de Contenido

Índice de Tablas	ix
Índice de Figuras	xi
1 Introducción	1
2 Revisión Bibliográfica	4
2.1 Sistema Escolar Chileno	4
2.2 <i>School Choice</i> : Revisión de Antecedentes	6
2.3 Ley de Inclusión Escolar	8
2.3.1 Educación Subvencionada Gratuita	9
2.3.2 Nuevo Sistema de Admisión Escolar	10
2.3.2.1 Algoritmo de Asignación	11
2.3.2.2 Los Efectos del SAE	12
2.3.3 Fin al Lucro	13
3 Datos y Estadísticas Descriptivas	14
3.1 Variables a Nivel de Estudiantes	17
3.2 Variables a Nivel de Establecimientos	22
3.3 Movilidad Escolar	29
4 Metodología	32
4.1 Marco Conceptual	32
4.1.1 Ciencia de datos	32
4.1.2 Aprendizaje Automático	33
4.1.3 Algoritmos de <i>Clustering</i>	35
4.1.3.1 K-MEANS	35
4.1.3.2 Modelos Mixtos Gaussianos (GMM)	36
4.1.3.3 <i>Clustering</i> Jerárquico Aglomerativo (HAC)	38
4.1.4 Elección del Número de <i>Clusters</i>	40
4.1.4.1 El Método <i>Elbow</i>	41
4.1.4.2 El Método de Silueta	41
4.2 Implementación	43
5 Resultados	47
5.1 Iquique - Alto Hospicio	49
5.1.1 Cambios Forzados	49
5.1.1.1 Escenario Pre SAE	49
5.1.1.2 Escenario SAE	52

5.1.2	Cambios No Forzados	56
5.1.2.1	Escenario Pre SAE	56
5.1.2.2	Escenario SAE	59
5.2	Coquimbo - La Serena	65
5.2.1	Cambios Forzados	65
5.2.1.1	Escenario Pre SAE	65
5.2.1.2	Escenario SAE	67
5.2.2	Cambios No Forzados	72
5.2.2.1	Escenario Pre SAE	72
5.2.2.2	Escenario SAE	74
5.3	Rancagua - Machalí	80
5.3.1	Cambios Forzados	80
5.3.1.1	Escenario Pre SAE	80
5.3.1.2	Escenario SAE	82
5.3.2	Cambios No Forzados	88
5.3.2.1	Escenario Pre SAE	88
5.3.2.2	Escenario SAE	90
5.4	Puerto Montt - Puerto Varas	97
5.4.1	Cambios Forzados	97
5.4.1.1	Escenario Pre SAE	97
5.4.1.2	Escenario SAE	99
5.4.2	Cambios No Forzados	105
5.4.2.1	Escenario Pre SAE	105
5.4.2.2	Escenario SAE	108
6	Conclusiones	115
	Bibliografía	120
	Apéndices	129
A	Descripción de Variables	129
B	Representatividad de Variables	132
C	Tratamiento de Datos Faltantes	134
D	Optimización del Número de <i>Clusters</i>	134
D.1	Iquique - Alto Hospicio	134
D.2	Coquimbo - La Serena	139
D.3	Rancagua - Machalí	143
D.4	Puerto Montt - Puerto Varas	148
D.5	Punta Arenas	152
E	Resultados de Escenarios SAE por Año	155
F	Resultados por Aglomeración Urbana y Caso	160
G	Códigos	168

Índice de Tablas

3.1	Conurbaciones de las 4 regiones consideradas.	14
3.2	Valores de Ingreso Hogar según Dependencia Administrativa	18
3.3	Valores para la variable Estudio de Padres	20
3.4	Matrícula total por aglomeración por año.	22
3.5	Rendición de prueba SIMCE por nivel y año.	27
3.6	Disponibilidad de datos SAE por región.	30
4.1	Medida de disimilitud en modelos de <i>clustering</i> aglomerativo jerárquico.	39
4.2	Escenarios en los que se aplica el modelo.	44
5.1	Atributos para modelo de <i>clustering</i>	47
5.2	Número óptimo de <i>clusters</i> para cada caso estudiado	48
5.3	Centroides de <i>clusters</i>	49
5.4	Centroides de <i>clusters</i>	52
5.5	Centroides de <i>clusters</i>	54
5.6	Centroides de <i>clusters</i>	57
5.7	Centroides de <i>clusters</i>	59
5.8	Centroides de <i>clusters</i>	61
5.9	Centroides de <i>clusters</i>	65
5.10	Centroides de <i>clusters</i>	67
5.11	Centroides de <i>clusters</i>	70
5.12	Centroides de <i>clusters</i>	73
5.13	Centroides de <i>clusters</i>	74
5.14	Centroides de <i>clusters</i>	77
5.15	Centroides de <i>clusters</i>	80
5.16	Centroides de <i>clusters</i>	83
5.17	Centroides de <i>clusters</i>	85
5.18	Centroides de <i>clusters</i>	88
5.19	Centroides de <i>clusters</i>	91
5.20	Centroides de <i>clusters</i>	93
5.21	Centroides de <i>clusters</i>	98
5.22	Centroides de <i>clusters</i>	100
5.23	Centroides de <i>clusters</i>	102
5.24	Centroides de <i>clusters</i>	105
5.25	Centroides de <i>clusters</i>	108
5.26	Centroides de <i>clusters</i>	111
6.1	Variables de establecimientos.	129

6.2	Variables de estudiantes.	130
6.3	Variables de cambio de escuela.	131
6.4	Número óptimo de <i>clusters</i> para cada caso estudiado	154
6.5	Calidad de resultados de <i>clustering</i> considerando Inercia y Coeficiente de Silueta.154	

Índice de Figuras

2.1	Número de EE según dependencia administrativa en periodo 1990-2006	5
3.1	Distribución de estudiantes por género.	17
3.2	Histograma de promedio general de notas estudiantes.	18
3.3	Beneficiarios SEP por año según dependencia administrativa y región.	19
3.4	Distribución del ingreso del hogar de estudiantes según dep. administrativa. .	19
3.5	Distribución del nivel de estudio de padres según dependencia administrativa.	21
3.6	Distancia hogar-colegio según dependencia administrativa.	21
3.7	Matrícula total según dependencia administrativa por año.	22
3.8	Distribución de EE según dependencia administrativa por año.	23
3.9	Establecimientos y matrícula que cambia de dependencia administrativa en Iquique - Alto Hospicio.	24
3.10	Establecimientos y matrícula que cambia de dependencia administrativa en Coquimbo - La Serena.	24
3.11	Establecimientos y matrícula que cambia de dependencia administrativa en Rancagua - Machalí.	24
3.12	Establecimientos y matrícula que cambia de dependencia administrativa en Puerto Montt - Puerto Varas.	25
3.13	Distribución de copago en EE particulares subvencionados por año.	25
3.14	Distribución de copago en EE particulares subvencionados por año.	26
3.15	Distribución de IVE según dependencia administrativa por año.	27
3.16	Puntaje promedio SIMCE según dependencia administrativa por año.	28
3.17	Distribución de establecimientos según GSE y dependencia administrativa. .	28
3.18	Cambio de escuela (%) con respecto a matrícula total según nivel, dependencia administrativa y año. Casos forzado y no forzado.	30
3.19	Número de postulantes por nivel de ingreso.	31
4.1	Ejemplo de entrenamiento de K-MEANS. Cada iteración reasigna puntos hasta convergencia.	36
4.2	Modelo Gausiano Mixto de dos componentes: datos y superficies equiprobables.	39
4.3	Ejemplo de dendograma para cada medida de disimilitud.	40
4.4	Ejemplo de Método <i>elbow</i>	41
4.5	Ejemplo de Puntaje de silueta para todo los datos.	42
4.6	Puntaje de silueta con respecto a cada <i>cluster</i>	43
5.1	Diagrama de resultados para cada una de las cuatro aglomeraciones urbanas consideradas.	49
5.2	Variables a nivel de establecimiento.	50

5.3	Variables a nivel de estudiante.	51
5.4	Variables a nivel de establecimiento.	53
5.5	Variables a nivel de estudiante.	54
5.6	Variables a nivel de establecimiento.	55
5.7	Variables a nivel de estudiante.	56
5.8	Variables a nivel de establecimiento.	57
5.9	Variables a nivel de estudiante.	58
5.10	Variables a nivel de establecimiento.	60
5.11	Variables a nivel de estudiante.	61
5.12	Variables a nivel de establecimiento.	62
5.13	Variables a nivel de estudiante.	63
5.14	Variables a nivel de establecimiento.	66
5.15	Variables a nivel de estudiante.	67
5.16	Variables a nivel de establecimiento.	68
5.17	Variables a nivel de estudiante.	69
5.18	Variables a nivel de establecimiento.	70
5.19	Variables a nivel de estudiante.	71
5.20	Variables a nivel de establecimiento.	73
5.21	Variables a nivel de estudiante.	73
5.22	Variables a nivel de establecimiento.	75
5.23	Variables a nivel de estudiante.	76
5.24	Variables a nivel de establecimiento.	78
5.25	Variables a nivel de estudiante.	79
5.26	Variables a nivel de establecimiento.	81
5.27	Variables a nivel de estudiante.	82
5.28	Variables a nivel de establecimiento.	83
5.29	Variables a nivel de estudiante.	84
5.30	Variables a nivel de establecimiento.	86
5.31	Variables a nivel de estudiante.	87
5.32	Variables a nivel de establecimiento.	89
5.33	Variables a nivel de estudiante.	90
5.34	Variables a nivel de establecimiento.	91
5.35	Variables a nivel de estudiante.	92
5.36	Variables a nivel de establecimiento.	94
5.37	Variables a nivel de estudiante.	95
5.38	Variables a nivel de establecimiento.	97
5.39	Variables a nivel de estudiante.	98
5.40	Variables a nivel de establecimiento.	100
5.41	Variables a nivel de estudiante.	101
5.42	Variables a nivel de establecimiento.	103
5.43	Variables a nivel de estudiante.	104
5.44	Variables a nivel de establecimiento.	106
5.45	Variables a nivel de estudiante.	107
5.46	Variables a nivel de establecimiento.	109
5.47	Variables a nivel de estudiante.	110
5.48	Variables a nivel de establecimiento.	111
5.49	Variables a nivel de estudiante.	112

6.1	Inercia y coeficiente de silueta. Caso Iquique - Alto Hospicio, cambios Forzados, Pre SAE.	134
6.2	Coeficiente de silueta por <i>cluster</i> . Caso óptimo. Iquique - Alto Hospicio, cambios Forzados, Pre SAE.	135
6.3	Inercia y coeficiente de silueta. Caso Iquique - Alto Hospicio, cambios Forzados, SAE Preferencias.	135
6.4	Coeficiente de silueta por <i>cluster</i> . Caso óptimo. Iquique - Alto Hospicio, cambios Forzados, SAE Preferencias.	135
6.5	Inercia y coeficiente de silueta. Caso Iquique - Alto Hospicio, cambios Forzados, SAE Admisión.	136
6.6	Coeficiente de silueta por <i>cluster</i> . Caso óptimo. Iquique - Alto Hospicio, cambios Forzados, SAE Admisión.	136
6.7	Inercia y coeficiente de silueta. Caso Iquique - Alto Hospicio, cambios No Forzados, Pre SAE.	136
6.8	Coeficiente de silueta por <i>cluster</i> . Caso óptimo. Iquique - Alto Hospicio, cambios No Forzados, Pre SAE.	137
6.9	Inercia y coeficiente de silueta. Caso Iquique - Alto Hospicio, cambios No Forzados, SAE Preferencias.	137
6.10	Coeficiente de silueta por <i>cluster</i> . Caso óptimo. Iquique - Alto Hospicio, cambios No Forzados, SAE Preferencias.	137
6.11	Inercia y coeficiente de silueta. Caso Iquique - Alto Hospicio, cambios No Forzados, SAE Admisión.	138
6.12	Coeficiente de silueta por <i>cluster</i> . Caso óptimo. Iquique - Alto Hospicio, cambios No Forzados, SAE Admisión.	138
6.13	Inercia y coeficiente de silueta. Caso Coquimbo - La Serena, cambios Forzados, Pre SAE.	139
6.14	Coeficiente de silueta por <i>cluster</i> . Caso óptimo. Coquimbo - La Serena, cambios Forzados, Pre SAE.	139
6.15	Inercia y coeficiente de silueta. Caso Coquimbo - La Serena, cambios Forzados, SAE Preferencias.	140
6.16	Coeficiente de silueta por <i>cluster</i> . Caso óptimo. Coquimbo - La Serena, cambios Forzados, SAE Preferencias.	140
6.17	Inercia y coeficiente de silueta. Caso Coquimbo - La Serena, cambios Forzados, SAE Admisión.	140
6.18	Coeficiente de silueta por <i>cluster</i> . Caso óptimo. Coquimbo - La Serena, cambios Forzados, SAE Admisión.	141
6.19	Inercia y coeficiente de silueta. Caso Coquimbo - La Serena, cambios No Forzados, Pre SAE.	141
6.20	Coeficiente de silueta por <i>cluster</i> . Caso óptimo. Coquimbo - La Serena, cambios No Forzados, Pre SAE.	141
6.21	Inercia y coeficiente de silueta. Caso Coquimbo - La Serena, cambios No Forzados, SAE Preferencias.	142
6.22	Coeficiente de silueta por <i>cluster</i> . Caso óptimo. Coquimbo - La Serena, cambios No Forzados, SAE Preferencias.	142
6.23	Inercia y coeficiente de silueta. Caso Coquimbo - La Serena, cambios No Forzados, SAE Admisión.	142

6.24	Coeficiente de silueta por <i>cluster</i> . Caso óptimo. Coquimbo - La Serena, cambios No Forzados, SAE Admisión.	143
6.25	Inercia y coeficiente de silueta. Caso Rancagua - Machalí, cambios Forzados, Pre SAE.	143
6.26	Coeficiente de silueta por <i>cluster</i> . Caso óptimo. Rancagua - Machalí, cambios Forzados, Pre SAE.	144
6.27	Inercia y coeficiente de silueta. Caso Rancagua - Machalí, cambios Forzados, SAE Preferencias.	144
6.28	Coeficiente de silueta por <i>cluster</i> . Caso óptimo. Rancagua - Machalí, cambios Forzados, SAE Preferencias.	144
6.29	Inercia y coeficiente de silueta. Caso Rancagua - Machalí, cambios Forzados, SAE Admisión.	145
6.30	Coeficiente de silueta por <i>cluster</i> . Caso óptimo. Rancagua - Machalí, cambios Forzados, SAE Admisión.	145
6.31	Inercia y coeficiente de silueta. Caso Rancagua - Machalí, cambios No Forzados, Pre SAE.	145
6.32	Coeficiente de silueta por <i>cluster</i> . Caso óptimo. Rancagua - Machalí, cambios No Forzados, Pre SAE.	146
6.33	Inercia y coeficiente de silueta. Caso Rancagua - Machalí, cambios No Forzados, SAE Preferencias.	146
6.34	Coeficiente de silueta por <i>cluster</i> . Caso óptimo. Rancagua - Machalí, cambios No Forzados, SAE Preferencias.	146
6.35	Inercia y coeficiente de silueta. Caso Rancagua - Machalí, cambios No Forzados, SAE Admisión.	147
6.36	Coeficiente de silueta por <i>cluster</i> . Caso óptimo. Rancagua - Machalí, cambios No Forzados, SAE Admisión.	147
6.37	Inercia y coeficiente de silueta. Caso Puerto Montt - Puerto Varas, cambios Forzados, Pre SAE.	148
6.38	Coeficiente de silueta por <i>cluster</i> . Caso óptimo. Puerto Montt - Puerto Varas, cambios Forzados, Pre SAE.	148
6.39	Inercia y coeficiente de silueta. Caso Puerto Montt - Puerto Varas, cambios Forzados, SAE Preferencias.	148
6.40	Coeficiente de silueta por <i>cluster</i> . Caso óptimo. Puerto Montt - Puerto Varas, cambios Forzados, SAE Preferencias.	149
6.41	Inercia y coeficiente de silueta. Caso Puerto Montt - Puerto Varas, cambios Forzados, SAE Admisión.	149
6.42	Coeficiente de silueta por <i>cluster</i> . Caso óptimo. Puerto Montt - Puerto Varas, cambios Forzados, SAE Admisión.	149
6.43	Inercia y coeficiente de silueta. Caso Puerto Montt - Puerto Varas, cambios No Forzados, Pre SAE.	150
6.44	Coeficiente de silueta por <i>cluster</i> . Caso óptimo. Puerto Montt - Puerto Varas, cambios No Forzados, Pre SAE.	150
6.45	Inercia y coeficiente de silueta. Caso Puerto Montt - Puerto Varas, cambios No Forzados, SAE Preferencias.	150
6.46	Coeficiente de silueta por <i>cluster</i> . Caso óptimo. Puerto Montt - Puerto Varas, cambios No Forzados, SAE Preferencias.	151

6.47	Inercia y coeficiente de silueta. Caso Puerto Montt - Puerto Varas, cambios No Forzados, SAE Admisión.	151
6.48	Coefficiente de silueta por <i>cluster</i> . Caso óptimo. Puerto Montt - Puerto Varas, cambios No Forzados, SAE Admisión.	151
6.49	Inercia y coeficiente de silueta. Caso Punta Arenas, cambios Forzados, Pre SAE.	152
6.50	Inercia y coeficiente de silueta. Caso Punta Arenas, cambios Forzados, SAE Preferencias.	152
6.51	Inercia y coeficiente de silueta. Caso Punta Arenas, cambios Forzados, SAE Admisión.	153
6.52	Inercia y coeficiente de silueta. Caso Punta Arenas, cambios No Forzados, Pre SAE.	153
6.53	Inercia y coeficiente de silueta. Caso Punta Arenas, cambios No Forzados, SAE Preferencias.	153
6.54	Inercia y coeficiente de silueta. Caso Punta Arenas, cambios No Forzados, SAE Admisión.	154
6.55	Comparación de comportamiento de la movilidad escolar, entre los procesos SAE de admisión 2018 y 2019, según aglomeración, criterio y tipos de cambio estudiado.	155

Capítulo 1

Introducción

La sociedad chilena presenta uno de los índices de segregación más altos del mundo (OCDE, 2018), siendo su sistema escolar una gran fuente de reproducción de la inequidad. El sistema escolar presenta altos niveles de segregación por nivel socioeconómico (Mizala y Torche, 2012; Gutiérrez et al., 2017), la que se expresa en la composición socioeconómica de los establecimientos educativos (EE) y como consecuencia de ello en sus resultados académicos (MINEDUC, 2016).

Frente a esto, el año 2015 se aprueba la Ley de Inclusión Escolar (Biblioteca del Congreso Nacional de Chile, 2015), la cual, entre otras cosas, crea el Sistema de Admisión Escolar (SAE), el que gestiona de manera centralizada la admisión escolar. Este mecanismo asigna estudiantes a establecimientos educacionales según las preferencias declaradas por las familias, lo que elimina la selección por parte de los colegios. Este sistema conlleva una modificación importante en la demanda educacional y sobre todo en la forma en la que las familias escogen colegios para sus hijos. Además, la ley de inclusión elimina el copago que realizan las familias y prohíbe el lucro en EE que reciban aportes del estado. Desde su implementación en 2015, se ha instalado en la opinión pública un debate sobre la efectividad del SAE en términos de segregación y justicia^{1 2}.

Algunos trabajos recientes han estudiado los efectos que ha tenido el SAE en el cambio de escuelas. Eyzaguirre et al. (2019a) identifican diferencias en las preferencias de estudiantes prioritarios³, los que postulan más a EE municipales, sin copago, con mayor Índice de Vulnerabilidad Escolar (IVE) y peores resultados académicos (pruebas SIMCE⁴). Por otro lado, Carrasco y Honey (2019) muestran que si bien hay un leve aumento en el acceso a escuelas de alta calidad, los patrones de desigualdad persisten, señalando además factores extrínsecos, como la oferta territorial disponible y el aumento de subvención a estudiantes prioritario, que limitan los efectos esperados del SAE. En este sentido, es interesante estudiar la manera en que los padres escogen escuelas en Chile y si la elección cambia (o no) bajo el

¹<http://www.facso.uchile.cl/noticias/136328/las-consecuencias-de-la-ley-de-inclusion-escolar>

²<https://www.ucsc.cl/blogs-academicos/los-desafios-inmediatos-de-la-ley-de-inclusion-escolar/>

³Los estudiantes prioritarios son los que pertenecen a familias que se encuentran dentro del 40 por ciento más vulnerables del país.

⁴Prueba estandarizada chilena.

nuevo escenario del SAE.

Un número importante de trabajos han estudiado los factores que son relevantes en la elección de colegios por parte de las familias, los que reconocen la importancia de atributos propios de los hogares y los estudiantes (Sanclemente, 2008; Makovec et al., 2010; Lara et al., 2011), de los establecimientos, y de la relación propia estudiante-colegio (Gallego y Hernando, 2009; Lara et al., 2011). No obstante, ninguno se ha enfocado en determinar los patrones (*clusters*) que se observan en los cambios de escuela, desde una perspectiva multidimensional. En este sentido, analizar el cambio de colegios a partir de los *clusters*, vistos como perfiles de cambio, y como estos se modifican, permitiría profundizar en la problemática desde una dimensión no abordada con anterioridad.

Asimismo, el aspecto geográfico del cambio de escuelas tampoco ha sido suficientemente estudiado. Urzua (2019), analiza los efectos del SAE en las regiones de Biobío y Valparaíso, encontrando un aumento en la segregación escolar, lo que difiere de los resultados de Carrasco y Honey (2019). Estas diferencias resalta la relevancia de estudiar cada territorio por su propia cuenta, antes de sacar conclusiones sobre los efectos del SAE en términos generales. Por otro lado, Rodríguez et al. (2016) determinan el acceso a escuelas efectivas en 22 ciudades del país, dando cuenta de la relevancia del factor geográfico a la hora de focalizar políticas públicas, aunque su énfasis no está puesto en *school choice*.⁵

Por otro lado, aún no existe suficiente evidencia sobre el impacto del nuevo sistema de admisión en dos dimensiones fundamentales: (i) el contraste del nuevo escenario educativo con respecto a la elección de escuelas previo al SAE, y (ii) análisis de sus efectos a nivel regional y comunal. Lo primero permite identificar de qué manera el nuevo proceso de elección de EE ha mostrado ser distinto bajo el régimen SAE en sus primeros años de aplicación, mientras que la consideración de la dimensión geográfica ayuda a analizar cada territorio en base a su propia realidad, dada la heterogeneidad del territorio chileno este análisis mostrará diferencias locales importantes a la hora de proponer mejoras al sistema.

Considerando los antecedentes anteriores, el objetivo de este trabajo es caracterizar la elección escolar en términos de atributos de EE y estudiantes, y analizar cómo este escenario se ha modificado producto de la implementación del SAE. El estudio se focaliza en cinco aglomeraciones urbanas del país: Iquique - Alto Hospicio, Coquimbo - La Serena, Rancagua - Machalí, Osorno - Puerto Montt, y Punta Arenas. Esto para proveer evidencia sobre el impacto de esta nueva ley, aportar a la comprensión de los cambios de la demanda por colegios, y del nuevo escenario educativo escolar.

Específicamente, esta investigación busca: (i) identificar factores relevantes en el cambio de escuelas en Chile, (ii) entrenar un modelo que identifique patrones en los cambios de escuela en la situación previa al SAE y en el nuevo escenario con SAE, (iii) analizar las diferencias entre la situación pre y post SAE, y (iv) generar recomendaciones de política pública para mejorar el sistema.

Este trabajo, por una parte, tiene relevancia teórica, pues entrega evidencia sobre el funcionamiento del modelo de educación de mercado planteado por Friedman (1962) y Chubb y

⁵Rama de la investigación educativa que estudia movilidad de estudiantes en un sistema escolar.

Moe (1991), el cual aboga por la libre oferta y demanda educativa. Desde un análisis teórico, el modelo propone que los EE compiten por las preferencias de las familias, las que escogen colegio para sus hijos; esta competencia incentiva a las escuelas a mejorar para mantener y atraer estudiantes. Bajo este nuevo escenario sin selección por parte de los colegios, por primera vez existen las condiciones para que las familias actúan según sus verdaderas preferencias, lo cual acerca la realidad a los supuesto teóricos del modelo, lo que permite evaluar con mayor precisión su operación. No obstante, fenómenos de autoselección⁶ están presentes en la sociedad chilena, por lo que este factor debe ser considerada en el posterior análisis.

Por otro lado, en aspectos metodológicos, esta investigación aporta con un nuevo enfoque en el estudio de movilidad escolar, a través de la implementación de un modelo de aprendizaje no supervisado para la identificación de patrones en cambios de escuela. Esto va en línea con Kleinberg et al. (2015), quienes promueven el uso de herramientas de ciencia de datos para abordar problemáticas de política pública, siendo este trabajo una adecuación de dichas herramientas en política pública educativa.

Finalmente, en un sentido práctico, el conocer a nivel regional los patrones de cambios de escuela según características de EE y estudiantes, así como las nuevas tendencias introducidas por el SAE, aportará a la generación de políticas públicas que fortalezcan el sistema escolar y su equilibrio en términos de oferta y demanda educativa.

El resto de esta tesis está organizada de la siguiente manera: el Capítulo 2 presenta la revisión de la literatura sobre el tema, el Capítulo 3 expone en detalle los datos utilizados y su estadísticas descriptivas; posteriormente, se describe la metodología en el Capítulo 4, y los resultados y su análisis se exponen en el capítulo 5. Finalmente el capítulo 5 presenta la discusión y conclusiones. Además, en el apartado final se encuentran las referencias bibliográficas y apéndices que ofrecen información complementaria.

⁶Estos se discuten en el siguiente capítulo.

Capítulo 2

Revisión Bibliográfica

Siguiendo los objetivos planteados, resulta de interés abordar la literatura en tres ámbitos: caracterización del modelo educacional chileno; elección de escuela (*school choice*), con enfoque en el caso de Chile; y finalmente, la ley de inclusión escolar, haciendo énfasis en el nuevo sistema de admisión escolar (SAE).

2.1. Sistema Escolar Chileno

La estructura actual del sistema escolar en Chile tiene sus inicios en la década de 1980, periodo en el que sufrió fuertes modificaciones, definiendo nuevas pautas de financiamiento y administración, muy en línea con lo que propone Friedman (1955) delegación por parte del estado de la educación hacia entidades privadas con fines de lucro, lo que transforma a la educación en un cuasi mercado (Chubb y Moe, 1991).

Se implementa un nuevo sistema de financiamiento basado en la subvención por estudiante (*vouchers*), mediante el cual los administradores de establecimientos educativos reciben una cierta cantidad de dinero por cada estudiante que asiste a clases. Esto aplica tanto para los EE municipales como para los subvencionados. Simultáneamente, la administración de colegios públicos pasó a manos de las municipalidades, concretando así la descentralización y desvinculación del Estado en lo que respecta a educación escolar.

Como resultado de estas políticas educativas, tres tipos de establecimientos (según dependencia administrativa) conforman el sistema: municipales, particulares subvencionados y particulares privados. Los dos primeros reciben financiamiento estatal, mientras que el último no reciben ningún tipo de subsidio.

Uno de los supuestos del sistema de *vouchers* es la competencia que se debería generar entre EE para atraer estudiantes, lo que llevaría a los colegios a mejorar su calidad. Siguiendo esta perspectiva de mercado, las escuelas deberían competir por las preferencias de los padres quedando en manos de las familias el poder escoger para sus hijos aquellos establecimientos que maximicen su bienestar.

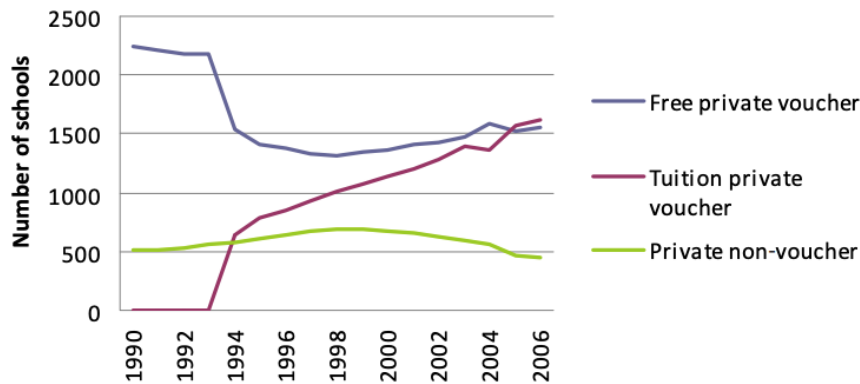


Figura 2.1: Número de EE según dependencia administrativa en periodo 1990-2006
Fuente:Elacqua (2009)

En este sentido, con la finalidad de poder medir la calidad de la educación que cada establecimiento entrega y así fomentar la competencia entre ellos (Inzunza y Campos, 2016), se implementa en 1988 la prueba Sistema Nacional de Medición de la Calidad de la Educación (SIMCE), evaluación estandarizada de carácter censal que entrega evidencia sobre la calidad promedio de cada EE en términos de aprendizaje (Cox, 1997).

El mismo año, se aumenta fuertemente la subvención y la inversión pública en materia educativa. Además, en 1993 se amplía la posibilidad de cobrar a los padres por sobre la subvención que el estado entrega (financiamiento compartido) esto se aplica para los colegios particulares subvencionados y solo para la enseñanza media en los colegios municipales, aunque muy pocos implementan el cobro. Los colegios que cobran un copago mensual a los apoderados mantienen la subvención del estado aunque el monto se reduce levemente de acuerdo a una tabla.

En este nuevo escenario educativo, la cantidad de colegios particulares subvencionados creció enormemente, pasando de abarcar un 15 % de la matrícula escolar previo a la reforma (CASEN, 2002), al 42 % de la matrícula nacional en 2005 y al 60 % el año 2014 (Canales et al., 2016). Esto en desmedro de la matrícula en colegios municipales, que cae de aproximadamente del 73 % en 1981 al 49 % en 2005 y a menos del 40 % actualmente. Dicha tendencia se observa en la Figura 2.1.

El diseño del sistema de subsidio a la demanda contemplaba que la libertad de elección por parte de los padres causaría competencia entre EE, y mejoras en la calidad de la educación ofrecida, pues los padres rechazarían colegios con bajos resultados, los que terminarían saliendo del mercado (Lara et al., 2011).

Sin embargo, al comparar escuelas subvencionadas con municipales, los efectos de mejora en desempeño académico son pequeños e incluso no significativos, si se controla por nivel socioeconómico (Lara et al., 2011). Más aún, en oposición a lo esperado, varios trabajos concluyen que uno de los principales supuestos del sistema de *vouchers*, esto es, que la competencia por estudiantes entre establecimientos conlleva un aumento en la calidad de la educación impartida, no se cumple (Contreras et al., 2010). Es decir, las diferencias numéricas que existen no son reales, pues son explicadas por el capital cultural del estudiante.

Además, se produce una relevante la segregación por NSE (Villalobos y Valenzuela, 2012), pues los estudiantes de niveles socioeconómicos (NSE) bajos y medios (en parte) se concentran en EE municipales, los de NSE medio lo hacen en EE particular subvencionados y aquellos estudiantes de familias de NSE alto se ubican primariamente en EE privados (Hsieh y Urquiola, 2006). Esto conlleva la reproducción de la estructura social a través del sistema escolar.

Esto también se refleja en términos académicos, establecimientos particulares pagados obtienen los puntajes más altos, mientras que los colegios municipales los más bajos. En efecto, Mizala y Torche (2012) revelan una elevada correlación entre calidad del colegio y el ingreso del hogar de sus alumnos, lo que reafirma la existencia de una alta estratificación del sistema escolar chileno. Esto implica, en términos de selección, que las escuelas en vez de competir por mejorar la calidad, compiten seleccionando alumnos de mayor NSE. De esta manera, la configuración del sistema escolar chileno, y en especial el subsidio a la demanda, promueve la segregación socioeconómica.

Por otra parte, entra en vigencia el año 2008 la ley de Subvención Escolar Preferencial (SEP) (Biblioteca del Congreso Nacional de Chile, 2014), la que aumenta la subvención para el 40 % de los estudiantes más vulnerables y entrega un monto adicional a EE con mayor concentración de éstos. Ley que tiene por objetivo contribuir a la igualdad de oportunidades mejorando la equidad y calidad de la educación, considerando el mayor costo de educar a estudiantes provenientes de familias con menos recursos (Mizala y Torche, 2013). Aunque inicialmente se consideraron sólo los niveles desde prekínder a octavo básico¹, a partir del año 2012, el beneficio se comienza a extender de manera progresiva hasta cuarto medio.

2.2. *School Choice*: Revisión de Antecedentes

El cambio de escuela es un tema ampliamente investigado en educación, teniendo sus orígenes en Friedman (1955, 1962). Este estudia la manera de entregar a los padres la elección del colegio que deseen para su hijo. Evidencia nacional e internacional muestra que familias con mayores recursos tienen una amplia capacidad de elegir, mientras que en sectores vulnerables, esta de elección nunca ha sido realmente libre (Abdulkadiroğlu y Sonmez, 2003). Es interesante entonces, estudiar los factores que inciden en la movilidad escolar, con foco especial en el caso chileno.

Numerosos autores han investigado los determinantes de la elección de escuelas en Chile. En primera instancia, se ha demostrado que la distancia del hogar a la escuela (Elacqua, 2005; Gallego y Hernando, 2009) y calidad del establecimiento² (Gallego y Hernando (2009), Chumacero et al. (2011)) son fundamentales a la hora de elegir. En términos de calidad, EE con menor desempeño están asociados con una mayor tasa de cambio (Paredes et al., 2013), mientras que familias con mayores recursos están dispuestas a recorrer mayores distancias. Más aún, existe una *tradeoff* entre calidad y distancia al EE (Gómez et al., 2012).

¹El sistema escolar chileno considera doce niveles divididos en enseñanza básica (de primero a octavo) y enseñanza media (de primero a cuarto).

²Medido según pruebas estandarizadas.

Factores económicos y administrativos también han mostrado influir en la decisión de cambio de EE. Paredes et al. (2013) constata que frente a opciones de menor costo, aumenta la probabilidad de cambio de EE. Por otro lado, se ha mostrado que estudiantes en colegios particulares subvencionados tienen mayor probabilidad de permanecer en esta dependencia administrativa en contraste al sector municipal (Makovec et al., 2010).

En esta misma línea, Canals (2015) identifica diferencias en la tasa de cambio según rendimiento académico de los estudiantes, los grados de enseñanza en que se encuentran, los resultados SIMCE de los EE, sus dependencias administrativas y distribución geográficas. Constata además que la distancia hogar-colegio aumenta con la progresión académica y que aquellos que se cambian son los que también viven más lejos de sus colegios (y por ende, tienen un *choice set* más grande), similar conclusión obtiene (Makovec et al., 2010).

En términos socioeconómicos, Córdoba (2006) discute la elección de escuelas en familias de menores recursos, encontrando que las principales razones para elegir colegio son la cercanía, la condición de gratuidad y la percepción de calidad, está última siendo evaluada de manera subjetiva por las familias, considerando criterios relacionados con nivel de aprendizaje, cuerpo docente, orden y disciplina, e infraestructura.

La disponibilidad de información también ha tenido un rol en la elección de escuelas en Chile. Gómez et al. (2012) evalúa el efecto de hacer pública la información sobre calidad de EE en la elección de escuelas por parte de los padres. Comparando la elección de colegio en 1996 (información escasa) con la de 2003 (información muy difundida), encuentra que la calidad de los EE es un importante factor en la elección de EE, pasando a ser aún más relevante a partir del momento que se hace pública la información sobre calidad. Sin embargo, incluso antes de hacer públicos los resultados SIMCE, los padres actuaban como si los conocieran, lo que sugiere la existencia de variables observables por los padres que tienen alta correlación con la calidad de los EE, medida por pruebas estandarizadas.

En lo que respecta a selección, mientras sólo en 5% de EE públicos hay algún tipo de prueba de ingreso, el 48% de los estudiantes de establecimientos particulares subvencionados lo debe rendir una prueba de ingreso (Contreras et al., 2010), lo que contribuye a la segregación del sistema según se expuso en la sección precedente.

A pesar de los argumentos anteriores, Elacqua (2005) muestra que, si bien las familias argumentan que les importa la calidad del EE, esto no se condice con la manera en la que escogen en la realidad, lo que indica la presencia de un fenómeno de autoselección por parte de los hogares. Asimismo, Hernández y Raczynski (2015) advierte que tanto estratos socioeconómicos bajo como medios provocan la segregación que el sistema evidencia, planteando que las familias de dichas clases tienen un concepto de calidad más amplio que los meros resultados de pruebas estandarizadas, y además manejan poca información objetiva de las escuelas, apareciendo los círculos y redes sociales como principales medios de información.

El mismo autor identifica una dinámica de segmentación social a través la elección de escuelas, mediante el cobro de mensualidad y la predominante mirada negativa de establecimientos municipales, lo que lleva a las familias de estratos bajos a adoptar conductas de autoexclusión y segregación, dadas sus evidentes restricciones en la toma de decisiones. Asimismo, Canales et al. (2016), muestra cualitativamente que las familias de clase media en

Chile escogen colegios basándose principalmente en criterios sociales, tratando de desmarcarse de grupos socioeconómicos inferiores. Esto conlleva el descarte de escuelas públicas, gratuitas y sin selección.

Tomando en cuenta todo lo anterior, Gallego y Hernando (2009) mediante el modelo de utilidad aleatoria propuesta por McFadden (1974) y la especificación de Berry et al. (1995), determinan que los atributos más valorados son la calidad del colegio (medida por el SIMCE) y la distancia a este. Concluyen además que la selección ocurre más por el lado de las familias que por lado de los colegios, apoyando la idea de una autoselección y la consiguiente segregación socioeconómica que sufre el sistema educativo chileno.

Lo anterior sugiere una reducción en la presión sobre los EE por mejorar su calidad, lo que podría incluso aumentar la segregación del sistema. Se hace así necesario avanzar hacia un diseño sistémico que termine con el financiamiento compartido y la selección por parte de escuelas y liceos, junto con un fortalecimiento de la educación pública.

Pese a lo anterior, vale señalar que el diseño de políticas que fomentan la elección de los padres ha sido fructífera en ciertos ámbitos: hay mayor búsqueda de información, se recorren mayores distancias y hay mayor disposición a invertir. No obstante, la respuesta por parte de los EE ha sido mucho más lenta, siendo muy bajo el acceso a colegios de calidad (Elacqua y Martínez (2015)), sobre todo en sectores vulnerables (Rodríguez et al. (2016)).

El análisis de los cambios de estudiantes entre colegios es importante por los efectos a nivel sistémico que estos pueden ocasionar. De esta manera, la literatura de *school choice* plantea diversos factores que deben considerarse a la hora de estudiar el cambio de escuelas.

2.3. Ley de Inclusión Escolar

El sistema educativo chileno presenta una elevada segregación por NSE en las escuelas, medida según el Índice de Segregación de Duncan (Valenzuela et al., 2013).

Una serie de trabajos han identificado los costos de esta segregación escolar en Chile, los que incluyen baja cohesión y movilidad social, y falta de igualdad de oportunidades (Bellei, 2013).

En un contexto de descontento general con la educación chilena, la ex presidenta Michelle Bachelet, durante su segundo gobierno, impulsa un gran número de propuestas legislativas y medidas gubernamentales, que conforman una Reforma Educacional. Esta reforma apunta a avanzar hacia el fin de las desigualdades estructurales del sistema educativo, para que el Estado logre proveer una educación de calidad a toda la población, para lo cual propone mejorar sustancialmente la calidad, fortalecer la educación pública y reducir la desigualdad. Todo de manera transversal, desde la educación parvularia hasta la educación superior (MINEDUC, 2015). Sus medidas principales en relación a la educación preescolar y escolar son las siguientes:

- (i) Educación parvularia: Aumento de la cobertura y aseguramiento de la calidad mediante

mayor fiscalización, apoyo a los establecimientos y fortalecimiento de la formación de profesionales del área, además de la creación de la Subsecretaría de Educación Preescolar y la Intendencia de Educación Preescolar.

- (ii) Educación general: Ley de Inclusión Escolar, Ley Carrera Docente y nueva institucionalidad de la educación pública (fin a la municipalización).
- (iii) Ley Sistema de Desarrollo Profesional Docente, que involucra profesores de educación parvularia, básica y media.

La Ley de Inclusión Escolar, promulgada en mayo de 2015 (Biblioteca del Congreso Nacional de Chile, 2015), es uno de los pilares de esta reforma. Esta tiene como objetivos mejorar equidad del sistema escolar e igualar las condiciones en las que operan tanto los EE municipales como los particulares subvencionados. La ley propone cambios estructurales al sistema escolar chileno: (i) término gradual del financiamiento compartido hasta alcanzar gratuidad en todo el sistema subvencionado, (ii) fin a la selección mediante un nuevo sistema de admisión centralizado y (iii) fin al lucro de los establecimientos que reciben aportes del Estado.

Tanto el copago como la capacidad de los establecimientos particular subvencionados generan segmentación por NSE sin mejorar los resultados educativos (Carrasco et al., 2014; Mizala, 2014). Finalmente, es necesario prohibir el lucro en los establecimientos que reciban recursos públicos para así incentivar que los recursos que la sociedad asigna a educación sean reinvertidos en los EE para su fortalecimiento y mejora continua. Cada una de estas medidas se detalla a continuación.

2.3.1. Educación Subvencionada Gratuita

En lo que respecta a la gratuidad de colegios los establecimientos particular subvencionados, se puso en marcha un proceso de transición que suprime paulatinamente el financiamiento compartido, siendo el copago que las familias realizan en muchos colegios subvencionados y unos pocos municipales, reemplazado por recursos públicos. Este proceso parte el 2016 congelando los tramos de cobro que los colegios pueden hacer a las familias, a partir del 2017 comienza su disminución según el aporte estatal vaya en aumento (MINEDUC, 2017).

En la misma línea, se fortalecen los recursos entregados por la ley SEP y se implementa un aporte por Gratuidad a partir de 2016 para los EE que ya sean sin fines de lucro y gratuitos. Esto apunta a generar incentivos en los EE por adoptar la gratuidad, también se da la opción de que establecimientos que no desean ser gratuitos se transformen en particulares pagados, para la cual no hay un plazo máximo establecido (Biblioteca del Congreso Nacional de Chile, 2015).

2.3.2. Nuevo Sistema de Admisión Escolar

La Ley de Inclusión Escolar apunta a una admisión más transparente y no discriminatoria en los EE que reciban aportes del Estado. Con este objetivo se establece que los colegios no podrán rechazar postulantes si existen vacantes. Para los casos de existir más postulantes que vacantes, se implementa un mecanismo de asignación aleatorio adaptado del algoritmo de Aceptación Diferida (AD) del trabajo seminal de Gale y Shapley (1962). El cual considera los siguientes grupos prioritarios establecidos por la ley:

- (i) Estudiantes con hermanos en el colegio.
- (ii) Estudiantes hijos de funcionarios del establecimiento.
- (iii) Estudiantes que han estado con anterioridad en el EE y no habían sido expulsados.

Además, el mecanismo incluye cupos especiales para:

- (i) Estudiantes prioritarios (40 % más pobre).
- (ii) Estudiantes con necesidades educativas especiales o discapacidades.
- (iii) Establecimientos con proyectos educativos artístico, deportivo o de alto rendimiento académico (hasta 30 % de la matrícula).

Para su funcionamiento, se reciben las preferencias de los padres por colegios (deben marcar al menos dos preferencias, sin un límite máximo), las que sirven de insumo al algoritmo, el que, considerando los criterios, genera un *matching* de estudiantes a escuelas. Al momento de realizarse la asignación, empates dentro de cada grupo prioritario son resueltos de manera aleatoria (Correa et al., 2019).

En lo que respecta a su puesta en marcha, el mismo año 2015 se promulga el Decreto con Fuerza de ley N°3 (Biblioteca del Congreso Nacional de Chile, 2015), que estipula que para un primer periodo, sólo la Región de Magallanes y Antártica ingresa al nuevo sistema de admisión, lo que corresponde a las postulaciones del año 2016 para el año escolar 2017. En una segunda etapa, al año siguiente, se adhieren al SAE cuatro regiones más: Tarapacá, Coquimbo, O'Higgins y Los Lagos. Posteriormente, para el proceso 2018, se agregan todo las regiones restantes excepto la Región Metropolitana, la cual se ingresa al sistema este año 2019 ³.

En términos de cobertura de niveles, el primer año que una región ingresa al SAE, este sólo aplica de forma obligatoria en el ingreso a los niveles de prekínder, kinder, primero básico, séptimo básico y primero medio. A partir su segundo año de funcionamiento en la región, el ingreso a todos los niveles del sistema se debe hacer mediante SAE⁴.

³Al momento de finalizada esta investigación, estos datos aún no son de conocimiento público, por lo que están fuera del alcance se este trabajo.

⁴A partir del proceso de admisión 2018, se suprime la participación en el SAE de los niveles de ingreso Prekínder y Kinder cuando el establecimiento educacional no ofrece continuidad hacia enseñanza básica.

Por otra parte, aquellos EE con proyectos educativos deportivos, artísticos y/o de alta exigencia, están autorizados a seleccionar, según el criterio de su proyecto, hasta un 30% (MINEDUC, 2017) de su matrícula. Aún así, incluso en este escenario los establecimientos no podrán tomar pruebas de contenido académico.

Finalmente, se establece la prohibición de expulsar estudiantes por rendimiento, incluso en los establecimientos privados pagados. Además, todos los colegios deben garantizar la oportunidad a todos los estudiantes de repetir un curso en enseñanza básica y un curso en enseñanza media.

De esta forma, todos los establecimientos educacionales que reciban aportes financieros del Estado deberán adscribir a este nuevo sistema de admisión, para que así acercarse hacia una sociedad en la que todos los estudiantes tengan las mismas oportunidades.

2.3.2.1. Algoritmo de Asignación

La base del mecanismo de asignación utilizado por el SAE es el célebre algoritmo de Aceptación Diferida (Gale y Shapley, 1962), el cual toma preferencias de agentes sobre los elementos de un conjunto, y genera un *matching* que posee propiedades de estabilidad y eficacia deseables en múltiples situaciones⁵, estudiadas posteriormente en profundidad por Roth (1982). Este mismo autor estudia el algoritmo en el contexto de asignación de médicos recién graduados a hospitales para sus residencias (Roth, 1984), a partir del cual se populariza su uso en diversos contextos (Roth, 2008).

En lo que respecta a educación, una serie de reformas en asignación escolar alrededor del mundo se han implementado con AD como base. Su primera gran implementación se lleva a cabo en Nueva York (Abdulkadiroğlu, Pathak, y Roth, 2005) y posteriormente se aplica en Boston (Abdulkadiroğlu, Pathak, Roth, y Sönmez, 2005), en desmedro del también famoso *Boston Mechanism*⁶. Posteriormente, variaciones de AD han sido utilizadas en ciudades como Barcelona (Calsamiglia et al., 2017)) y Amsterdam (de Haan et al., 2016). En Chile, al 2019, ya se ha utilizado en cuatro procesos de admisión, estando bajo continuo escrutinio público y evaluación⁷.

Hasta ahora, los estudios concernientes a sistemas de admisión mediante AD se han centrado en aspectos algorítmicos de su implementación, enfocados sobre todo en segregación y bienestar social. Mientras que la presente investigación lo hace en sus efectos sobre las preferencias de los padres y como éstas se ven modificadas (o no) producto de la entrada en vigencia de un nuevo sistema de admisión.

⁵Gale y Shapley (1962) demuestran que si las preferencias son estrictas, siempre existe un *matching* estable óptimo para los agentes.

⁶El cual no es ni estable ni a prueba de estrategias (Kojima y Ünver, 2011).

⁷Para mayores detalles sobre el aspecto algorítmico de su implementación consultar Correa et al. (2019)

2.3.2.2. Los Efectos del SAE

Una serie de trabajos han estudiado los efectos que ha tenido el SAE sobre el sistema escolar chileno desde su primera implementación en 2015, basado en los datos de preferencias y asignación provistos por el sistema.

En términos socioeconómicos, Eyzaguirre et al. (2019a) identifica sesgos en las preferencias por colegio en todas las regiones del país, mostrando que estudiantes prioritarios en promedio postulan más a EE municipales y menos a EE con copago (aunque sean gratuitos para ellos), lo que hace surgir nuevamente la hipótesis de autoexclusión social. Además, infiere que estudiantes prioritarios postulan a colegios con puntajes SIMCE considerablemente más bajos y con mayor IVE, lo que se observa en todos los niveles escolares.

Al analizar la asignación, el mismo trabajo concluye que proporcionalmente más alumnos no prioritarios fueron admitidos en EE subvencionados, con mayor copago, SIMCE más alto e IVE más bajo. Aún así, concluyen que (por ahora) el SAE avanza a una distribución menos segregada de estudiantes prioritarios. Cuando se intentan explicar las causas de estos sesgos, Eyzaguirre et al. (2019b) apuntan a dos factores: (i) la segregación residencial que limita la *choice set* de cada estudiante y por ende el acceso a escuelas efectivas (Rodríguez et al., 2016), y (ii) un déficit de información que afecta a las familias más vulnerables, lo que se condice con los resultados de Aguirre (2018).

Por otro lado, aunque Carrasco y Honey (2019) observan una leve mejora en términos de acceso a EE de calidad, concluye que la distribución de cupos en los EE más deseados permanece inalterada, destacando que se observa una desigualdad en estructura de preferencias (al igual que Eyzaguirre et al. (2019b)). Carrasco et al. (2019) resaltan las ventajas del SAE en su proceso de postulación, identificando, por lo demás, tres procesos de adaptación cultural en estos primeros años del sistema: (i) nuevo equilibrio social de oportunidades, (ii) paso de una experiencia presencial a una virtual, y (iii) adaptación a un sistema complejo en tecnología y reglas. Estos fenómenos podrían ayudar a explicar efectos contraintuitivos que se observan hasta ahora, más aún, estos hacen un llamado a ser precavidos sobre las conclusiones preliminares respecto de efectos del SAE, dada la adaptación que se está viviendo.

Asimismo, Urzua (2019) no advierte cambios significativos en las variables que determinan la asignación final de estudiantes a colegios, y plantea que el SAE tiene efectos ambiguos en términos de segregación. Reafirma que las familias consideran diversos factores a la hora de escoger escuela, mostrando que el SIMCE no siempre parece relevante, y que ciertamente lo es menos que la distancia hogar colegio. Plantea además que los efectos ambiguos observados pueden deberse a que el algoritmo es muy complicado de entender para las familias.

Finalmente, desde un enfoque más técnico, Correa et al. (2019) estudian la última implementación del SAE, proponiendo mejoras algorítmicas que impactan en dos dimensiones los resultados: (i) aumentando las chances de que hermanos queden en el mismo EE y (ii) asegurando que los postulantes que se cambian de colegio al menos tengan cupo en su EE de origen. Además, muestra que si lo que se busca es beneficiar a estudiantes prioritarios, la actual cuota obligatoria de 15% es insuficiente, lo que confirma lo observado por trabajos anteriores.

Es claro que la evidencia sobre el SAE es aún muy prematura para sacar sus conclusiones, en parte debido a que sólo se han analizado los datos del nuevo sistema de admisión, en vez de tomar una perspectiva más amplia comparando el actual escenario con lo que era el *status quo* del sistema anteriormente. Por otro lado, la variante geográfica sólo ha sido integrada parcialmente en el análisis, siendo está relevante dada las alta heterogeneidad entre las comunas chilenas. En tanto, los estudios anteriores sólo han integrado algunas de las variables relevantes en *school choice* (ver secciones anteriores), siendo aún menos considerada la interacción entre estas. Es por ello que la presente investigación, mediante una nueva metodología, busca ahondar en la comprensión del fenómeno tomando en cuenta todos los elementos antes identificados.

2.3.3. Fin al Lucro

Para cumplir con este tercer objetivo, se crea el concepto de Fines Educativos, el que agrupa once clases de gastos ligados al Proyecto Educativo Institucional (PEI) de cada EE, que abarcan todos los gastos necesarios para la realización del proyecto educativo, excluyendo así el retiro directo o indirecto de ganancias.

De esta manera, la ley estipula que a partir de marzo de 2016, los fondos públicos sólo se pueden usar en alguna de las once categorías correspondientes a Fines Educativos. Por consiguiente, cualquier tipo de excedente generado en base a estos recursos, debe ser reinvertido directamente en el fortalecimiento del PEI (OPECH, 2015).

Para resguardar el cumplimiento de esta nueva medida, la ley confiere a la Superintendencia de Educación Escolar la facultad de fiscalizar a los EE, los que deben llevar cuentas por separado de los recursos públicos, desde ahora destinados sólo a Fines Educativos, y de aquellos que tengan otro origen y puedan dirigirse a otros usos MINEDUC (2017).

Capítulo 3

Datos y Estadísticas Descriptivas

Este trabajo analiza a todos los estudiantes en establecimientos que ofrecen educación en los niveles de sexto básico, séptimo básico, octavo básico o primero medio, en el periodo 2013-2019. Se consideran sólo las cinco aglomeraciones urbanas de mayor población en las regiones de Tarapacá, Coquimbo, O'Higgins y Los Lagos. La tabla 3.1 muestra las comunas que forman parte de este trabajo¹ y su número de habitantes.

Región	Comunas	Habitantes	Total Conurbación
Tarapacá	Iquique	191.468	299.843
	Alto Hospicio	108.375	
Coquimbo	Coquimbo	227.730	448.784
	La Serena	221.054	
O'higgins	Rancagua	241.774	294.279
	Machalí	52.505	
Los Lagos	Puerto Montt	245.902	290.480
	Puerto Varas	44.578	

Tabla 3.1: Conurbaciones de las 4 regiones consideradas.

Fuente: Instituto Nacional de Estadísticas (2017)

Se trabaja sólo con estos grados pues los cambios en las transiciones séptimo básico-octavo básico y octavo básico-primero medio deben hacerse obligatoriamente mediante el SAE, por lo cual son un buen escenario de comparación con respecto a lo que ocurría antes de implementarse el nuevo sistema. Además, en estos casos se cuenta con los datos previos de aquellos estudiantes que se cambian de escuela, lo que permite considerar mayores dimensiones en el análisis, esto no es posible con el ingreso a prekínder, kinder y primero básico, que también deben obligatoriamente utilizar el SAE.

El estudio se centra en las cuatro regiones antes señaladas pues se dispone de los datos de los procesos de postulación SAE 2017 y 2018. A pesar de que el 2018 se agreguen más regiones y niveles, para contar con al menos dos años de postulaciones se mantienen sólo

¹Inicialmente también se consideró la comuna de Punta Arenas, sin embargo, esta mostró no presentar la estructura suficiente para obtener en ella resultados significativos e interpretables. Se discute más al respecto en el Capítulo de Resultados.

estas regiones como parte de la investigación. Además, como fuente de validación externa a la elección de aglomeraciones urbanas, las conurbaciones escogidas son las mismas con las consideradas en el informe SAE 2017 como zonas urbanas (Sistema de Admisión Escolar, 2018).

En lo que respecta a la elección de años, con los datos 2013-2016 se modelará la situación previa al SAE, mientras que con los datos 2017-2018 se trabajará el escenario en el que el nuevo sistema ya está en funcionamiento. Es importante señalar que al referirse a un determinado proceso SAE, este comprende el proceso de admisión para el año siguiente. Por ejemplo, el proceso SAE 2018 corresponde a la admisión de estudiantes para el año escolar 2019.

En lo que se refiere a las fuentes de los datos, se utiliza información provenientes de la plataforma de datos abiertos² del Ministerio de Educación (MINEDUC), los que permiten caracterizar a establecimientos y estudiantes. Además, se cuenta con la información de postulaciones hechas mediante el SAE en los procesos de admisión 2017 y 2018, que permiten modelar la demanda educativa luego de implementada la reforma. Por otro lado, se trabaja con los resultados SIMCE y del cuestionario que los padres responden al momento en que sus hijos rinden la prueba, provistos por la Agencia de la Calidad, con los cuáles se pueden agregar las dimensiones socioeconómica y de rendimiento académico al estudio. Finalmente, se utilizan estimaciones de distancia efectiva hogar-colegio proporcionadas por el Centro de Investigación Avanzada en Educación (CIAE), lo que permite contar con una variable geográfica para el análisis. En el Apéndice A se adjunta una Tabla con cada variable utilizada, su descripción y escala de valores.

Con el fin de precisar la muestra de EE y estudiantes, se excluyen aquellos ubicados en zonas rurales, de modo que el análisis sobre elección de escuelas se realice en contextos donde exista una mayor y más diversa oferta educativa, lo que usualmente no ocurre en áreas rurales, en las que hay limitadas opciones de establecimiento. Además, para delimitar el mercado educativo de cada aglomeración, sólo se consideran estudiantes matriculados en algún establecimiento de ésta, excluyendo a aquellos que, aunque vivan en la conurbación, estudian en una comuna aledaña fuera de esta. Quedan fuera también aquellos estudiantes que al momento de cambiarse, ya sea en sexto u octavo básico, están matriculados en un establecimiento fuera de la aglomeración considerada.

De esta manera, la muestra queda compuesta, por 107 establecimientos y 226.000 estudiantes por año. Vale señalar que esta se restringe a aquellas observaciones que contengan información en las seis variables fundamentales del modelo (Ver Tabla 4.2). Además, para lograr una mayor representatividad, se utilizan técnicas de tratamiento de datos faltantes, según el marco establecido por Batista y Monard (2004), para lo cual, se consideran criterios pertinentes a cada una de las variables, En el Apéndice C se detallan estos procedimientos.

A continuación, se explican las principales variables utilizadas en este trabajo, detallando su uso y estadísticas descriptivas. Se presentan primero las variables a nivel de estudiante y luego las que caracterizan a los establecimientos. Finalmente, se exponen aquellas que conciernen la interacción estudiante- escuela. Cada variable está acotada a las cinco regiones consideradas y únicamente al periodo 2013-2018. En el Apéndice B se presenta la tabla de

²www.datos.mineduc.cl

representatividad de cada variable utilizada.

3.1. Variables a Nivel de Estudiantes

Género del Estudiante

En este estudio esta variable toma el valor 1 para masculino y 2 en el caso femenino. La Figura 3.2 muestra el porcentaje de cada género.

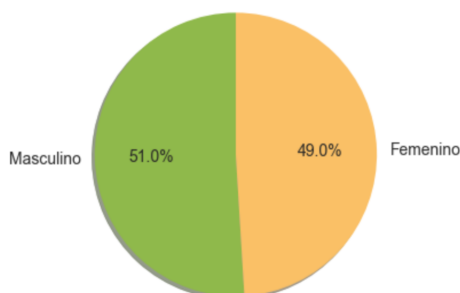


Figura 3.1: Distribución de estudiantes por género.
Fuente: Elaboración propia.

Esta misma proporción se mantiene al considerar únicamente la población de estudiantes que cambian de establecimiento educativo.

Promedio de Notas

Esta variable es el promedio final de notas que un estudiante obtiene en el año escolar, el que varía entre 1,0 y 7,0. Para efectos de este estudio se considera el percentil de la nota con respecto a estudiantes del mismo nivel y establecimiento durante el año escolar, lo que permite la comparación a través de todo el sistema.

La Figura 3.2 representa el histograma de promedio general de notas y el diagrama de caja con bigote. Se observa que el histograma tiene forma de campana, la mediana es 5,8 y la mayoría de los estudiantes tiene promedio entre 5,3 y 6,2.

Beneficiario SEP

Esta variable indica si el estudiante es beneficiario de la ley SEP, lo que significa que es alumno prioritario (40% más vulnerable) y está matriculado en algún establecimiento municipal o particular subvencionado adscrito a la SEP. A partir del año 2016 se consideran como beneficiarios SEP, además, a los alumnos preferentes que cumplan las mismas condiciones anteriores. Se entiende que un alumno es preferente si no es prioritario y pertenece al 80% más vulnerable de la población según el Registro Social de Hogares.

La Figura 3.3 muestra el porcentaje de estudiantes beneficiarios SEP por año en EE municipales y particular subvencionados, a nivel regional. No se incluyen los EE particulares

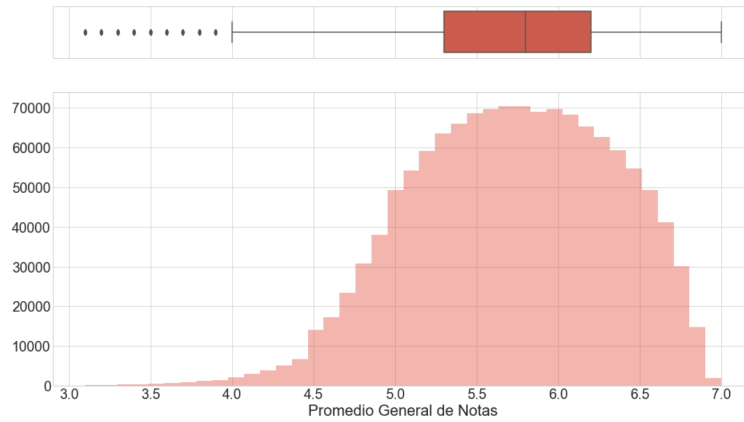


Figura 3.2: Histograma de promedio general de notas estudiantes.
Fuente: Elaboración propia.

pagados pues no hay estudiantes SEP en esta dependencia administrativa. En esta figura, se observa consistentemente a través del tiempo que los establecimientos municipales tienen mayor proporción de estudiantes SEP que los particular subvencionados. Por otro lado, es claro el efecto de la ampliación de cobertura a partir del año 2016.

En términos geográficos, los valores más elevados corresponden a la aglomeración Puerto Montt - Puerto Varas.

Ingreso Familiar

Esta variable se obtiene a partir del cuestionario para padres aplicado a los estudiantes que rinden la prueba SIMCE. La Figura 3.4 muestra la distribución del ingreso del hogar de estudiantes según dependencia administrativa, donde la escala indica ingresos desde menos de \$100.000 (valor 1) hasta más de \$2.200.000 (valor 15, ver Apéndice A para mayor detalle).

Los tramos de ingreso predominantes corresponden a ingresos entre \$100.000 y \$400.000 (valores 2, 3 y 4), en los cuales la matrícula se reparte de manera similar entre municipales y subvencionados. A medida que aumentan los ingresos, la matrícula de EE particular subvencionados aumenta en desmedro de la municipal. A partir de \$1.200.000 (valor 10), emerge la matrícula de establecimientos privados, teniendo su máximo en el tramo de ingreso más alto.

La Tabla 3.2 detalla los valores exactos para cada categoría y dependencia administrativa.

Dependencia \ Ingreso	1	12	23	34	45	56	67	78	89	10	11	12	13	14	15
Privados	58	45	71	98	149	196	367	590	581	559	563	514	754	779	5483
P. Subvencionados	2335	9782	16888	14243	10803	8893	9368	6899	4148	2485	1783	1267	1271	913	2053
Municipales	4007	12869	15821	9089	5085	3372	2604	1352	533	302	194	106	88	42	119

Tabla 3.2: Valores de Ingreso Hogar según Dependencia Administrativa

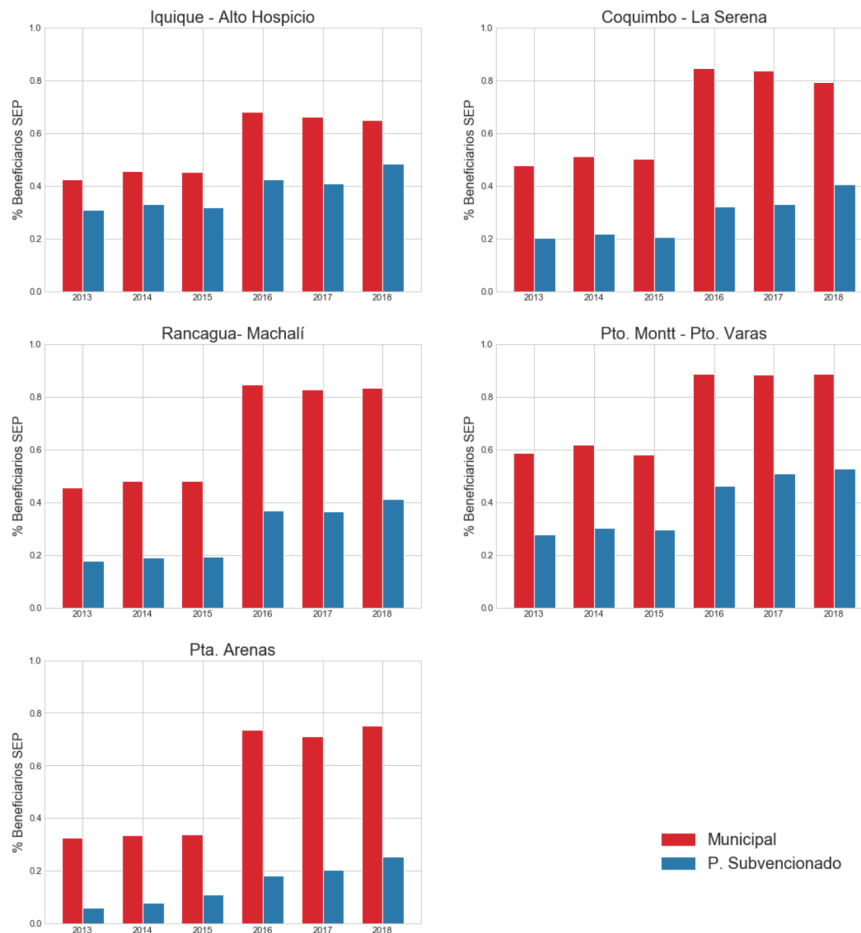


Figura 3.3: Beneficiarios SEP por año según dependencia administrativa y región.
Fuente: Elaboración propia.

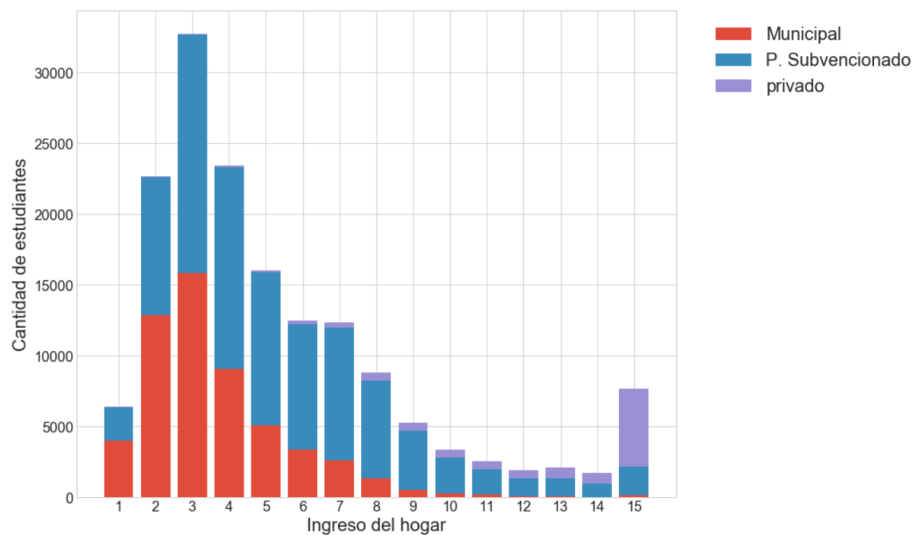


Figura 3.4: Distribución del ingreso del hogar de estudiantes según dep. administrativa.
Fuente: Elaboración propia.

Nivel de Estudios de los Padres

Esta variable toma valores entre 1 y 20, los que indican el nivel de escolaridad de cada padre, partiendo desde primero básico hasta el grado de doctorado, la Tabla 3.3 detalla la

descripción de cada una de estas categorías. Es importante señalar que el valor numérico de la variable no representa exactamente los años de estudio del padres. Se trabaja con el nivel de estudios promedio de ambos padres.

Valor	Descripción
1	No estudió
2	1 ^{er} año de educación básica
3	2 ^o año de educación básica
4	3 ^o año de educación básica
5	4 ^o año de educación básica
6	5 ^o año de educación básica
7	6 ^o año de educación básica
8	7 ^o año de educación básica
9	8 ^o año de educación básica
10	I año de educación media
11	II año de educación media
12	III año de educación media
13	IV año de educación media científico-humanista
14	IV o V año de educación media técnico-profesional
15	Educación incompleta en un centro de FT o IP
16	Titulado de un centro de FT o IP
17	Educación incompleta de una universidad
18	Educación universitaria completa
19	Grado de magíster universitario
20	Grado de doctor universitario

Tabla 3.3: Valores para la variable Estudio de Padres

La Figura 3.5 muestra la distribución del nivel de estudio de padres según dependencia administrativa, en la que se observa que los niveles que más se observan corresponden a estudios completos de enseñanza media científico-humanista y técnico-profesional (valores 13 y 14 respectivamente). A menor nivel de estudios se aprecia una mayor representación de matrícula municipal, mientras que a partir del valor 13, predominan los establecimientos particular subvencionados, en desmedro de la matrícula pública. En lo que respecta a los establecimientos particulares pagados, vemos que estos se concentran en estudios universitarios completos (valor 18).

Distancia Hogar-Colegio

Variable calculada en kilómetros, la cual proviene de dos fuentes: del CIAE se obtiene la distancia efectiva hogar-colegio de los estudiantes el año 2016, estos se cruzan con matrículas de del periodo 2013-2016 para mejorar la representatividad de la variable. En segundo lugar, desde los datos SAE, se toman las coordenadas de latitud y longitud de los hogares de los postulantes, para calcular, en conjunto con las coordenadas de los EE, las distancia euclidiana³ entre sus hogares y establecimientos de origen. Por consiguiente, las distancias Pre SAE no son comparables con las post SAE, pues se calculan de manera distinta, por lo que esto se debe tener presente en el análisis posterior.

³en línea recta.

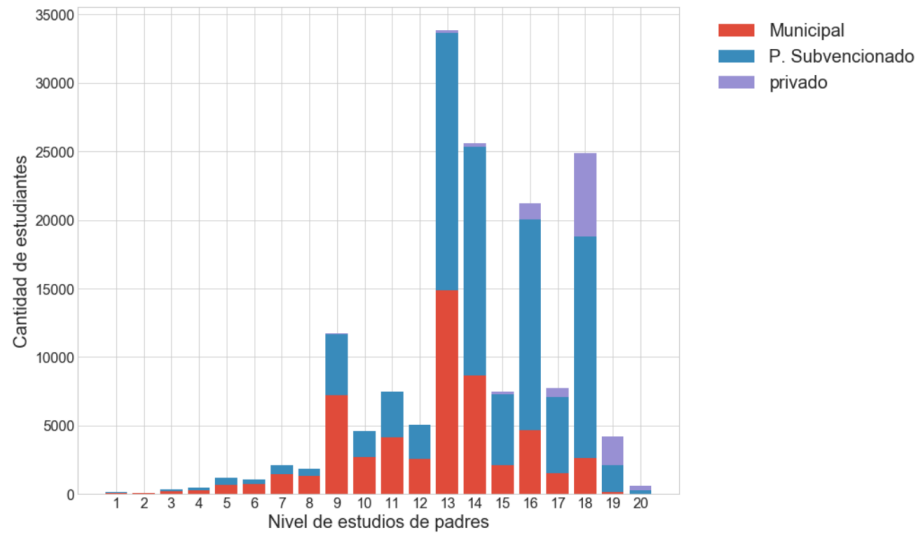


Figura 3.5: Distribución del nivel de estudio de padres según dependencia administrativa.
Fuente: Elaboración propia

La Figura 3.6 muestra el histograma normalizado de distancias hogar-colegio según dependencia administrativa (se considera sólo hasta el percentil 95), sobre el cual se ubica un *boxplot* con las medidas de posición de los datos agregados.

Se observa que la mediana de estudiantes vive a 2,2 km de su establecimiento, siendo la media aproximadamente 3 km. La mayoría de los estudiantes viven entre 1 y 4 km de su EE. Según dependencia administrativa, los estudiantes de EE municipales son los que recorren menores distancias, seguidos por los establecimientos particulares subvencionados. Los estudiantes de colegios particulares pagados son los que se desplazan trayectos más largos.

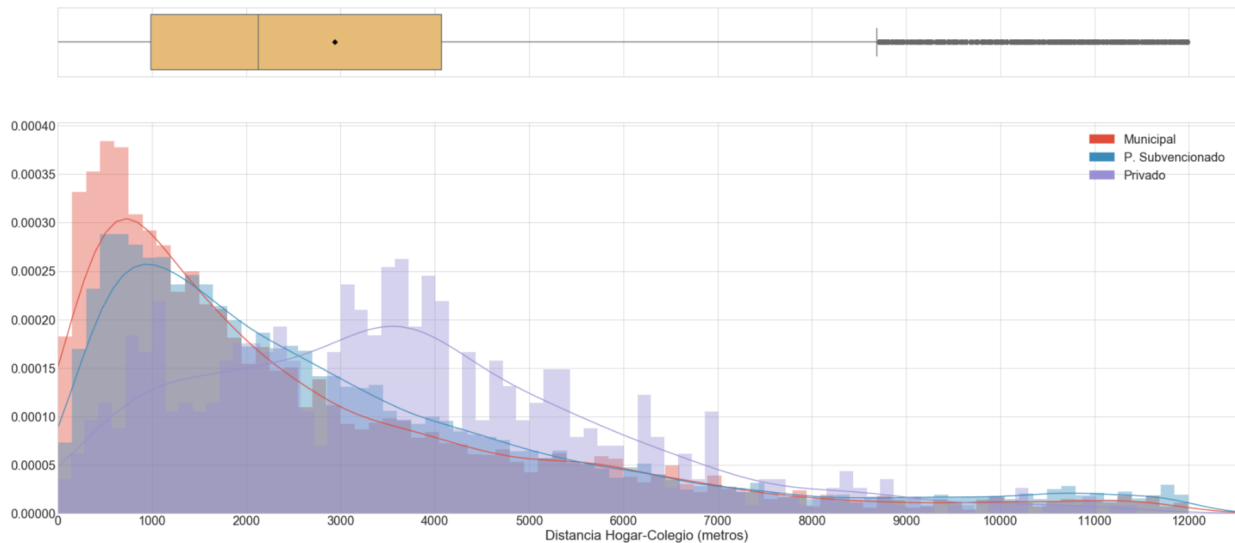


Figura 3.6: Distancia hogar-colegio según dependencia administrativa.
Fuente: Elaboración propia

3.2. Variables a Nivel de Establecimientos

Matrícula

En la Figura 3.4 se aprecia el comportamiento de la matrícula en cada aglomeración a lo largo de los años. Se observa que Coquimbo - La Serena tiene el mayor número de estudiantes, seguida por Iquique - Alto Hospicio, Rancagua - Machalí y Puerto Montt - Puerto Varas (los tres muy similares). Lo anterior está correlacionado con la cantidad de habitantes por conurbación (Tabla 3.1). Por otro lado, Coquimbo-La Serena presenta el mayor aumento de matrícula (+ 10 %).

Aglomeración \ Año	2013	2014	2015	2016	2017	2018
Iquique - Alto Hospicio	9.477	8.960	9.333	9.321	9.563	9.867
Coquimbo - La Serena	12.400	12.715	12.497	12.644	13.042	13.747
Rancagua- Machalí	8.966	8.852	8.861	8.749	8.791	9.102
Pto. Montt - Pto. Varas	8.757	8.897	8.878	8.787	9.218	9.183
Total Matrícula	48.048	47.546	47.834	47.746	48.974	50.384

Tabla 3.4: Matrícula total por aglomeración por año.

Fuente: Elaboración propia.

Al segmentar la matrícula por dependencia administrativa (Figura 3.7) se advierte en el año 2018 una baja en la matrícula de los EE particulares subvencionados al mismo tiempo que aumenta la de particulares pagados. Esto puede ser un efecto de la Ley de Inclusión Escolar, que obliga a los EE subvencionados a transformarse en entidades sin fines de lucro o en su defecto pasar a ser particulares pagados, lo que ha generado el paso de EE subvencionados de copago alto a particulares pagados. Por otro lado, se observa una disminución leve en la matrícula de EE municipales, lo que podría explicarse por las preferencias de las familias por establecimientos particular subvencionados sobre municipales (ver Capítulo 2).

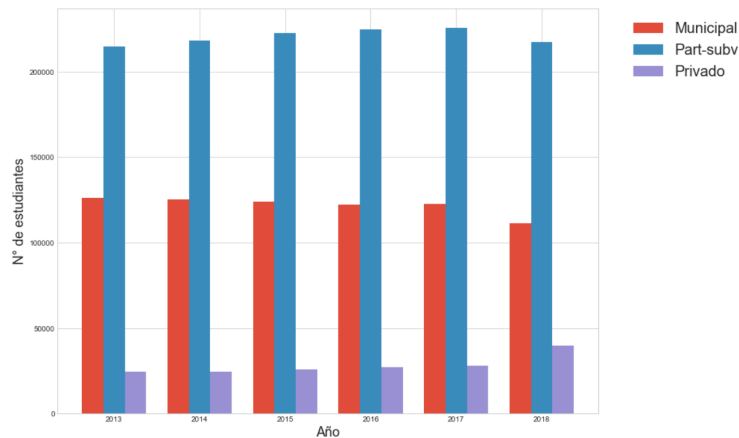


Figura 3.7: Matrícula total según dependencia administrativa por año.

Fuente: Elaboración propia.

Dependencia Administrativa

Variable que indica el régimen administrativo bajo el cual funciona el establecimiento. Este puede ser: municipal, particular subvencionado, particular pagado, Corporación de Administración Delegada o Servicio Local de Educación (SLE), esta última corresponde a una nueva unidad administrativa fruto del proceso de desmunicipalización que se lleva a cabo en todo el país.

Con el fin de mantener lo esencial de la categorización, los establecimientos de corporaciones de administración delegada, que representan el 0,5% del total de establecimiento de la muestra, se agrupan con los particulares subvencionados, ya que ambos son de administración privada y reciben recursos del estado. Por otra parte, los establecimientos de SLE se unen al grupo de colegios municipales, ya que esta era su dependencia de origen.

La Figura 3.8 muestra la distribución de EE según dependencia administrativa por año, en la cual destaca una disminución en EE subvencionados y un aumento en establecimientos privados el año 2018, lo que es consistente con la Figura 3.7.

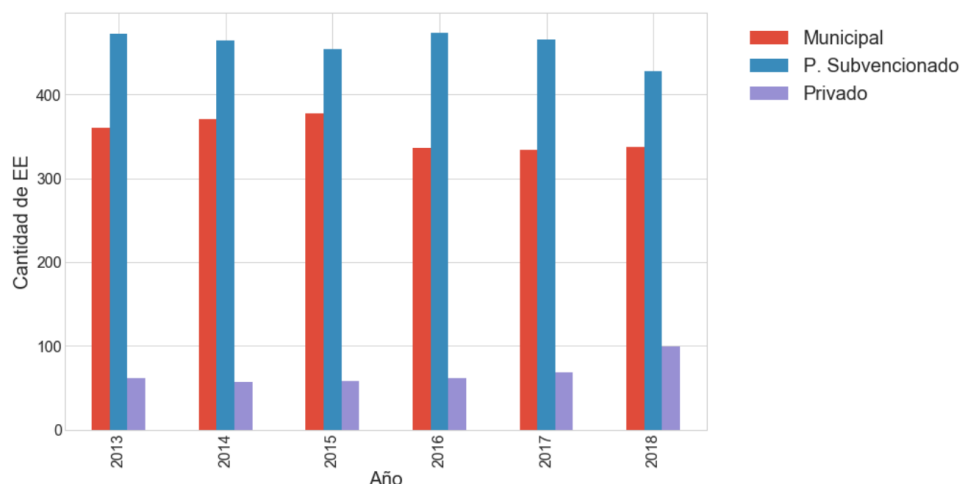


Figura 3.8: Distribución de EE según dependencia administrativa por año.

Fuente: Elaboración propia

Por otro lado, los incentivos para sostenedores cambian dadas las restricciones introducidas por la ley de inclusión escolar hacia los establecimientos: prohibición del lucro en EE que reciban aportes del estado y gratuidad progresiva, congelando el copago cobrado a las familias. Estas han provocado que algunos establecimientos de copago alto decidan cambiar su dependencia administrativa, convirtiéndose en particulares privados. Las Tablas 3.9, 3.10, 3.11 y 3.12 muestran dichos cambios.

En todas las aglomeraciones urbanas se observa que en el año 2017 ocurre un cambio significativo de dependencia administrativa de establecimientos, siendo especialmente fuerte este efecto en Coquimbo - La Serena, donde casi el 8% de la matrícula, correspondiente a once establecimientos, migra hacia dependencia particular privada.

	Año	Total establecimientos	Total matrícula	Establecimientos que cambian de dependencia	Estudiantes en establecimientos que cambian de dependencia	Porcentaje de estudiantes en establecimientos que cambian de dependencia
0	2013	141	63864	0	0	0
1	2014	137	63473	0	0	0
2	2015	135	64761	3	124	0.19
3	2016	134	65261	1	133	0.2
4	2017	135	66279	13	2982	4.5
5	2018	133	67317	0	0	0

Figura 3.9: Establecimientos y matrícula que cambia de dependencia administrativa en Iquique - Alto Hospicio.
Fuente: Elaboración propia

	Año	Total establecimientos	Total matrícula	Establecimientos que cambian de dependencia	Estudiantes en establecimientos que cambian de dependencia	Porcentaje de estudiantes en establecimientos que cambian de dependencia
0	2013	235	84772	0	0	0
1	2014	236	86413	0	0	0
2	2015	236	87886	1	878	1
3	2016	232	88590	2	87	0.1
4	2017	232	90695	11	7034	7.76
5	2018	228	92633	0	0	0

Figura 3.10: Establecimientos y matrícula que cambia de dependencia administrativa en Coquimbo - La Serena.
Fuente: Elaboración propia

	Año	Total establecimientos	Total matrícula	Establecimientos que cambian de dependencia	Estudiantes en establecimientos que cambian de dependencia	Porcentaje de estudiantes en establecimientos que cambian de dependencia
0	2013	140	62460	0	0	0
1	2014	139	62457	1	767	1.23
2	2015	139	62942	0	0	0
3	2016	138	62786	0	0	0
4	2017	133	63012	5	1981	3.14
5	2018	133	63517	0	0	0

Figura 3.11: Establecimientos y matrícula que cambia de dependencia administrativa en Rancagua - Machalí.
Fuente: Elaboración propia

Copago

Corresponde al pago que realizan las familias en establecimientos que reciben aportes del Estado. Este se calcula para cada EE a partir de su ingreso anual declarado, dividido en asistencia promedio y en cantidad de meses del año (12). La Figura 3.13 muestra la progresión anual del copago en EE particulares subvencionados. Sólo se considera esta dependencia pues los EE municipales son gratuitos, y por otro lado, no hay información precisa sobre el cobro realizado por EE particulares pagados a las familias, el cual de todas formas es muy superior.

Año	Total establecimientos	Total matrícula	Establecimientos que cambian de dependencia	Estudiantes en establecimientos que cambian de dependencia	Porcentaje de estudiantes en establecimientos que cambian de dependencia
0 2013	156	59568	0	0	0
1 2014	160	60579	0	0	0
2 2015	157	60726	0	0	0
3 2016	149	61016	0	0	0
4 2017	148	60971	1	250	0.41
5 2018	147	61614	0	0	0

Figura 3.12: Establecimientos y matrícula que cambia de dependencia administrativa en Puerto Montt - Puerto Varas.
Fuente: Elaboración propia

Se observa claramente el efecto de la Ley de Inclusión Escolar, la que desde su entrada en vigencia en 2015, provoca una disminución abrupta en el copago, cuyo valor mediano baja a 0 a partir de 2016. El copago promedio también disminuye significativamente. Cabe notar, además, la presencia de establecimiento en el cuartil más alto que aún cobran una elevada mensualidad a las familias.

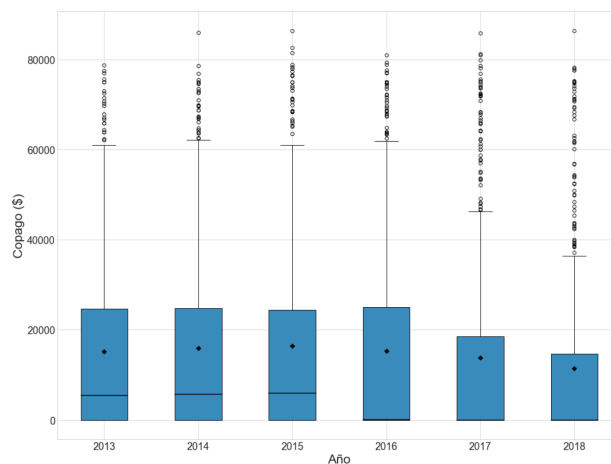


Figura 3.13: Distribución de copago en EE particulares subvencionados por año.
Fuente: Elaboración propia

La Figura 3.14 muestra la segmentación por tramos de pago, también para los establecimientos particulares subvencionados.

Se observa que hasta el año 2015, el tramo mayor de cobro (más de \$50.000 mensuales) iba levemente el aumento. Tendencia que se observa que es interrumpida abruptamente a partir del año 2016, con la entrada en vigencia de la Ley de Inclusión Escolar, lo que genera un sostenido aumento de EE subvencionados gratuitos, simultaneo a una reducción de establecimientos con copago, sobre todo en los tramos de cobro inferior.

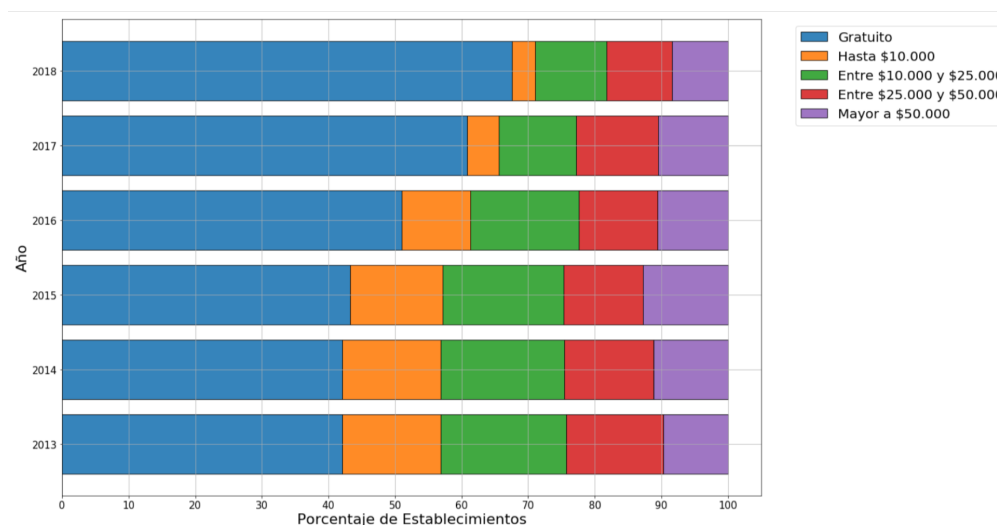


Figura 3.14: Distribución de copago en EE particulares subvencionados por año.

Fuente: Elaboración propia

Índice de Vulnerabilidad Escolar

El Sistema Nacional de Asignación con Equidad⁴ (SINAE) clasifica a los estudiantes en 3 niveles, según condiciones de pobreza y riesgo de fracaso escolar. A partir de esta categorización, se calcula para cada escuela el Índice de Vulnerabilidad Escolar (IVE), correspondiente al porcentaje de alumnos del EE que pertenece a alguna de las tres categorías anteriores. Este índice se utiliza para cuantificar el grado de vulnerabilidad de los estudiantes de un colegio.

La Figura 3.15 muestra diagramas de caja de la variable IVE anual, segmentada en dependencias municipal y particular subvencionada. No se consideran EE particulares pagados ya que sus IVE son 0. Se observa una mayor vulnerabilidad en establecimientos municipales y una distribución más concentrada en los mismos. En lo que respecta a EE particulares subvencionados, se aprecia un leve aumento en su vulnerabilidad promedio, aunque su valor mediano permanece inalterado. Además, estos últimos exhiben una mayor varianza en su distribución.

Por otro lado, el SINAE provee de este índice tanto para enseñanza básica como para enseñanza media, dependiendo de los niveles que tenga cada establecimiento. En el caso de aquellos que imparten ambos ciclos, para efectos de este estudio, se considera como IVE del EE, el promedio ponderado por matrícula de cada ciclo educativo.

Puntaje SIMCE

Resultados a nivel estudiante y de colegio de la prueba estandarizada SIMCE, la cual es de carácter censal y se rinde en tres niveles distintos, variando año a año entre 4º básico (siempre rinde la prueba), 6º básico, 8º básico y 2º medio. La Tabla 3.5 muestra los años y cursos, concernientes a este trabajo, para los que estos datos están disponibles. Se trabaja con

⁴Organismo dependiente del MINEDUC.

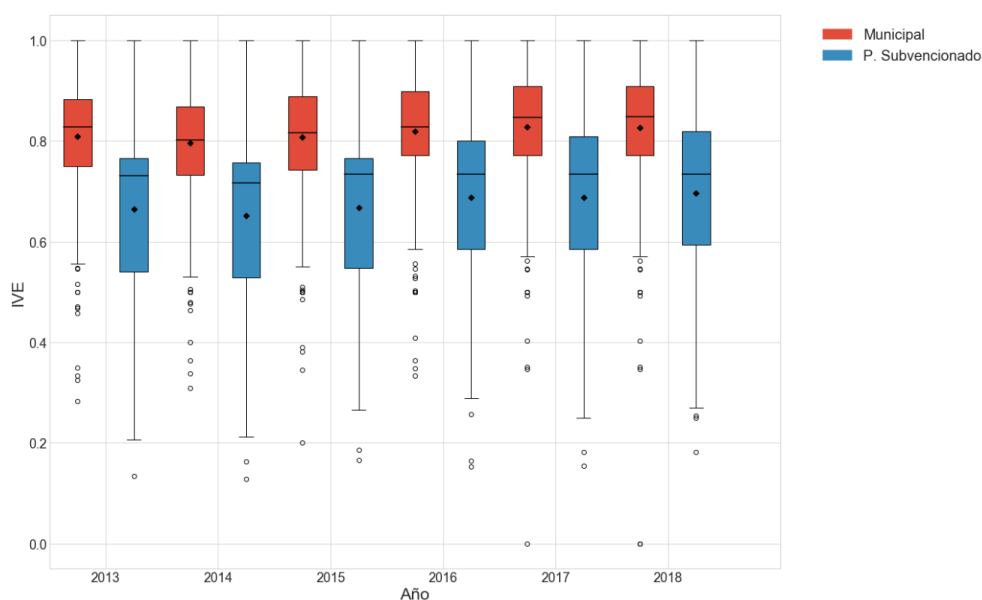


Figura 3.15: Distribución de IVE según dependencia administrativa por año.

Fuente: Elaboración propia

el puntaje promedio lenguaje - matemáticas a nivel alumno. Cabe señalar que esta variable se utiliza tanto a nivel estudiante, como a nivel de establecimiento, el que corresponde al puntaje promedio de sus estudiantes.

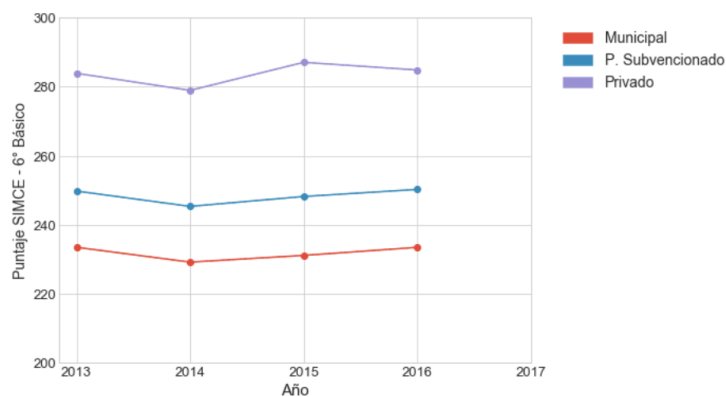
Curso \ Año	2012	2013	2014	2015	2016	2017	2018
6° Básico	-	✓	✓	✓	✓	-	✓
8° Básico	-	✓	✓	✓	-	✓	-

Tabla 3.5: Rendición de prueba SIMCE por nivel y año.

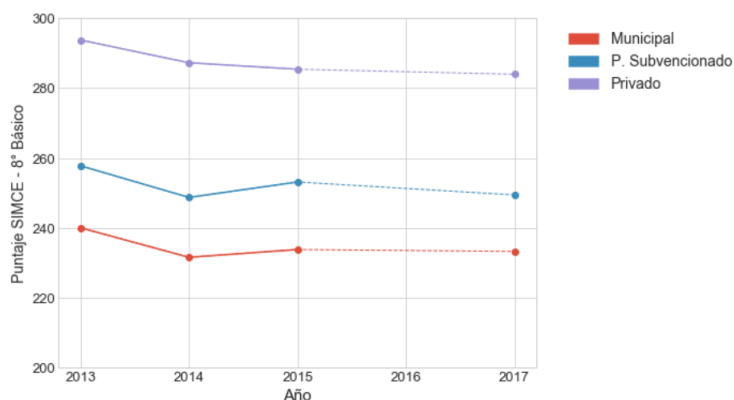
La Figura 3.16 muestra el puntaje promedio entre pruebas de Matemáticas y Lenguaje, según dependencia administrativa y año, para los niveles de sexto y octavo básico. Se observa un cierto estancamiento durante el periodo considerado. Sin embargo, es marcada la diferencia de resultados según dependencia. Establecimientos particulares pagados obtienen los más altos puntajes, seguidos por EE particular subvencionados, los que, por su parte, consistentemente se ubican por sobre las escuelas municipales. Estos resultados están fuertemente correlacionados con el NSE de las familias y por ende con las dependencias de los EE

Grupo Socioeconómico GSE

Variable categórica a nivel de EE que indica el grupo socioeconómico promedio al cual pertenecen sus estudiantes. Esta considera cinco tramos: bajo, medio bajo, medio, medio alto y alto. La Figura 3.17 muestra la distribución de establecimientos según GSE y dependencia administrativa de manera acumulada para el periodo 2013 - 2018. Se observa una predominancia de los GSE medio y medio bajo (valores 2 y 3). Para el nivel bajo la mayoría de EE son municipales, para los tres grupos intermedio son particular subvencionados los que más



(a) Sexto Básico



(b) Octavo Básico

Figura 3.16: Puntaje promedio SIMCE según dependencia administrativa por año.

Fuente: Elaboración propia

presencia tienen, mientras que sólo en el GSE más alto, están y predominan establecimientos particulares pagados.

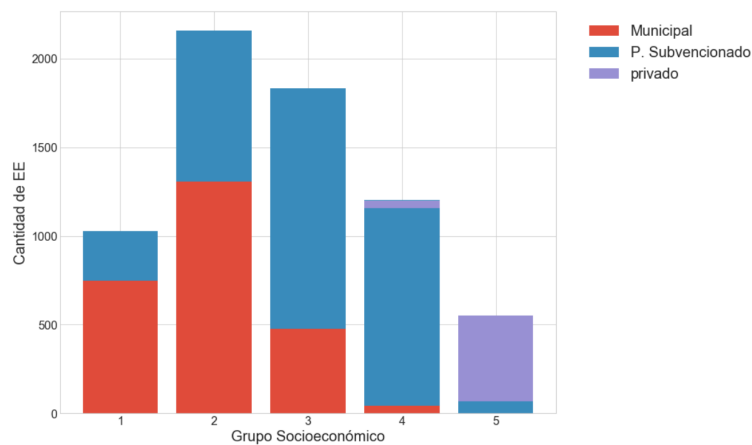


Figura 3.17: Distribución de establecimientos según GSE y dependencia administrativa.

Fuente: Elaboración propia

3.3. Movilidad Escolar

Dado que el foco del estudio está en la movilidad escolar y cómo esta ha sido modificada producto del SAE, se elabora una variable que indica, en el escenario previo al SAE (2013-2016), los cambios de EE. Además, se considera que todos los postulantes a algún proceso del SAE, cambian también de establecimiento. Esto es útil a la hora de enfocar el análisis en los cambios efectivos de escuela. En la Figura 3.18 se presenta el cambio de escuela con respecto a matrícula total según nivel, dependencia administrativa⁵ y año, para los casos forzado y no forzado, donde forzado significa que en el establecimiento no ofrece el grado al cual es promovido el estudiante el año siguiente (establecimientos que terminan en 6° básico o en 8° básico).

En el caso de 6° básico, los cambios forzados en las dependencias municipal y particular subvencionada son mínimos, mientras que los no forzados varían entre el 15% y 20% de la matrícula, siendo siempre mayor la proporción de cambio en establecimientos municipales.

Para 8° básico el escenario es distinto, en el que es claro que los cambios son mayoritariamente forzados y comprenden gran parte de la matrícula en el caso de las escuelas municipales, donde se aprecia un cambio de hasta el 90% de la matrícula total. Por otro lado, los cambios no forzados son similares por dependencia administrativa. Además, se aprecia una disminución en los cambios forzados para ambas dependencias. Al comparar ambos niveles, se observa que más cambios, independiente de su tipo o dependencia, ocurren en 8° básico. Finalmente, se observa que los establecimientos particulares pagados, para ambos niveles, sólo presentan cambios no forzados, los que se han mantenido invariantes en el tiempo.

La elección de los niveles de cambio de sexto y octavo básico responde a la mayor presencia de datos históricos en estos niveles, además de la información previa de estos estudiantes (ingreso del hogar, por ejemplo). Esto enriquece el análisis, pues permite abordar la problemática desde una mayor cantidad de dimensiones.

Es importante recalcar que los establecimientos particulares pagados no forman parte de los procesos SAE. Estos quedan fuera de la Ley de Inclusión Escolar pues no reciben aportes financieros del estado, por lo que lo que ocurre con este tipo de establecimientos queda fuera del foco de este trabajo. Sin embargo, estos sí se consideran, para no sesgar lo que ocurre en los cambios de escuela que pasan por el SAE, dado que se ha constatado la presencia de cambios de escuela desde y hacia EE particulares pagados, y con respecto a todas las otras dependencias administrativas.

Sistema de Admisión Escolar

El SAE se ha implementado ya en cuatro procesos desde el año 2016. Entregado una base de datos con toda la información de postulaciones y oferta educativa. La Tabla 3.6 muestra la disponibilidad de estos datos por región. La información del proceso de admisión 2019 aún no están disponibles.

⁵1: Municipal, 2: Subvencionado y 3: Privado

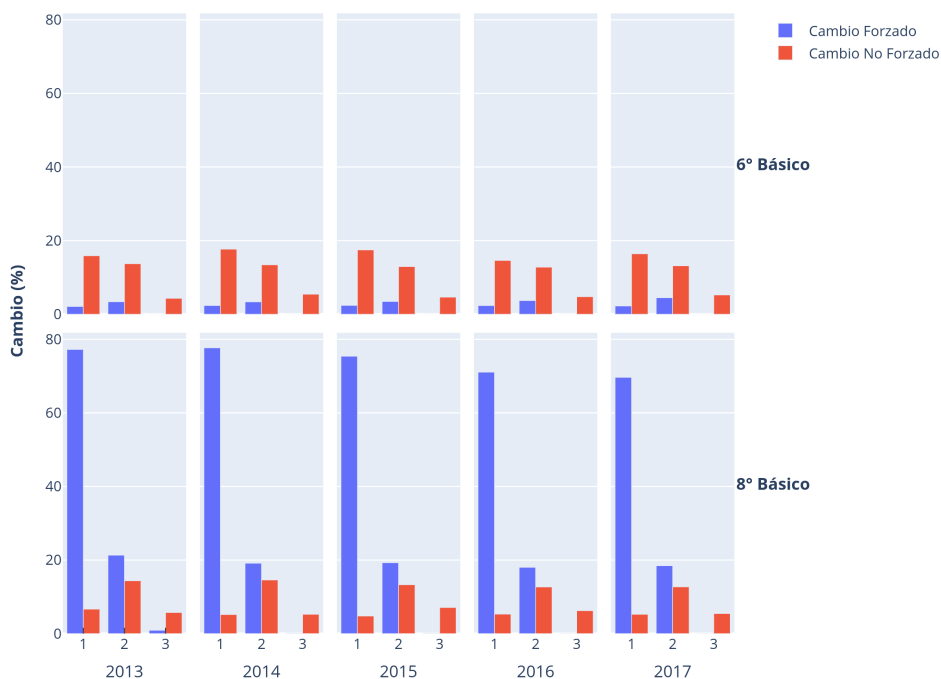


Figura 3.18: Cambio de escuela (%) con respecto a matrícula total según nivel, dependencia administrativa y año. Casos forzado y no forzado.

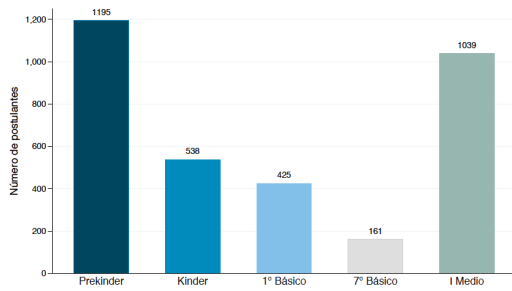
Fuente: Elaboración propia.

Proceso \ Región	Tarapacá	Coquimbo	O'Higgins	Los Lagos	Magallanes	Total
SAE 2016	-	-	-	-	3.358	3.358
SAE 2017	9.059	20.113	22.448	20.382	4.819	76.821
SAE 2018	13.542	27.045	29.111	25.264	5.185	100.147

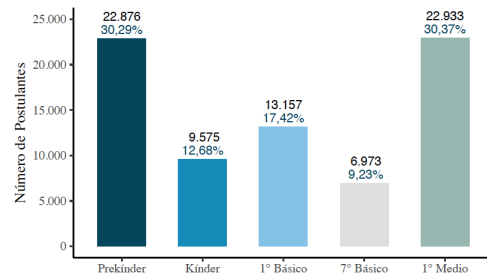
Tabla 3.6: Disponibilidad de datos SAE por región.

En lo que respecta a la demanda educativa registrada durante estos periodos, la Figura 3.19 muestra la cantidad de postulantes para cada proceso SAE, según el nivel y grado al que postulan. Primero medio es un nivel con mucho cambio de EE, comprendiendo alrededor de un tercio de los cambios en todo el sistema. Por otro lado, 7° básico representa a los más el 10% de los cambios totales cada año.

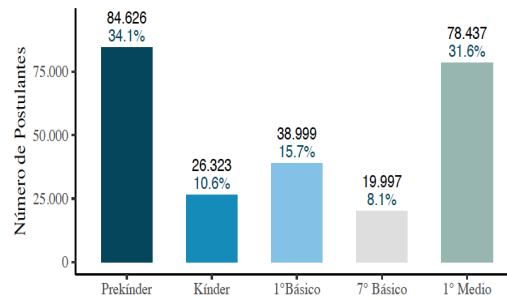
Este capítulo ha buscado caracterizar el escenario educativo en el cual ocurre la reforma en las postulaciones producida por el SAE y la Ley de Inclusión Escolar.



(a) SAE 2016



(b) SAE 2017



(c) SAE 2018

Figura 3.19: Número de postulantes por nivel de ingreso.

Capítulo 4

Metodología

Esta tesis busca identificar patrones de movilidad escolar en diversos contextos educativos, para lo cual se aplica un modelo de agrupamiento de datos, que forma conjuntos, denominados clústers, de tal manera que los objetos (datos) de cada uno sean lo más similar posible entre sí, en tanto que se diferencien los más posible del resto de los grupos. A continuación se describe el marco conceptual bajo el cual se concibe el modelo, así como el uso de estos métodos en problemáticas similares.

4.1. Marco Conceptual

La metodología utilizada en esta investigación forma parte de un conjunto de disciplinas que, si bien es estudiada desde los 70' en el contexto educativo (Shavelson, 1979), en los últimos años han tomado mayor fuerza dado el aumento en capacidad de almacenamiento y cómputo de datos (Magana y Boutin, 2018). Dichas técnicas permiten estudiar aspectos del sistema educativo previamente difíciles de abordar.

4.1.1. Ciencia de datos

La ciencia de datos, o *Data Science*, es un campo multidisciplinario que busca extraer conocimiento a partir de conjuntos de datos, con o sin estructura. Está basada en varias técnicas que utilizan métodos propios de algorítmica, modelos, procesos, sistemas, entre otros; mediante los cuales pretende resolver problemas basados en datos, aprovechando tanto hardware como algoritmos y software (Cao, 2017).

Por otro lado, la ciencia de datos se enmarca en la intersección de diversas disciplinas como la estadística, análisis de datos y aprendizaje automático, con el objetivo de comprender y analizar fenómenos de los cuales existen datos disponibles. En términos teóricos aborda temáticas propias de matemáticas, estadística, *computer science* y teoría de la información (Cao, 2017).

Su uso en contexto educativo abarca un amplio número de campos, en los que se aprovecha la existencia de datos. Baker (2013) explica que las diversas herramientas de esta área traen beneficios significativos, en tanto proveen capacidad (i) predictiva, (ii) de descubrimiento de estructuras y relaciones, y (iii) de destilación de información. Procesos en los que diversos agentes educativos han tomado parte usando información cuantitativa para tomar decisiones, asesorar instituciones y evaluar iniciativas y políticas educativas, entre otras (Klašnja-Milićević et al., 2017; Bernhardt, 2013).

4.1.2. Aprendizaje Automático

El *machine learning*, o aprendizaje automático, es la disciplina que estudia el uso de algoritmos y modelos estadísticos que pueden, mediante computadores, realizar tareas específicas sin instrucciones explícitas, basándose en el reconocimiento de patrones subyacentes en los datos, por ende, es en parte una rama de la Ciencia de Datos. Los algoritmos de aprendizaje automático construyen modelos matemáticos basados en muestras de datos (datos de entrenamiento), para lograr hacer predicciones o tomar decisiones sin ser explícitamente programados para hacerlo, para lo cual se utilizan variadas técnicas de optimización, mediante las que el algoritmo “aprende” a partir de la información capturada en los datos (Hurwitz y Kirsch, 2018).

Los modelos de aprendizaje automático se pueden clasificar en tres grandes categorías:

- Aprendizaje Supervisado: Se trabaja con datos que contienen etiquetas u información sobre resultados esperados, a partir de los cuales el modelo puede “aprender”.
- Aprendizaje No supervisado: Se trabaja con datos que no contienen etiquetas o resultados esperados, por lo que este tipo de modelos sólo busca descubrir estructuras y patrones en los datos.
- Aprendizaje por Refuerzo: este tipo de modelos tiene por objetivo “aprender” maximizando alguna noción de recompensa acumulada.

Todos estos tipos de modelos ha probado ser útil en una amplia gama de aplicaciones, tales como filtros de correo electrónico, visión por computador y procesamiento de lenguaje natural (Hurwitz y Kirsch, 2018), campos en los cuales ha resultado muy difícil e ineficaz desarrollar algoritmos de manera convencional para desarrollar dichas tareas (Lee et al., 2018).

En educación, estos modelos han sido aplicados en múltiples situaciones: para detectar patrones de comportamiento en estudiantes, predecir y entender resultados académicos, investigar similitudes y diferencias entre establecimientos educativos (Baker, 2010), etc.

En particular, el modelo utilizado en esta tesis es del segundo tipo, pues lo que se busca es encontrar patrones en el contexto de movilidad escolar, sin tener información a priori sobre posibles resultados (etiquetas).

Los algoritmos de aprendizaje no supervisado construyen modelos matemáticos a partir de un conjunto de datos que contiene únicamente información concerniente a atributos que exhibe la situación estudiada, sin conocer ningún tipo de resultados o comportamiento espe-

rado (Xu y Tian, 2015). Este tipo de modelos se utilizan para encontrar estructura o patrones en los datos, lo que a su vez permite agrupar observaciones según los criterios de similitud deseados (*clusters*).

El *clustering*, o análisis de *clusters*, es uno de los métodos de aprendizaje no supervisado más utilizados (Xu y Tian, 2015). Esta corresponde a la tarea de segmentar un conjunto de datos de tal manera que aquellos de un mismo grupo sean más similares entre sí (es algún sentido) que respecto a otros grupos. Esta agrupación da paso a categorías, las que son asignadas como etiquetas a los datos. Además de crear los *clusters*, se busca lograrlo de tal manera que estos sean significativos en el contexto estudiado.

Métodos de *clustering* en educación ha sido primariamente utilizados para identificar perfiles de aprendizaje (Bovo et al., 2013), modelar comportamiento de estudiantes (Peckham y McCalla, 2012) y para diseñar estrategias para abordar procesos de aprendizaje (Gašević et al., 2017). También se ha utilizado para predecir desempeño académico, basado en características demográficas, promedio de notas, variables socioemocionales entre otras (Wook et al., 2009).

Por ejemplo, Psaromiligkos et al. (2011) determina perfiles de aprendizaje de estudiantes a través del registro de la actividad online en tareas de aprendizaje. Tie et al. (2010) mediante datos de encuestas a estudiantes de cursos de computación, diseña estrategias para enseñar habilidades computacionales básicas.

En lo que respecta a movilidad estudiantil, Suartana y Hidayat (2017) utiliza algoritmos de *clustering* para procesar y seleccionar estudiantes en un proceso de admisión universitaria, mejorando su eficiencia. En el ámbito escolar, Harris et al. (2007) estudia la segregación escolar en Inglaterra con especial énfasis en las variables geodemográficas y étnicas, para lo cual utiliza el marco de trabajo de Vickers et al. (2005), consistente en *clusters* geodemográficos determinados a partir de variables censales sobre estructura demográfica, composición de hogares, empleo y vivienda.

En el contexto chileno, Vargas (2018) determina *clusters* de estudiantes y de establecimientos de la Región Metropolitana, encontrando que los atributos más importantes son, para EE, la dependencia administrativa y el copago, mientras que para los estudiantes, lo son el género, NSE y tipo de establecimiento.

En relación directa con el presente trabajo, Arce-trigatti et al. (2016) estudia la diferenciación de escuelas públicas en New Orleans (Estados Unidos), definiendo el mercado educativo en base a *clusters* de establecimientos. Estos se generan en bases a atributos tales como: preparación para la universidad (*college prep*), orientación curricular, actividades extraprogramáticas, cantidad de horas de instrucción, proceso de selección, entre otros. Sin embargo, no identifican si la oferta educativa se adapta a los que las familias buscan, y al no incluir características de estudiantes, no logran cuantificar el efectos de las escuelas en sus estudiantes.

Sin embargo, ningún trabajo ha estudiado movilidad escolar mediante la agrupación de cambios de establecimiento, menos aún, considerando las características de establecimientos de origen y destino. Por ende, este trabajo tiene una contribución metodológica, en tanto

aborda la problemática de *school choice* desde una perspectiva anteriormente no considerada.

Además, uno de las mayores dificultades encontradas en trabajos que buscan patrones en datos en contexto educativo, es que no disponen de los suficientes datos para obtener resultados precisos (Jyothi et al., 2012). Lo cual no es un inconveniente en esta tesis, al contar con los datos de cambio de escuela de los últimos 7 años.

4.1.3. Algoritmos de *Clustering*

En esta sección se presentan los tres tipos de modelos de agrupamientos considerados para la problemática de este trabajo.

4.1.3.1. K-MEANS

El algoritmo K-MEANS es un método de agrupación de datos en un número específico de *clusters* disjuntos, el cual es propuesto por primera vez por J. MacQUEEN (1967). El modelo busca minimizar algún tipo de métrica entre el centroide¹ de cada *cluster* y sus elementos mediante un proceso iterativo que reasigna puntos a grupos hasta que la medida de distancia considerada sea la menor posible.

K-MEANS es uno de los algoritmos más utilizados en aprendizaje no supervisado. Sus principales ventajas son su fácil comprensión e implementación, además de ser eficiente para instancias usuales. Sin embargo, este modelo sólo aplica si la noción de distancia existe, por lo que no es directamente aplicable a datos categóricos. Por otro lado, es necesario especificar el número de *cluster*, y es sensible a *outliers*. Además, el modelo sólo encuentra *cluster* hipersféricos, por lo que no es ideal para situaciones donde la geometría de los datos no cumpla esta regla. Por ende, escenarios ideales para K-MEANS, son aquellos con datos con bajo número de dimensiones, con atributos numéricos y continuos. A pesar de esto, no existe evidencia clara de que algún otro algoritmo tenga un mejor desempeño en general.

Variadas aplicaciones se han encontrado para este algoritmo, como lo son la agrupación temática de documentos de texto, identificación de sectores más susceptibles a crímenes, segmentación de clientes, detección de fraude, análisis de transporte público, entre muchos otros.

Formalmente, dado un conjunto de datos finito con m atributos, $\mathcal{D} = \{x_1, \dots, x_n\}$, tal que $x_i \in \mathbb{R}^m$ y el número de *clusters* buscados, k , el algoritmo aprende una partición de \mathcal{P} de tal manera que los datos en cada agrupación sean lo más similar posible, lo que se traduce en el siguiente problema:

$$\min_{C_1 \cup C_2 \cup \dots \cup C_k = \mathcal{D}} \sum_{i=1}^k \sum_{x \in C_i} \left\| x - \frac{1}{|C_i|} \sum_{x_j \in C_i} x_j \right\|^2$$

¹Centro geométrico del conjunto de datos.

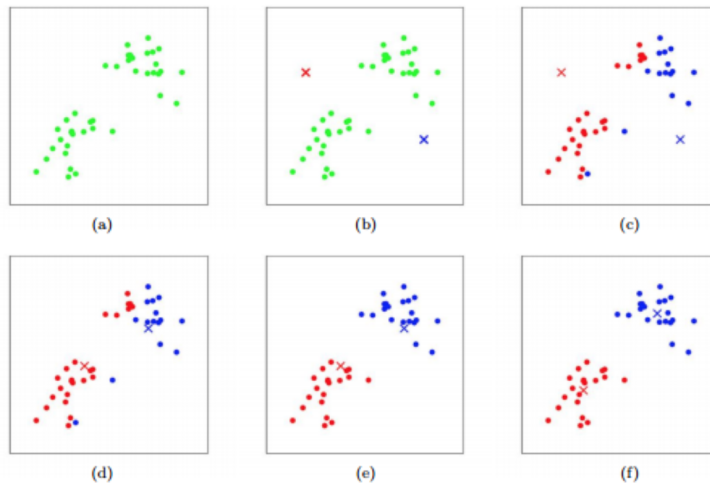


Figura 4.1: Ejemplo de entrenamiento de K-MEANS. Cada iteración reasigna puntos hasta convergencia.

Fuente: The Learning Machine (2019)

Donde $\|\cdot\|$ corresponde a la distancia euclidiana.

Resolviendo este problema de optimización, el algoritmo predice k centroides y etiquetas $c^{(i)}$ para cada $x_i \in \mathcal{D}$. Sin embargo, sólo está garantizada su convergencia a mínimos locales, ya que el problema es NP-difícil (Mahajan et al., 2009), aunque existen implementaciones de heurísticas eficientes que convergen rápidamente a algún óptimo local. Es por esto, que se el algoritmo se ejecuta múltiples veces con distintas condiciones iniciales, de tal manera que aumente la probabilidad de alcanzar el máximo global.

La Figura 4.2 muestra una instancia de K-MEANS con datos en dos dimensiones. En esta, los datos son representados por puntos y los centroides por cruces. Se muestran dos iteraciones del algoritmo, en cada una de la cuales se asignan los puntos al *cluster* de centroide más cercano, para posteriormente ajustar dicho centroide a la media de los nuevos puntos del grupo.

El pseudocódigo de la implementación básica del algoritmo se puede ver en la Figura Algoritmo 1.

4.1.3.2. Modelos Mixtos Gaussianos (GMM)

Este tipo de modelos asume la existencia de un determinado número de distribuciones gaussianas, cada una de las cuales representa un *cluster*. Así, para un conjunto de datos, el modelo GMM determina la probabilidad de pertenencia de cada punto a cada una de estas distribuciones (*clusters*). Por consiguiente, cada *cluster* es descrito en términos de una densidad gaussiana, mediante un centroide y una matriz de covarianza. Entrenar un modelo GMM implica entonces encontrar los parámetros de estas distribuciones que mejor modelen los datos.

Algoritmo 1 K-MEANS(X, k)

- 1: **Entrada:** Datos (X) y número de *clusters* (k)
 - 2: **Inicializar:** Centros aleatorios $c_1, \dots, c_k \in \mathbb{R}^m$ ▷ m número de atributos
 - 3: **Repetir:**
 - 4: Encontrar centro más cercano:
 - 5: $\forall x \in X : b(x) = \operatorname{argmin}_{j=1, \dots, k} d(x, c_j)$ ▷ $d(\cdot, \cdot)$ distancia euclidiana
 - 6: Actualizar centros c_1, \dots, c_k :
 - 7: $\forall j \in 1, \dots, k : \operatorname{mean}(x \in X \mid b(x) = j)$ ▷ $\operatorname{mean}(\cdot)$ media
 - 8: **Hasta:** centros no cambien \vee
puntos permanezcan en el mismo *cluster* \vee
máximo número de iteraciones se alcanza.
 - 9: **Salida:** Centros c_1, \dots, c_k y etiquetas $b(x), \forall x \in X$.
-

Este proceso utiliza el enfoque de esperanza-maximización, el que realiza lo siguiente:

- Hacer suposición inicial de parámetros del modelo.
- Repetir hasta convergencia:
 - Paso Esperanza (E): Para cada punto, hallar los pesos que determinan la probabilidad de pertenencia a cada *cluster*.
 - Paso Maximización (M): Para cada *cluster*, actualizar su ubicación (centro), normalización y forma basado en todos los puntos, haciendo uso de los pesos.

El resultado de este procedimiento es que cada agrupación ya no es asociado con una esfera rígida como en K-MEANS, sino que lo hace con un modelo gaussiano "suave". Si embargo, este modelo también podría converger a máximos locales. por lo que en la práctica se deben utilizar múltiples inicializaciones aleatorias para asegurar la obtención del óptimo.

Formalmente, un modelo mixto gaussiano de K componentes está definido, según los parámetros de cada una de estas: media μ_k , varianza σ_k (matriz de covarianza Σ_k en el caso multivariado) y peso ϕ_k , tal que $\sum_k \phi_k = 1$. De esta manera, el modelo queda definido de la siguiente manera:

$$p(x) = \sum_{i=1}^K \phi_i \mathcal{N}(x \mid \mu_i, \sigma_i)$$
$$\mathcal{N}(x \mid \mu_i, \sigma_i) = \frac{1}{\sigma_i \sqrt{2\pi}} \exp\left(-\frac{(x - \mu_i)^2}{2\sigma_i^2}\right)$$
$$\sum_k \phi_k = 1$$

Buscando estimar estos parámetros, se siguen los siguientes pasos:

Inicialización de parámetros

Para todo $k = 1, 2, \dots, K$:

- Elección sin reposición de $\hat{\mu}_k$ a partir de $X = \{x_1, x_2, \dots, x_N\}$ (conjunto de datos).
- Establecer parámetros $\hat{\sigma}_k = \frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})^2$, donde $\bar{x} = \frac{1}{N} \sum_{i=1}^K x_i$
- Establecer parámetros $\hat{\phi}_k = \frac{1}{K}$

Paso de Esperanza (E)

Calcular $\forall i, k$:

$$\hat{\gamma}_{ik} = \frac{\hat{\phi}_k \mathcal{N}(x_i | \hat{\mu}_k, \hat{\sigma}_k)}{\sum_{j=1}^K \hat{\phi}_j \mathcal{N}(x_i | \hat{\mu}_j, \hat{\sigma}_j)}$$

Donde $\hat{\gamma}_{ik}$ es la probabilidad que x_i provenga de la componente C_k , es decir, $\hat{\gamma}_{ik} = p(C_k | x_i, \hat{\phi}, \hat{\mu}, \hat{\sigma})$

Paso de Maximización (M)

Usar los $\hat{\gamma}_{ik}$ para calcular lo siguiente $\forall k$:

$$\begin{aligned} \hat{\phi}_k &= \sum_{i=1}^N \frac{\hat{\gamma}_{ik}}{N} \\ \hat{\mu}_k &= \frac{\sum_{i=1}^N \hat{\gamma}_{ik} x_i}{\sum_{i=1}^N \hat{\gamma}_{ik}} \\ \hat{\sigma}_k^2 &= \frac{\sum_{i=1}^N \hat{\gamma}_{ik} (x_i - \hat{\mu}_k)^2}{\sum_{i=1}^N \hat{\gamma}_{ik}} \end{aligned}$$

El caso multivariado es análogo.

Al igual que en K-MEANS, dado que el número de *clusters* debe ser especificado, se entrena el modelo para muchos valores distintos, escogiendo aquel que mejor “equilibre” ajuste y número de componentes.

4.1.3.3. *Clustering* Jerárquico Aglomerativo (HAC)

Al contrario de los modelos anteriores en los que se debe especificar el número de *clusters* buscado, los modelos de agrupamiento jerárquico no tienen esta restricción. En cambio, estos requieren únicamente especificar una métrica de disimilitud entre grupos de observaciones. Como lo sugiere su nombre, estos modelos forman representaciones jerárquicas en los que los *clusters* son creados a cada nivel mezclando las observaciones correspondientes de niveles inferiores. En el nivel más bajo, cada *cluster* está formado por una única observación, en tanto que el nivel más alto, comprende un único grupo formado por todos los datos considerados.

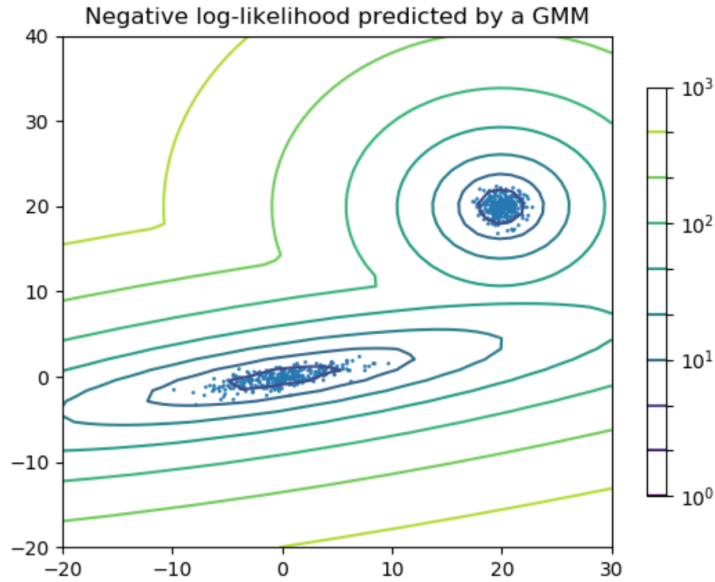


Figura 4.2: Modelo Gaussiano Mixto de dos componentes: datos y superficies equiprobables.
Fuente: Marwala (2018)

Métrica	Expresión	Descripción
Single linkage (SL)	$d_{SL}(G, H) = \min_{i \in G, j \in H} d_{i,j}$	La menor disimilitud entre puntos
Complete linkage (CL)	$d_{CL}(G, H) = \max_{i \in G, j \in H} d_{i,j}$	La mayor disimilitud entre puntos
Group average (GA)	$d_{GA}(G, H) = \frac{1}{ G \cdot H } \sum_{i \in G} \sum_{j \in H} d_{i,j}$	La disimilitud promedio

Tabla 4.1: Medida de disimilitud en modelos de *clustering* aglomerativo jerárquico.

La estrategia de *clustering* jerárquico aglomerativo comienza en el nivel más bajo, mezclando de manera recursiva pares de *clusters* en uno solo. Esto produce un agrupamiento en el siguiente nivel con un *cluster* menos. Los grupos que se escogen para ser fusionados son aquellos que presentan la menor disimilitud entre ellos.

Formalmente, sean G y H dos grupos en algún nivel del agrupamiento. Se calcula la disimilitud $d(G, H)$ entre ambos grupos a partir de las disimilitudes entre sus pares de puntos $d_{i,j}, \forall i \in G, j \in H$. Las principales medidas de disimilitud se resumen en la Tabla 4.1

Este esquema de agrupamiento se puede representar gráficamente mediante un árbol binario, de tal manera que la altura de cada nodo sea proporcional a la disimilitud entre los grupos hijos de dicho nodo. Los nodos terminales (observaciones individuales) se consideran el nivel cero del mapa. Este tipo de figuras recibe en nombre de dendograma. La figura 4.3 muestra ejemplos de dendograma para cada uno de los tipos de disimilitud especificados.

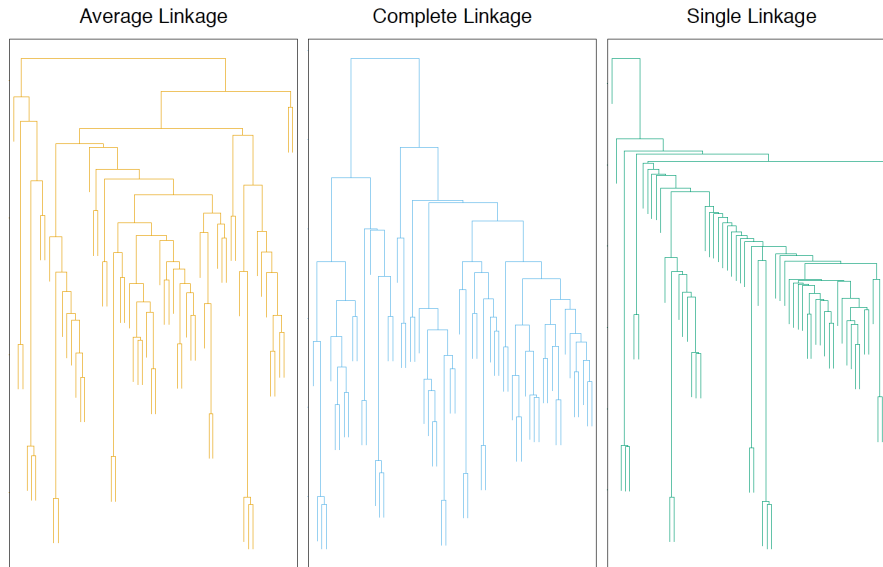


Figura 4.3: Ejemplo de dendrograma para cada medida de disimilitud.
Fuente: Hastie et al. (2001)

4.1.4. Elección del Número de *Clusters*

En el contexto de modelos de aprendizaje no supervisado, entrenar un modelo significa encontrar la mejor agrupación de los datos, dado un cierto número de *clusters*. Sin embargo, la elección de este número es en sí un proceso aparte, ya que, al no existir etiquetas que guíen los resultados, se deben utilizar diversos métodos para asegurar que la elección sea lo más significativa posible.

La correcta elección del número de *clusters* es normalmente ambigua, con interpretaciones que dependen de la forma y la escala de distribución de los datos, y del tipo de resultados que se buscan. Es por esto, que además es necesario el conocimiento y criterio experto a la hora de tomar la decisión.

Hay que tener presente que si se aumenta el número de agrupaciones sin penalización, siempre disminuirá la varianza interna de cada grupo, lo que llevaría al caso límite en el que cada punto es un *cluster*. Entonces, que se deben balancear la compresión total de los datos (*cluster* único), con el escenario opuesto de máxima precisión (tantos *clusters* como número de datos). Es por esto, que existen numerosos métodos para informar esta elección.

En este estudio, para evaluar la calidad de los *clusters*, se utilizan medidas que consideran dos criterios: (i) cantidad de varianza explicada y su aumento marginal, y (ii) balance entre cohesión intra-*clusters*, que cuantifica la cercanía de los puntos en un mismo *cluster*, y separación inter-*cluster*, que considera la distancia entre centroides. A continuación, se presentan los métodos utilizados.

4.1.4.1. El Método *Elbow*

El método *elbow*, propuesto por Thorndike (1953), identifica cuánta varianza es explicada como función del número de *clusters*, la cual se representa como porcentaje. Según este criterio, a partir del primero, cada *cluster* añadido contribuirá con más información (explica más varianza), hasta que en algún punto, la ganancia marginal disminuirá. Es este punto el que se escoge como número óptimo de *clusters*. Es decir, se debe escoger un número de *clusters* de tal manera que si se añade uno más, no mejora considerablemente la calidad de segmentación de los datos.

Formalmente, se grafican puntos $(k, Var(k))$ tal que:

$$Var(k) = \sum_{i=1}^n \min_{\mu_j \in C = \{C_1, \dots, C_k\}} (\|x_i - \mu_j\|^2)$$

Donde cada μ_j corresponde al centroide del *cluster* $C_j \in C$, siendo C el conjunto de *clusters*. Esta varianza se conoce como inercia, la cual se interpreta como una medida de coherencia interna de cada *cluster*. la Figura 4.4 muestra un ejemplo de este método: a la izquierda el gráfico de varianza explicada (inercia) y a la derecha la respectiva agrupación de los datos según el óptimo seleccionado (3).

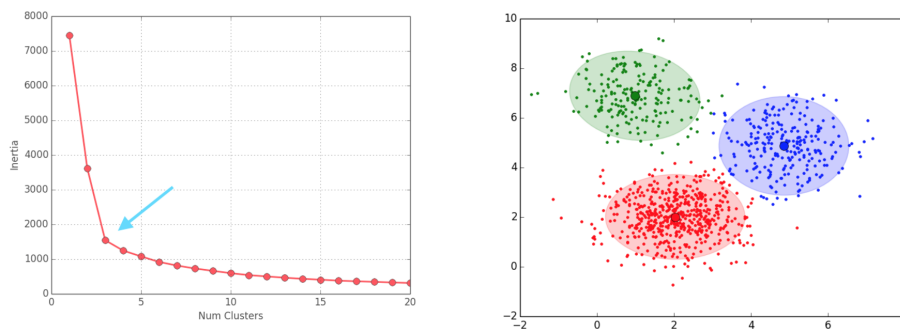


Figura 4.4: Ejemplo de Método *elbow*
Fuente: Moya (2016)

4.1.4.2. El Método de Silueta

La silueta promedio es otro criterio útil para escoger la cantidad de *clusters*. La silueta de los datos para una cierta agrupación, es una medida que indica, por un lado, qué tan cercanos están los puntos de cada *cluster*, y por otro, qué tan lejanos estos se encuentran del grupo más cercano.

Formalmente, para un conjunto de datos agrupados en K *clusters*, para cada punto $x_i \in C_i \in C$, se define $a(i)$ como la distancia media del punto a resto de puntos del mismo *cluster*, la cual se interpreta como una medida de que tan bien el punto x_i está asignado a su *cluster*. Valores pequeños de $a(x_i)$ son mejores.

$$a(x_i) = \frac{1}{|C_i| - 1} \sum_{x \in C_j, j \neq i} \|x_i - x\|$$

Por otro lado, para cada punto x_i , se define como medida de separación, a la menor distancia media del punto a los puntos de algún otro *cluster*.

$$b(x_i) = \min_{k \neq i} \frac{1}{|C_k|} \sum_{x \in C_k} \|x_i - x\|$$

En base a estos concepto, se define el puntaje de silueta para un punto x_i :

$$s(x_i) = \frac{b(x_i) - a(x_i)}{\max\{a(x_i), b(x_i)\}}, \text{ si } |C_i| > 1$$

Y $s(x_i) = 0$ cuando $|C_i| = 1$. Esta restricción de puntaje nulo para grupos de tamaño unitario es importante pues previene el aumento del número de *clusters*.

Finalmente, se define el coeficiente de silueta como la media de los puntajes de silueta sobre todos los puntos del conjunto de datos. Como número óptimo de *clusters*, se escoge aquel que maximice dicho coeficiente, es decir, aquel en el cual se equilibra la cohesión de los puntos de un mismo *cluster* y la separación entre estos. La Figura 4.5 muestra un ejemplo del coeficiente de silueta graficado, del cual se desprende que el número óptimo de *clusters* para la situación es 3.

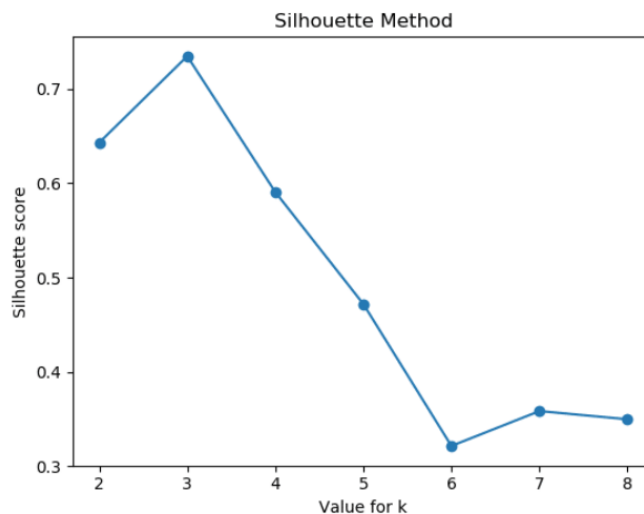


Figura 4.5: Ejemplo de Puntaje de silueta para todo los datos.

Fuente: Subramanian (2019)

Además, el puntaje de silueta se puede graficar para cada grupo, lo que entrega una visión más detallada sobre la calidad de cada agrupación. La Figura 4.6 muestra el puntaje de silueta para cada observación, las que se agrupan por *cluster* de manera ascendente (izquierda). En

la misma Figura se aprecia el conjunto de datos coloreados según *cluster*, con sus respectivos centroides (derecha).

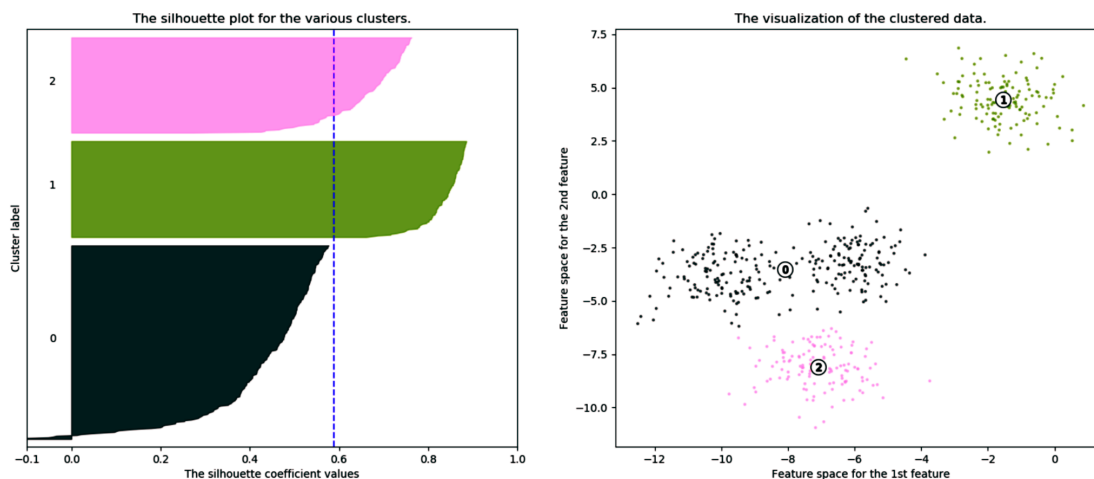


Figura 4.6: Puntaje de silueta con respecto a cada *cluster*.

Fuente: Dalwani (2015)

Cada uno de los métodos recién descritos son aplicables a cada tipo de modelo especificado. De esta manera, se escoge el número de *clusters* a utilizar en cada escenario, luego se entrena el modelo definitivo, y finalmente se obtienen las etiquetas para los datos, a través de las cuales se lleva a cabo el análisis de los resultados y patrones que emerjan.

4.2. Implementación

Utilizando los datos de establecimientos y estudiantes, explicados en el capítulo anterior, en una primera etapa, se determina la cantidad óptima de *clusters* para cada escenario, es decir, para cada una de las conurbaciones consideradas, tanto para el caso de cambios forzados de EE como el caso de cambios no forzados. Se incluyen además, el escenario previo al SAE y aquel con el nuevo sistema ya en marcha. La Tabla 4.2 resume los 24 escenarios que serán analizados en este trabajo. Luego, en una segunda etapa, con la finalidad de identificar tendencias de cambio según características de establecimientos de origen y destino, los modelos se entrenan y aplican en cada caso².

Un aspecto técnico relevante, es que dado que tanto KMEANS como GMM se basan en distancias para generar las agrupaciones, se debe estandarizar cada uno de los atributos para que estos tengan igual peso en los cálculos respectivos. Por otro lado, al ser la dependencia administrativa una variable categórica, este debe ser binarizado, es decir, se definen tres nuevas columnas, indicatrices de cada una de las dependencias consideradas, esto para todos los modelos. Una vez los modelos son entrenados, se aplica el proceso inverso a cada variable,

²En la etapa de preparación de datos se utiliza la librería Pandas, para los procesos de entrenamiento y evaluación del modelo se utiliza la librería Scikit-learn (Pedregosa et al., 2011). Todo lo anterior con base en el lenguaje de programación Python.

par así mostrar los centroides en su escala original, dando mayor sentido a los resultados presentados.

Conurbación	Casos	Escenarios	Total
Iquique - Alto Hospicio	- Forzado - No Forzado	- Pre SAE	24 casos totales
Coquimbo - La Serena		- SAE Preferencias	
Rancagua - Machalí		- SAE Admisión	
Puerto Montt - Puerto Varas			

Tabla 4.2: Escenarios en los que se aplica el modelo.

Fuente: Elaboración propia.

La segmentación anterior permite abordar con mayor profundidad la problemática, pues, por un lado, toma en cuenta la particularidad geográfica de cada región, y además evalúa de forma diferenciada contextos en los cuales la elección educativa es distinta (casos forzado y no forzado). Por otro lado, al abordar periodos con distinto régimen de funcionamiento (con SAE y sin SAE), se captura empíricamente las modificaciones que han ocurrido en la manera en la que las familias escogen colegios.

Cabe señalar que en el escenario pre SAE solo se observan las preferencias mediadas por los resultados de los procesos de postulación (no observables), por lo cual no se conocen las preferencias reales. En tanto que en los escenarios SAE las situaciones son distinguibles: se identifican por un lado las preferencias reales de las familias (SAE Preferencias) y por otro los resultados de la admisión (SAE Admisión). Así, en estricto rigor lo correcto es comparar los escenarios pre SAE y SAE Admisión. En tanto que el escenario SAE Preferencias entrega información de lo que los padres eligen, variante que no ha sido estudiada con anterioridad, por lo que destaca como uno de los aporte importante de esta investigación.

Estos patrones de cambios son determinados en base a las características de los establecimientos de origen y destino de aquellos estudiantes que cambian de EE. Por tanto, los *clusters* identificados corresponden a perfiles (EE origen, EE destino) con atributos particulares. En este sentido, los *clusters* se pueden interpretar como “tubos” que conectan ciertos tipos de establecimientos, a través de los cuales los estudiantes transitan en su proceso de cambio de escuela. En el Apéndice G se presentan los códigos más relevantes en el desarrollo de este trabajo.

Para cada caso de análisis se presenta la siguiente información:

- Figuras que representan distintas perspectivas de las aglomeraciones, cada una de las cuales se identifica con un color característico del tipo de cambio que señala:
 - Diagrama de coordenadas paralelas: gráfico con seis ejes verticales, los que presentan los valores promedio para cada *cluster* en las variables de copago, IVE y puntaje promedio SIMCE, tanto de los establecimientos de origen como de destino de los estudiantes de cada agrupación. Es importante tener en cuenta que sus ejes tienen distintas escalas, por lo que no tiene sentido considerar tendencias de crecimiento o decrecimiento en dirección horizontal, cuando se comparan ejes que representan variables diferentes.
 - Flujo de cambio entre dependencias administrativas: tipo particular de diagrama de coordenadas paralelas, que contiene dos ejes verticales, los que representan las

dependencias administrativas de los establecimientos de origen (izquierda) y destino (derecha). El grosor de cada flujo representa la cantidad relativa de estudiantes que pertenece al *cluster* respectivo.

- Gráfico circular: representa el tamaño relativo de cada aglomeración. Cada sector circular incluye además el número absoluto de estudiantes caracterizados por el *cluster*.
- Tabla que contiene en sus filas el centro de cada uno de los *clusters*. El copago se presenta en términos de percentiles con respecto a los establecimientos de la misma aglomeración, mientras que el promedio SIMCE se presenta también en percentiles, pero con respecto a los resultados de todo el país. Las variable IVE toma valores entre 0 y 1, donde el valor 0 representa nula vulnerabilidad y el valor 1 representa que el total de la población escolar es vulnerable. Finalmente, la variable de dependencia administrativa se presenta como una combinación de las dependencias que definen a cada *cluster*, tanto para establecimiento de origen como de destino.
- Gráfico radar de características de estudiantes: Gráfico circular de siete ejes repartidos homogéneamente, cada uno de los cuales representa una variable a nivel de estudiante. Cada *cluster* se grafica con su respectivo color, poniendo el valor promedio de dichas variables en cada eje. Estas variables son:
 1. No SEP: “1 - porcentaje de estudiantes del *cluster* que son beneficiarios SEP”. Valores hacia afuera representan menor composición de estudiantes beneficiarios SEP en el *cluster*. Es importante tener presente que esta ley amplió su cobertura a partir del año 2016, por lo que escenarios Pre SAE no son comparables en esta variable con los escenarios SAE.
 2. Percentil de promedio de notas: con respecto a los estudiantes del mismo establecimiento, nivel y año.
 3. Porcentaje Masculino: cantidad relativa de estudiantes de género masculino.
 4. Distancia hogar colegio: medida en metros.
 5. Puntaje SIMCE estudiante: promedio en pruebas de lenguaje y matemáticas en los casos que el estudiante rinda la prueba.
 6. Educación de los Padres: nivel de estudios alcanzado en promedio por ambos padres. Dado que es una variable categórica, al hablar de promedio de Estudio de Padres en un cierto *cluster*, se hace referencia al nivel de estudio en torno al cual se encuentran los padres de los estudiantes de dicha agrupación.
 7. Ingreso Hogar: ingreso promedio del hogar de cada estudiante.

En todos los caso, para las dimensiones de vulnerabilidad (IVE y porcentaje de Beneficiarios SEP), se representan en los gráficos los valores de dichas variables sustraídos a la unidad, es decir, “uno menos la variable”. Esto con la finalidad de que valores mayores para cada atributo tomen mayor sentido visual en el contexto estudiado. Por ejemplo, valores de “No Vulnerabilidad” altos van de la mano (muy frecuentemente) con mayor copago y mayores resultados SIMCE, lo que se visualiza claramente al quedar todos estos puntos sobre el resto de los valores de otros *clusters*.

Es importante recalcar que, si bien los colores identifican a los *clusters* según las transiciones entre dependencias administrativas que estos representen, *clusters* con el mismo color

no son iguales y comparables en todos sus aspectos. Por ejemplo, en la aglomeración urbana Iquique - Alto Hospicio, en el caso de cambios Forzados, tanto el *cluster* A del escenario Pre SAE como el de escenario SAE Preferencias se identifican con el color rojo y representan cambios entre establecimientos municipales, sin embargo, en el primero la movilidad es (en promedio) entre escuelas con IVE de 0,74 y 0,78, mientras que en el segundo caso, esta ocurre entre establecimientos con valores de 0,69 y 0,77 para la misma variable.

Finalmente, se deben considerar las limitaciones estadísticas de este tipo de modelos. Si bien al agrupar datos los *clusters* resultantes son más homogéneos que el conjunto total de datos, aún existe una varianza interna en cada uno que indica grados de desorden para cada atributo de la agrupación. Por ende, al resaltar como resultado el centroide de cada *cluster*, se hace referencia a este como un patrón representativo del grupo, pero no a una estimación robusta de las características de este. Por ejemplo, los cambios de un cierto *cluster* caracterizado por un IVE medio de establecimiento de origen de 0,6, puede tener una distribución de este atributo muy dispersa en el rango 0,3 - 0,8. No obstante, los patrones que se rescatan son significativos en el sentido que dan luces sobre tendencias previamente desconocidas en el escenario de movilidad escolar.

Capítulo 5

Resultados

La caracterización de perfiles de cambio se lleva a cabo considerando variables de los establecimientos de origen y destino de cada estudiante que cambia de escuela. De esta manera, los *clusters* resultantes, representan los flujos principales de estudiantes que se movilizan entre determinados perfiles de escuela. La Tabla 5.1 muestra los atributos considerados a la hora del entrenamiento del modelo. Se toman en cuenta las dimensiones de rendimiento académico (SIMCE), dependencia administrativa, copago y vulnerabilidad (IVE).

Variable	Descripción
SIMCE_rbd_origen	Puntaje promedio en prueba SIMCE de EE de origen
SIMCE_rbd_elegido	Puntaje promedio en prueba SIMCE de EE de destino
cod_depe_rbd_origen	Dependencia administrativa EE de origen
cod_depe_rbd_elegido	Dependencia administrativa EE de destino
copago_rbd_origen	Copago EE de origen
copago_rbd_elegido	Copago EE de destino
IVE_rbd_origen	Vulnerabilidad EE de origen
IVE_rbd_elegido	Vulnerabilidad EE de destino

Tabla 5.1: Atributos para modelo de *clustering*.

Esta elección de atributos se llevó a cabo mediante un análisis exploratorio de las diversas combinaciones de variables de las que se disponen, poniendo énfasis en la calidad de los *clusters* y su grado interpretación en el contexto estudiado. Así, se utilizan datos que representan cambios de EE, es decir, un punto en el *dataset* representa a un estudiante, en un año determinando, con las características de su escuela durante ese periodo, y además, las características de su escuela de destino (del año siguiente).

En lo que respecta a la elección de modelo, se realizan entrenamientos de prueba en varios escenarios y para múltiples configuraciones de parámetros para cada uno de los modelos (KMEANS, GMM, HAC), obteniéndose resultados muy similares en los casos óptimos respecto al número de *clusters*, aunque levemente superiores (del orden de 5%) para el modelo KMEANS. Por lo cual, dada su mayor simpleza, interpretabilidad y mejores resultados, se escoge KMEANS para implementar en todos los escenarios. Además, el obtener resultados similares para modelos con distintos paradigmas de base, indica que los resultados obtenidos son robustos.

De esta manera, utilizando la metodología presentada en el capítulo anterior, se optimiza el modelo KMEANS para cada uno de los escenarios estudiados¹ (Tabla 4.2). La Tabla 5.2 muestra los resultados de dicho proceso². Cabe recordar que “SAE Preferencias” corresponde al escenario de estudio que considera el periodo en que ya está implementado el SAE (procesos de ingreso 2018 y 2019), en el que además se considera con establecimiento de destino aquel marcado como primera preferencia por los padres. Asimismo, SAE Admisión considera, para los procesos SAE, el resultado del algoritmo de asignación. Se observa que el número óptimo de aglomeraciones para todos los casos y escenarios es de 4 o 5.

Conurbación	Caso	Escenario		
		Pre SAE	SAE Preferencias	SAE Admisión
Iquique - Alto Hospicio	Forzado	5	4	4
	No Forzado	4	4	4
Coquimbo - La Serena	Forzado	4	4	4
	No Forzado	4	5	5
Rancagua- Machalí	Forzado	4	4	4
	No Forzado	5	5	4
Pto. Montt - Pto. Varas	Forzado	4	4	4
	No Forzado	5	5	5

Tabla 5.2: Número óptimo de *clusters* para cada caso estudiado

Cabe señalar que se ha escogido considerar, por un lado, todos los años Pre SAE de manera simultánea, es decir, la base de datos de entrenamiento de los modelos contienen los eventos de cambio de escuela de todos los periodos Pre SAE (2013-2017). Esto con la finalidad de capturar lo esencial y permanente del comportamiento de elección de escuela. Por otro lado, también se consideran ambos periodos SAE simultáneamente, esto con la misma motivación anterior. Sin embargo, se analizaron ambos procesos SAE por separado para explorar diferencias significativas que evidenciaran la necesidad de considerar ambos años por separado. No obstante, se constata que existen diferencias importantes en los resultados. En el Apéndice E se presenta el análisis de ambos procesos por separado.

A continuación se presentan los resultados para cada uno de los casos y escenarios estudiados, cuya estructura se muestra en el diagrama de árbol de la Figura 5.1. Se incluyen conclusiones generales por caso y por aglomeración urbana. En el Apéndice F se exponen los resultados en conjunto por cada aglomeración y caso.

¹La aglomeración urbana de Punta Arenas fue descartada de este análisis por tener una cantidad muy baja de habitantes (130.000 aproximadamente, menos de la mitad que las otras aglomeraciones), lo que indica una menor oferta y demanda educativa en esta, por lo cual se obtienen patrones distintos y que responden a otros factores. En este sentido, según evidencian los resultados de la optimización del número de *clusters*, los datos de esta comuna no poseen la suficiente estructura para que se pueda definir un número razonable de agrupaciones, las que sean además interpretables. Más detalles en el Apéndice D.5.

²Detalles se presentan en el Apéndice D.

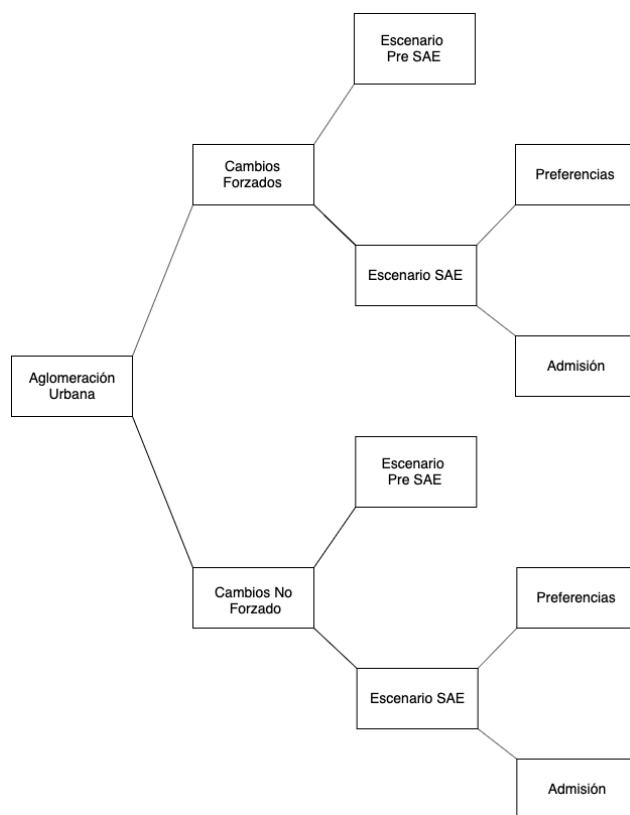


Figura 5.1: Diagrama de resultados para cada una de las cuatro aglomeraciones urbanas consideradas.

5.1. Iquique - Alto Hospicio

5.1.1. Cambios Forzados

5.1.1.1. Escenario Pre SAE

En este caso hay cinco *clusters*. A partir de la Figura 5.2 (b), se observa que estos quedan muy bien definidos por la relación entre dependencias administrativas de origen y destino, a excepción de *cluster* E, que incluye un pequeño flujo desde y hacia establecimientos particulares pagados. En las Figuras 5.2 (b) y (c), se observa que el *cluster* A tiene la mayor representación porcentual de estudiantes del escenario (37%), mientras que los *clusters* B y C tienen la menor (10% y 9% respectivamente). El *cluster* C comprende el 27% de los cambios y el *cluster* E el 18% de estos.

Cluster	Copago rbd origen	Copago rbd destino	IVE rbd origen	IVE rbd destino	Promedio SIMCE rbd origen	Promedio SIMCE rbd destino	Dependencia rbd origen	Dependencia rbd destino
A	P60 (Gratuito)	P60 (Gratuito)	0,74	0,78	P15	P6	Municipal	Municipal
B	P60 (Gratuito)	P73	0,76	0,69	P14	P40	Municipal	Part. Subv.
C	P63	P61 (Gratuito)	0,75	0,77	P41	P36	Part. Subv.	Municipal
D	P64	P68	0,76	0,71	P40	P45	Part. Subv.	Part. Subv.
E	P90	P90	0,38	0,34	P73	P76	Part. Subv.	Part. Subv.

Tabla 5.3: Centroides de clusters.

La Tabla 5.3 y la Figura 5.2 (a) muestra los centroides de cada *cluster*. Se observa que el Copago no cambia (entre el EE de origen y destino) en los *clusters* A, C, D y E, teniendo este último (*cluster* E) los valores más elevados. En tanto que la movilidad representada por el *cluster* B aumenta términos de copago, lo que es consistente con el cambio desde EE municipal hacia particulares subvencionados que representa. Asimismo, la vulnerabilidad (IVE) permanece constante en el *cluster* E, disminuye en los *clusters* B y D, y aumenta levemente en los *clusters* A y C. En términos de resultados SIMCE de establecimientos de origen y destino, cada uno de los *clusters* C, D y E presenta valores invariantes; por otro lado, en el *cluster* A hay una disminución desde rendimiento bajo hacia un rendimiento aún menor, en tanto que sólo en el *cluster* B estos puntajes aumentan.

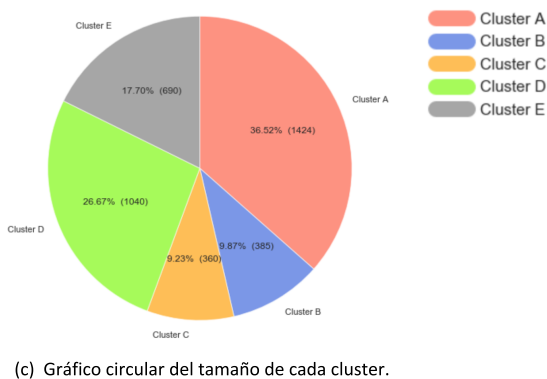
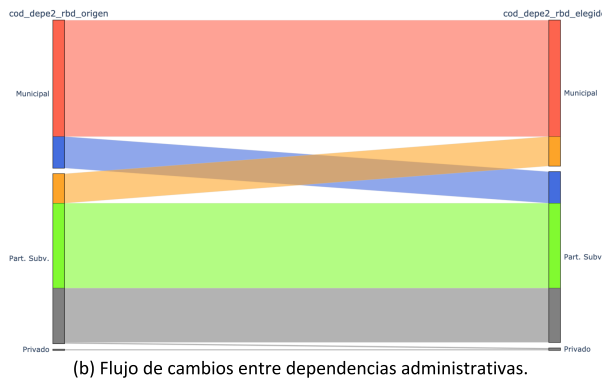
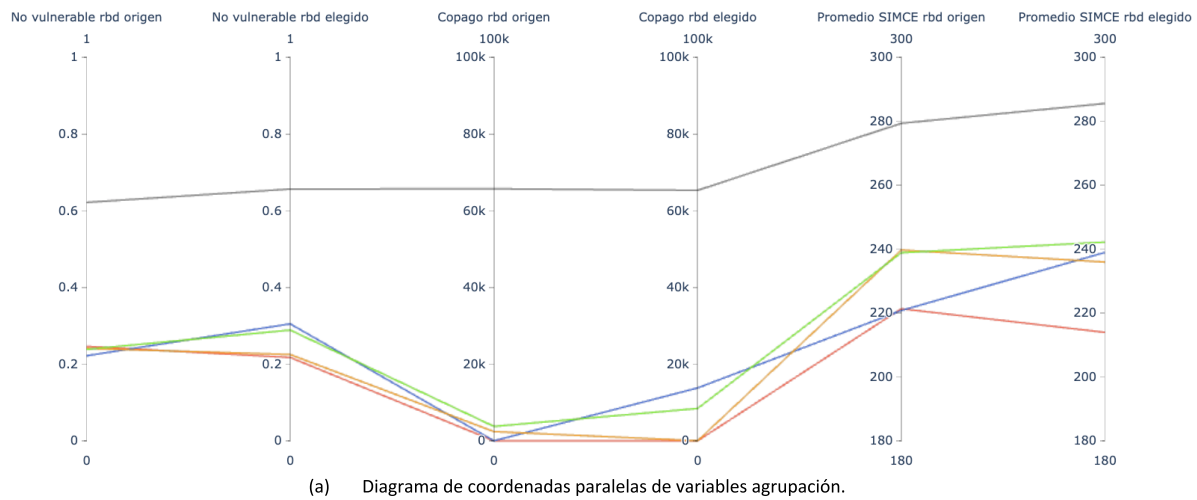


Figura 5.2: Variables a nivel de establecimiento.

Al comparar las variables a nivel de estudiante, en la Figura 5.3 se observa, para la variable de vulnerabilidad (No participa en la SEP), que los *clusters* A y B son los más vulnerables (cerca del 35% de los estudiantes no son Beneficiarios SEP), seguidos de cerca, en orden decreciente de vulnerabilidad, por el *cluster* C (40% de los estudiantes no son Beneficiarios SEP) y luego por el *cluster* D (50% de los estudiantes no son Beneficiarios SEP). El *cluster* E presenta nula composición de beneficiarios SEP.

Familias de estudiantes cuyos cambios son caracterizados por los *clusters* A, B, C y D tienen un Ingreso del Hogar similar (entre \$300.000 y \$400.000), en tanto que el mayor valor

para esta variable es en el *cluster* E (entre \$800.000 y \$1.000.000). Por otro lado, los padres de estudiantes cuyos cambios son caracterizados por los *clusters* A, B, C y D presentan en promedio educación hasta tercero medio, mientras que en aquellos cambios que son parte del *cluster* E, los padres poseen una educación promedio de titulados de CFT (centro de formación técnica) o IP (instituto profesional). Los mayores puntajes SIMCE lo obtienen los estudiantes del *cluster* E (cerca de 280 puntos), seguidos de los *clusters* C y D (cerca de 240 puntos), y terminando con los estudiantes de los *clusters* A y B (alrededor de 220 puntos). Además, los estudiantes de los *clusters* C, D y E recorren mayores distancias (Distancia Hogar Colegio, alrededor de 3 km), seguidos por los estudiantes de los *clusters* A y B (2 km). Todos los *clusters* exhiben valores promedios similares para las variables de género (% Masculino) y notas (Percentil Promedio de Notas), con valores en torno al 50% y 60% respectivamente.

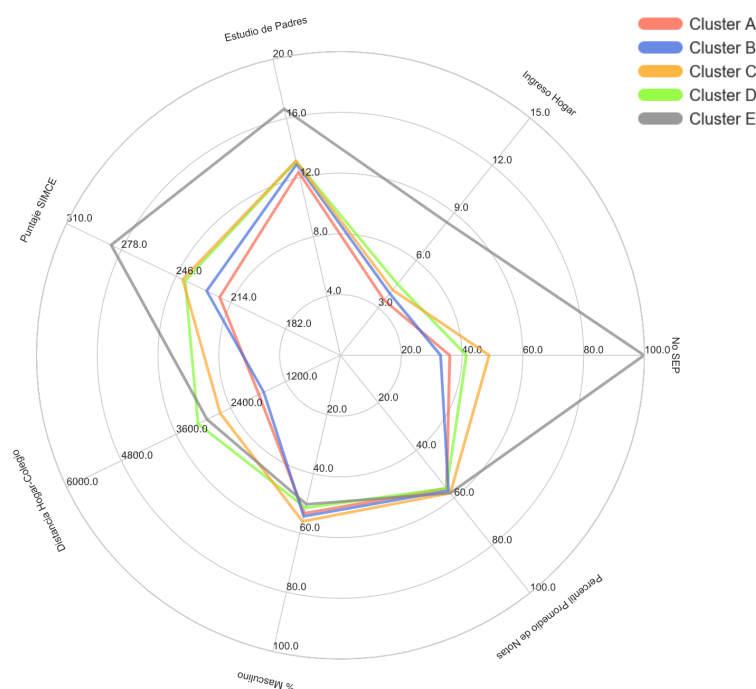


Figura 5.3: Variables a nivel de estudiante.

En este escenario se pueden destacar las siguientes características:

- El *cluster* E lo forman estudiantes de GSE alto: sobresalen las variables Puntaje SIMCE, Educación de los Padres, Ingreso del Hogar y No Vulnerabilidad. Además, el mismo *cluster* representa movilidad entre EE particulares subvencionados, con copago alto, baja vulnerabilidad y altos resultados SIMCE.
- Los *clusters* A y B son los de mayor vulnerabilidad: menor Puntaje SIMCE e Ingreso del Hogar, con más Beneficiarios SEP; asisten a establecimientos de bajo Copago, con alta Vulnerabilidad y con resultados SIMCE muy bajos. Estos son, además, los *clusters* con establecimientos de origen de dependencia municipal.

5.1.1.2. Escenario SAE

SAE Preferencias

En este caso hay cuatro *clusters*. A partir de la Figura 5.4 (b), se observa que estos quedan muy bien definidos por la relación entre dependencias administrativas de origen y destino, a excepción del *cluster* D, que incluye un pequeño flujo desde y hacia establecimientos particulares pagados. De las Figuras 5.4 (b) y (c), se observa que los *clusters* A y D tienen la mayor representación porcentual de estudiantes del escenario (35 % y 36 % respectivamente), mientras que los *clusters* B y C tienen la menor (15 % y 14 % respectivamente).

Cluster	Copago rbd origen	Copago rbd preferencia	IVE rbd origen	IVE rbd preferencia	Promedio SIMCE rbd origen	Promedio SIMCE rbd preferencia	Dependencia rbd origen	Dependencia rbd preferencia
A	P65 (Gratuito)	P65 (Gratuito)	0,69	0,77	P17	P8	Municipal	Municipal
B	P65 (Gratuito)	P69	0,7	0,65	P21	P46	Municipal	Part. Subv.
C	P65 (Gratuito)	P65 (Gratuito)	0,75	0,77	P52	P56	Part. Subv.	Municipal
D	P67	P69	0,73	0,70	P48	P52	Part. Subv.	Part. Subv.

Tabla 5.4: Centroides de *clusters*.

La Tabla 5.4 y la Figura 5.4 (a) muestra los centroides de cada *cluster*. Se observa que el Copago no varía en los *clusters* A, C, y D, mientras que la movilidad representada por el *cluster* B aumenta en términos de copago, lo que es consistente con el tipo de cambio que representa (desde municipal hacia particular subvencionado). La vulnerabilidad (IVE) disminuye en los *clusters* B y D, y aumenta en los *clusters* A y C, siendo leve el efecto en este último (*cluster* C). En términos de resultados SIMCE de los establecimientos de origen y destino, sólo en el *cluster* A se observa una baja en rendimiento, mientras que en los *clusters* B, C y D hay un aumento de puntaje, este efecto es elevado para el *cluster* B (cambio desde municipal hacia particular subvencionado).

Al comparar las variables a nivel de estudiante, en la Figura 5.5 se observa, para la variable de vulnerabilidad (No participa en la SEP), que los *clusters* A y B son los "menos no vulnerable" (cerca del 20 % de los estudiantes no son Beneficiarios SEP), es decir, son los más vulnerables, seguidos de cerca, en orden decreciente de vulnerabilidad, por los *clusters* C y D (35 % de los estudiantes no son Beneficiarios SEP). Por otro lado, los hogares de estudiantes de todos los *clusters* perciben un ingreso similar entre \$300.000 y \$500.000. Además, los padres de estudiantes de todos los *clusters* presentan en promedio educación hasta tercero medio (valor 12 variable Educación de los Padres).

Los mayores puntajes SIMCE lo obtienen los estudiantes de los *clusters* C y D (cerca de 240 puntos), seguidos de los *clusters* A y B (cerca de 220 puntos). Los estudiantes de todos los *clusters* recorren distancias entre 1,5 y 2 km (Distancia Hogar Colegio). Por otro lado, se constatan diferencias significativas en la variable de género (Porcentaje de Hombre): el 65 % de los estudiantes del *cluster* C son hombres, los *clusters* A y B están compuestos de un 55 % de hombres, y el *cluster* D por un 48 %. Finalmente, no se observan diferencias en el rendimiento académico de los estudiantes (Percentil Promedio de Notas), estando los estudiantes todos los *clusters* alrededor del percentil 58 en promedio.

En este escenario se pueden destacar las siguientes características:

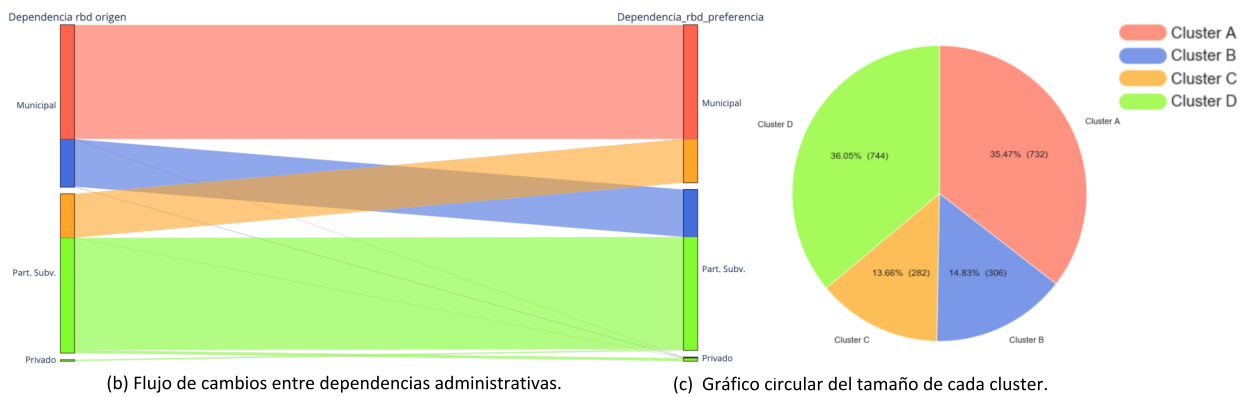
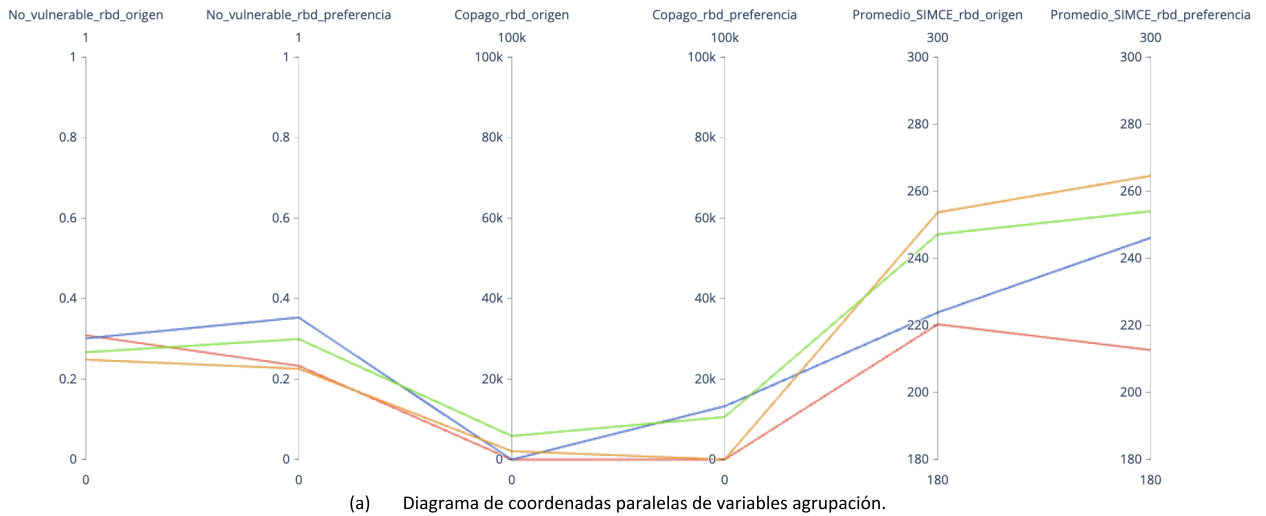


Figura 5.4: Variables a nivel de establecimiento.

- Hay similar preferencia por escuelas subvencionadas (*clusters* B y D) que por municipales (*clusters* A y C).
- La vulnerabilidad es homogénea y elevada en todos los *clusters* (IVE, Estudios Padres, Ingreso del Hogar).
- Se observa sesgo de género en el *cluster* C.

SAE Admisión

En este caso hay cuatro *clusters*. A partir de la Figura 5.6 (b), se observa que estos quedan muy bien definidos por la relación entre dependencias administrativas de origen y destino, a excepción del *cluster* D, que incluye un pequeño flujo desde y hacia establecimientos particulares pagados. En las Figuras 5.6 (b) y (c), se observa que el *cluster* D tiene la mayor representación porcentual de estudiantes del escenario (40%), mientras que los *clusters* B y C tienen la menor (13% y 10% respectivamente). En tanto, el *cluster* A comprende el 37% de los cambios.

La Tabla 5.5 y la Figura 5.6 (a) muestran los centroides de cada *cluster*. Se observa que el Copago no cambia en ninguno de los *clusters*. La vulnerabilidad (IVE) de los establecimientos

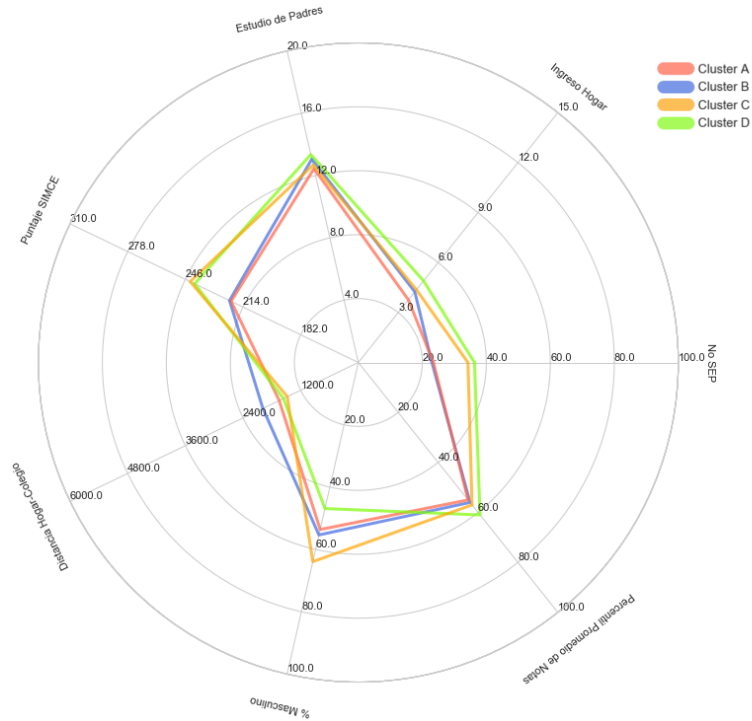


Figura 5.5: Variables a nivel de estudiante.

Cluster	Copago rbd origen	Copago rbd admitido	IVE rbd origen	IVE rbd admitido	Promedio SIMCE rbd origen	Promedio SIMCE rbd admitido	Dependencia rbd origen	Dependencia rbd admitido
A	P65 (Gratis)	P65 (Gratis)	0,69	0,76	P17	P7	Municipal	Municipal
B	P65 (Gratis)	P68	0,7	0,69	P21	P44	Municipal	Part. Subv.
C	P65 (Gratis)	P65 (Gratis)	0,72	0,76	P53	P48	Part. Subv.	Municipal
D	P66	P68	0,74	0,72	P46	P48	Part. Subv.	Part. Subv.

Tabla 5.5: Centroides de *clusters*.

de origen y destino permanece casi constante en los *clusters* B y D, mientras que en los *cluster* A y C aumenta. En términos de resultados SIMCE de los establecimientos de origen y destino, en los *clusters* A y C se observa una baja en rendimiento, mientras que en el *cluster* B hay un aumento importante de puntaje; en tanto que en el *cluster* D no se observa diferencia en rendimiento medido por el SIMCE.

Al comparar las variables a nivel de estudiante, en la Figura 5.7 se observa, para la variable de vulnerabilidad (No participa en la SEP), que los estudiantes que conforman los *clusters* A y B son los "menos no vulnerable" (en torno al 20 % de los estudiantes no son Beneficiarios SEP), es decir, son los más vulnerables, seguidos, en orden decreciente de vulnerabilidad, por los *clusters* C y D (40 % de los estudiantes no son Beneficiarios SEP). Asimismo, los hogares de estudiantes de todos los *cluster* perciben un ingreso similar entre \$400.000 y \$500.000. los padres de estudiantes de todos los *clusters* presentan en promedio educación hasta tercero medio (valor 12 variable Educación de los Padres).

Los mayores puntajes SIMCE los obtienen (en promedio) estudiantes de los *clusters* C y D (cerca de 246 puntos), seguidos de los *clusters* A y B (cerca de 220 puntos). Por otro lado, los estudiantes de los *clusters* A, C y D recorren distancias promedio de 1,6 km (Distancia Hogar Colegio), mientras que los estudiantes del *cluster* B lo hacen 2,2 km. en promedio.

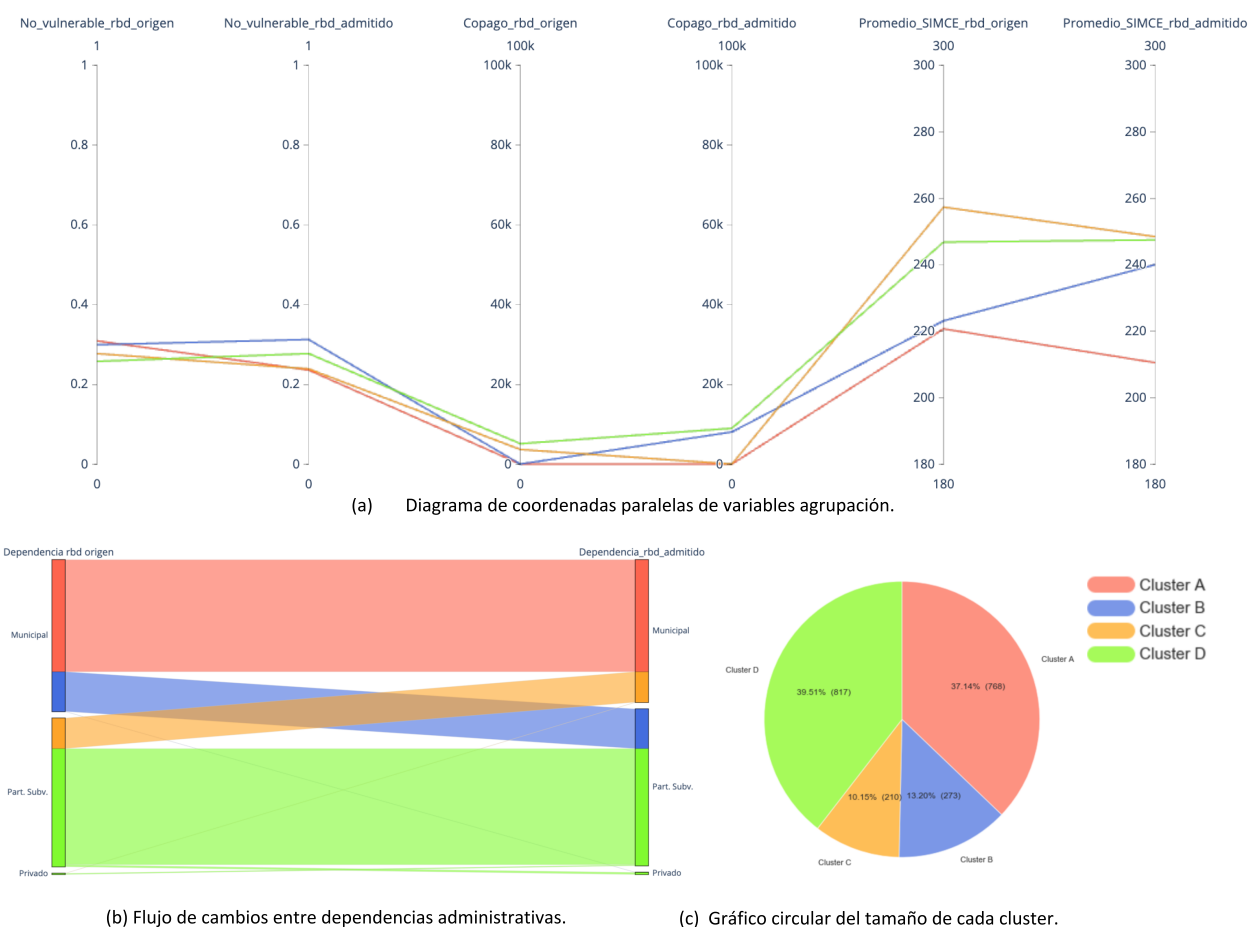


Figura 5.6: Variables a nivel de establecimiento.

Además, no se constatan diferencias significativas en la variable de género (Porcentaje de Hombres), teniendo todos los *clusters* una composición cercana a 55% de hombres. Tampoco se observan diferencias en el rendimiento académico de los estudiantes (Percentil Promedio de Notas), estando todos los *clusters* alrededor del percentil 58 en promedio.

En este escenario se pueden destacar las siguientes características:

- Hay similar admisión en escuelas subvencionadas (*clusters* B y D) que en municipales (*clusters* A y D).
- La vulnerabilidad es homogénea y elevada en todos los *clusters*, evidenciada por las variables IVE, Estudios Padres e Ingreso del Hogar.

Patrones de Cambio Forzado en Iquique - Alto Hospicio

Se observan patrones similares en los escenarios SAE Preferencias y SAE Admisión en términos de representatividad de los *clusters*. Sin embargo, las preferencias apuntan a establecimientos con mejor rendimiento, medido por el SIMCE que aquellos en lo que son asignados, es decir, no existe una oferta educativa que satisfaga la demanda en términos de resultados SIMCE.

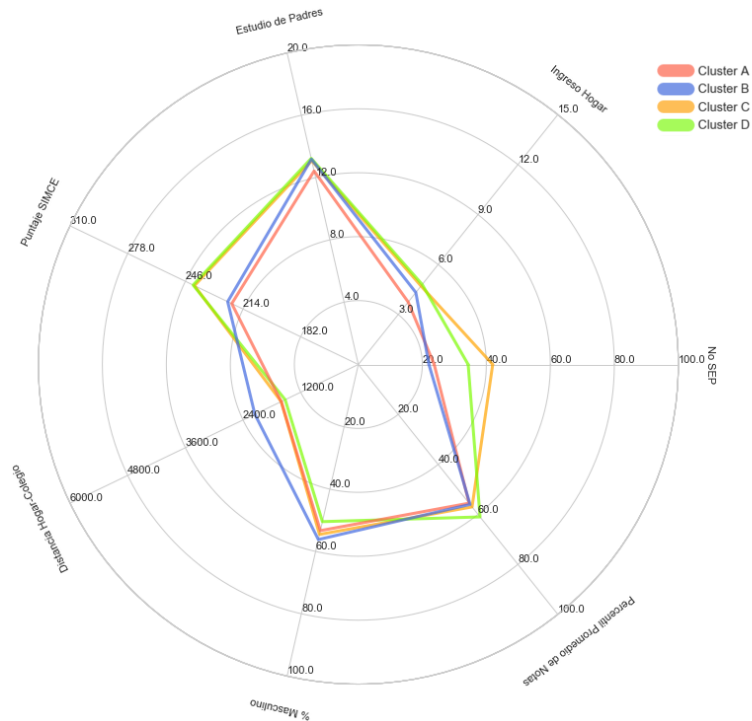


Figura 5.7: Variables a nivel de estudiante.

Al comparar el escenario Pre SAE con aquellos post SAE, la agrupación de cambios de estudiantes de GSE alto (*cluster E*), no es lo suficientemente fuerte para ser considerada como un *cluster* aparte en los escenarios SAE Preferencia y SAE admisión. Este perfil de cambio es integrado por el *cluster D*, lo que se evidencia en el aumento de su representatividad.

Siguiendo con la comparación del escenario previo al SAE con aquellos que ya lo implementan, se observa que los puntajes SIMCE de los establecimientos de destino son similares (*cluster A*) o mejoran (*clusters C, D y E*), tanto para el caso de Preferencias como en el de Admisión.

Resalta finalmente la desaparición del sesgo de género observado en el escenario SAE Preferencias (*clusters C y D*), al compararlo con la Admisión. Aunque siguen siendo considerablemente más hombres que mujeres los que cambian de escuela.

5.1.2. Cambios No Forzados

5.1.2.1. Escenario Pre SAE

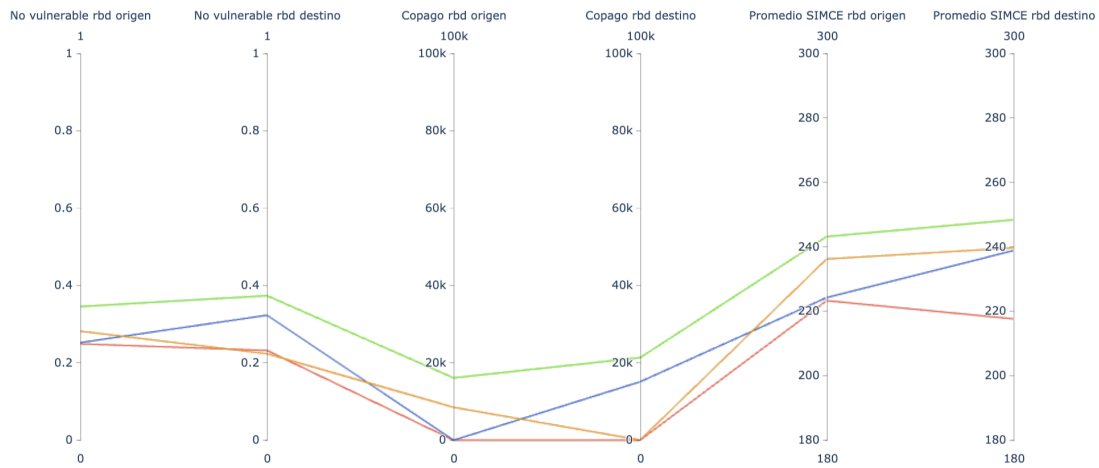
En este caso hay cuatro *clusters*. A partir de la Figura 5.8 (b), se observa que estos quedan bien definidos por la relación entre dependencias administrativas de origen y destino, a excepción de *cluster D*, que incluye el flujo desde y hacia establecimientos particulares pagados. En las Figuras 5.8 (b) y (c), se observa que el *cluster D* tiene la mayor representación porcentual de estudiantes del escenario (54%), mientras que los *cluster A y B* tienen la menor

(12% y 10% respectivamente). EL *cluster* C comprende el 24% de los cambios.

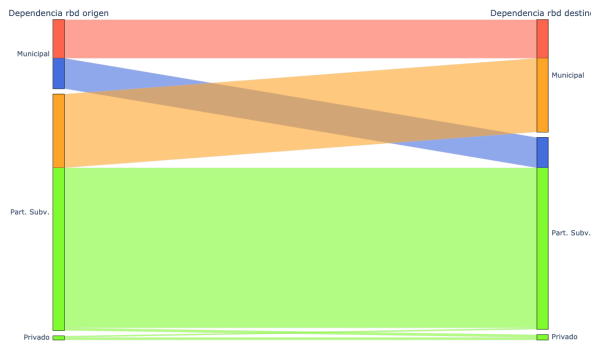
Cluster	Copago rbd origen	Copago rbd destino	IVE rbd origen	IVE rbd destino	Promedio SIMCE rbd origen	Promedio SIMCE rbd destino	Dependencia rbd origen	Dependencia rbd destino
A	P60 (Gratis)	P61 (Gratis)	0,75	0,77	P18	P10	Municipal	Municipal
B	P60 (Gratis)	P74	0,74	0,68	P19	P40	Municipal	Part. Subv.
C	P68	P61 (Gratis)	0,72	0,77	P38	P41	Part. Subv.	Municipal
D	P75	P78	0,65	0,63	P46	P52	Part. Subv.	Part. Subv.

Tabla 5.6: Centroides de *clusters*.

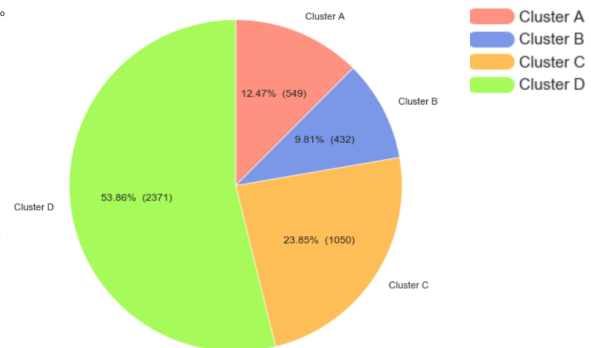
La Tabla 5.6 y la Figura 5.8 (a) muestra los centroides de cada *cluster*. Se observa que el Copago no cambia en el *cluster* A. La movilidad representada por los *clusters* B y D aumentan en términos de copago, mientras que el *cluster* C es el único que disminuye en esta variable. La vulnerabilidad (IVE) aumenta en los *clusters* A y C, mientras que en los *clusters* B y D hay una disminución, siendo muy leve la variación en este último (*cluster* D). Por otro lado, en términos de resultados SIMCE de los establecimientos de origen y destino, sólo el *cluster* C presenta valores constantes, en tanto que en el *cluster* A hay un cambio desde rendimiento bajo hacia un rendimiento aún menor, y al contrario, en los *clusters* B y D hay una mejora en rendimiento, medido por el SIMCE, siendo más pronunciado en el primero.



(a) Diagrama de coordenadas paralelas de variables agrupación.



(b) Flujo de cambios entre dependencias administrativas.



(c) Gráfico circular del tamaño de cada cluster.

Figura 5.8: Variables a nivel de establecimiento.

Al comparar las variables a nivel de estudiante, en la Figura 5.9 se observa, para la variable de vulnerabilidad (No participa en la SEP), que los estudiantes de los *clusters* A y B son

los "menos no vulnerable" (en torno al 35% de los estudiantes no son Beneficiarios SEP), es decir, son los más vulnerables, seguidos, en orden decreciente de vulnerabilidad, por los estudiantes de los *clusters* C (50% de los estudiantes no son Beneficiarios SEP) y D (60% de los estudiantes no son Beneficiarios SEP). También hay diferencias en el Ingreso del Hogar: hogares de estudiantes del *cluster* D reciben mayores ingresos (entre \$500.000 y \$600.000), seguido por los otros tres *clusters*, que reciben entre \$300.000 y \$500.000 en promedio. Con respecto a la variable Educación de los Padres, los padres de estudiantes de todos los *clusters* promedian educación media completa.

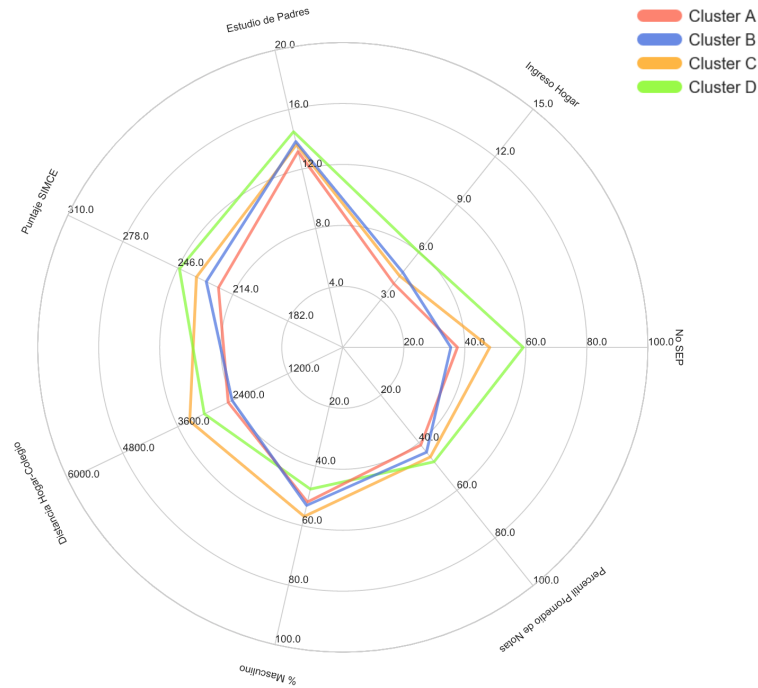


Figura 5.9: Variables a nivel de estudiante.

Los estudiantes de todos los *clusters* obtienen puntajes SIMCE entre 200 y 246 en promedio, presentando leves diferencias: los mayores puntajes en promedio los obtienen los estudiantes del *cluster* D, seguidos por aquellos del *cluster* C, luego por los del *cluster* B, y finalmente por los estudiantes del *cluster* A. Asimismo, estudiantes de los *clusters* C y D recorren mayores distancias (Distancia Hogar Colegio), de cerca de 3,5 km en promedio, mientras que los estudiantes de los *clusters* A y B recorren en torno a los 2,4 km.

No se constatan diferencias significativas en la variable de género (Porcentaje de Hombres), teniendo todos los *clusters* una composición entre el 48 y 58% de hombres. Tampoco se observan diferencias en el rendimiento académico de los estudiantes (Percentil Promedio de Notas), estando estudiantes de todos los *clusters* (en promedio) entre los percentiles 40 y 50.

En este escenario se pueden destacar las siguientes características:

- La mayoría de los cambios ocurren hacia escuelas subvencionadas (*clusters* B y D), mientras que la minoría ocurre hacia municipales (*clusters* A y C).
- La Vulnerabilidad del *cluster* D es levemente menor: menor IVE, mayor Ingreso del Hogar, menos Beneficiarios SEP y mayor Copago.

- Estudiantes que se cambian desde establecimientos particulares subvencionados (*clusters* C y D) recorren mayores distancias.

5.1.2.2. Escenario SAE

SAE Preferencias

En este caso hay cuatro *clusters*. A partir de la Figura 5.10 (b), se observa que estos quedan parcialmente definidos por la relación entre dependencias administrativas de origen y destino: el *cluster* A representa la movilidad de estudiantes desde establecimientos municipales, el *cluster* B aquella entre municipales y particulares subvencionados, el *cluster* C comprende parte de los cambios entre particulares subvencionados, y finalmente en *cluster* D, además de contener cambios entre particulares subvencionados, incluye los cambios que involucren EE particulares pagados. En las Figuras 5.10 (b) y (c), se observa que el *cluster* C tiene la mayor representación porcentual de estudiantes del escenario (40%), mientras que los *clusters* A, B y D comprenden el 19%, 22% y 19% de los cambios respectivamente.

Cluster	Copago rbd origen	Copago rbd preferencia	IVE rbd origen	IVE rbd preferencia	Promedio SIMCE rbd origen	Promedio SIMCE rbd preferencia	Dependencia rbd origen	Dependencia rbd preferencia
A	P65 (Gratuito)	P68	0,73	0,69	P24	P44	Municipal	Muni./Subv.
B	P65 (Gratuito)	P65 (Gratuito)	0,75	0,78	P42	P56	Part. Subv.	Municipal
C	P65 (Gratuito)	P67	0,74	0,71	P42	P48	Part. Subv.	Part. Subv.
D	P77	P82	0,46	0,36	P56	P66	Subv./Priv.	Subv./Priv.

Tabla 5.7: Centroides de *clusters*.

La Tabla 5.7 y la Figura 5.10 (a) muestran los centroides de cada *cluster*. Se observa que el Copago aumenta en los *clusters* A y D, disminuye en el *cluster* B y se mantiene en el *cluster* C. La vulnerabilidad (IVE) disminuye en los *clusters* A, C y D, mientras que aumenta solo en el *cluster* B. En términos de resultados SIMCE, en todos los *clusters* hay un aumento en puntajes, siendo en el *cluster* A el efecto más pronunciado (en términos de percentiles).

Al comparar las variables a nivel de estudiante, en la Figura 5.11 se observa, para la variable de vulnerabilidad (No participa en la SEP), que el *cluster* A es el "menos no vulnerable" (en torno al 15% de los estudiantes no son Beneficiarios SEP), es decir, es el más vulnerable, seguido, en orden decreciente de vulnerabilidad, por los *clusters* B y C (en torno al 30% de los estudiantes no son Beneficiarios SEP), mientras que el *cluster* D lo componen mayoritariamente estudiantes no vulnerables (85% de los estudiantes no son Beneficiarios SEP).

Hay leves diferencias en el Ingreso del Hogar de estudiantes de los *clusters* A, B y C, los que reciben (en promedio) entre \$300.000 y \$500.000. Hogares de estudiantes del *cluster* D reciben mayores ingresos, promediando entre \$600.000 y \$800.000. Por otro lado, con respecto a la variable Educación de los Padres, aquellos cuyos hijos forman los *clusters* A, B y C presentan en promedio educación media completa (valores 13 y 14), mientras que los padres de estudiantes del *cluster* D poseen son en promedio graduados de IP o CFT.

Los estudiantes de los *clusters* A, B y C obtienen puntajes SIMCE entre 230 y 240 puntos en promedio, presentando leves diferencias: los mayores puntajes son de estudiantes de los

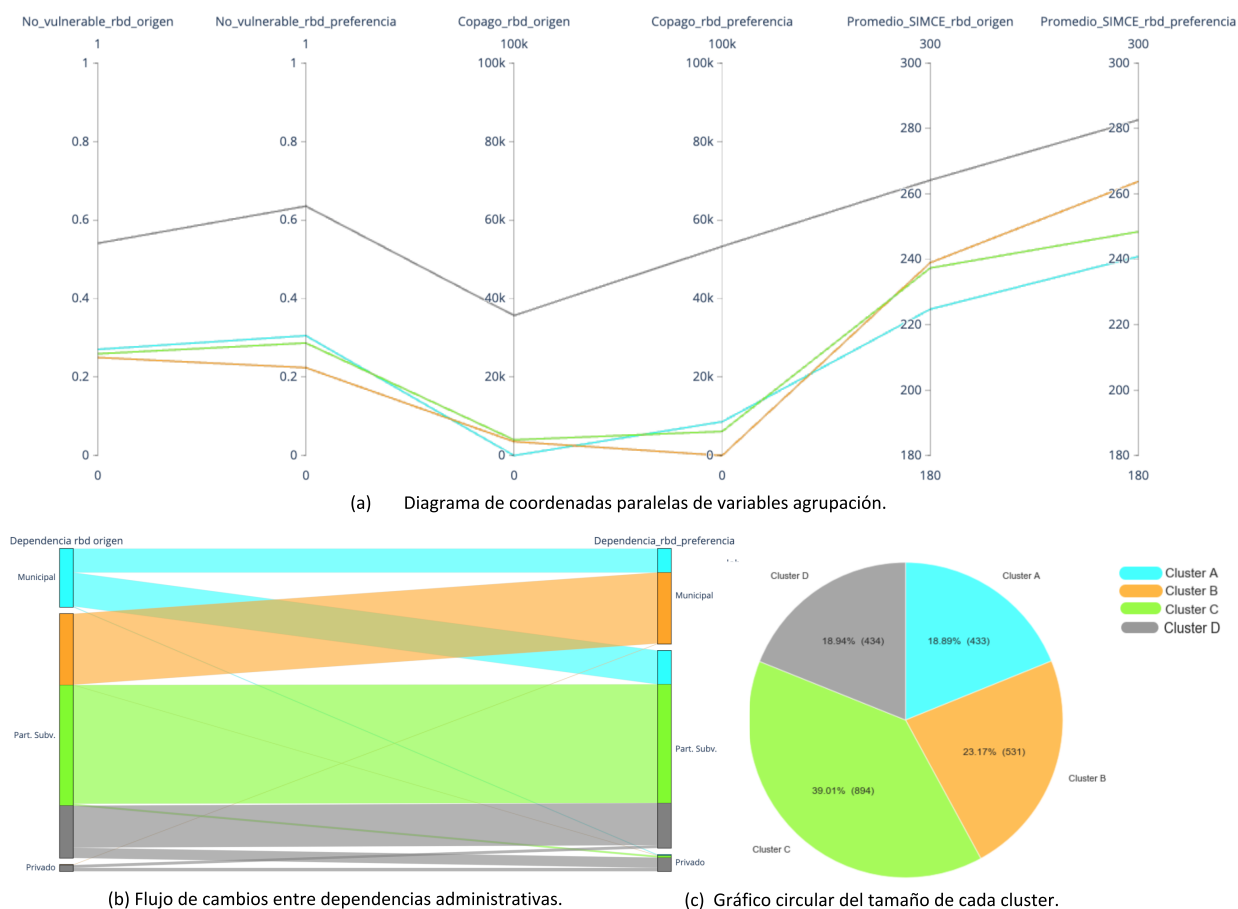


Figura 5.10: Variables a nivel de establecimiento.

clusters B y C, seguidos del *cluster* A. En tanto, los estudiantes del *cluster* D tienen mayores puntajes SIMCE, de 260 puntos en promedio. Por otro lado, los estudiantes de los *clusters* A y D recorren mayores distancias (Distancia Hogar Colegio), cerca de 3 km. Estudiantes del *cluster* B recorren en torno a los 2,3 km y los del *cluster* B, 2 km.

No se constatan diferencias significativas en la variable de género (Porcentaje de Hombres), teniendo los *clusters* A, B, C y D una composición de hombres de 55 %, 50 %, 45 % y 45 % respectivamente. Al contrario, si se observan diferencias en el rendimiento académico de los estudiantes (Percentil Promedio de Notas): estudiantes del *cluster* A tienen bajo rendimientos (percentil 32), seguidos del *cluster* C (percentil 40) y luego por el D (45 %); finalmente, estudiantes del *cluster* B tienen el mayor rendimiento, ubicándose sus alumnos en promedio en el percentil 50.

En este escenario se pueden destacar las siguientes características:

- Los *clusters* ya no están totalmente ordenados de acuerdo transición de dependencia administrativa.
- La mayoría de los cambios ocurren entre escuelas subvencionadas (*clusters* C y D).
- La vulnerabilidad del *cluster* D es menor: menor IVE, menos Beneficiarios SEP, mayor Copago y mayor rendimiento medido por el SIMCE.

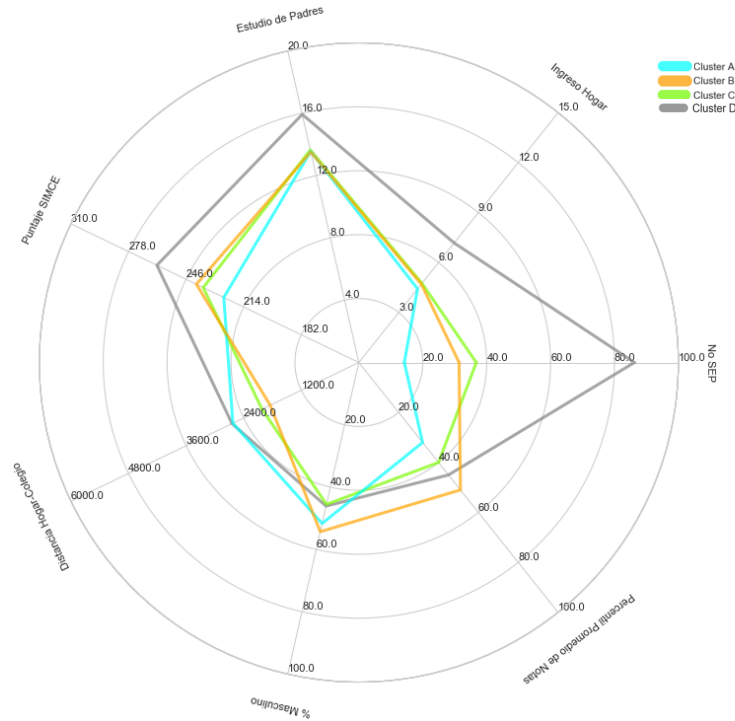


Figura 5.11: Variables a nivel de estudiante.

- Estudiantes que cambia desde EE municipales (*cluster A*) tienen bajo rendimiento académico y menores puntajes SIMCE.

SAE Admisión

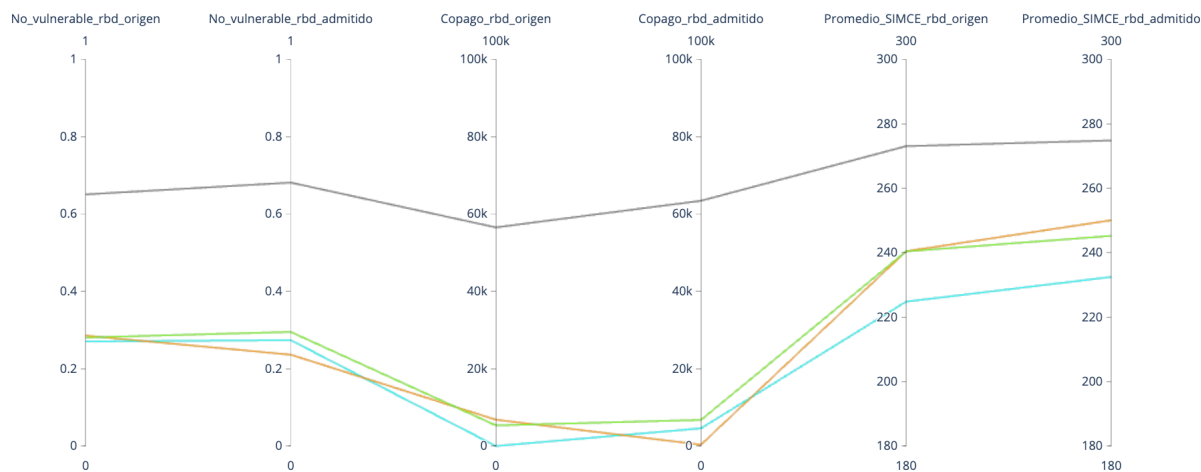
Este escenario hay 4 *clusters*. A partir de la Figura 5.12 (b), se observa que estos quedan parcialmente definidos por la relación entre dependencias administrativas de origen y destino: el *cluster A* representa la movilidad de estudiantes desde establecimientos municipales, el *cluster B* aquella entre municipales y particulares subvencionados, el *cluster C* comprende los cambios entre particulares subvencionados, y el *cluster D* comprende cambios entre particulares subvencionados y además a aquellos que involucran establecimientos particulares pagados. De las Figuras 5.12 (b) y (c), se observa que el *cluster C* tiene la mayor representación porcentual de estudiantes del escenario (54%), mientras que los *cluster A*, B y D comprenden el 19%, 18% y 9% los cambios respectivamente.

Cluster	Copago rbd origen	Copago rbd admitido	IVE rbd origen	IVE rbd admitido	Promedio SIMCE rbd origen	Promedio SIMCE rbd admitido	Dependencia rbd origen	Dependencia rbd admitido
A	P65 (Gratuito)	P66	0,73	0,73	P24	P37	Municipal	Muni./Subv.
B	P67	P65 (Gratuito)	0,71	0,76	P44	P50	Part. Subv.	Municipal
C	P66	P67	0,72	0,7	P44	P45	Part. Subv.	Part. Subv.
D	P82	P83	0,35	0,32	P60	P62	Subv./Priv.	Subv./Priv.

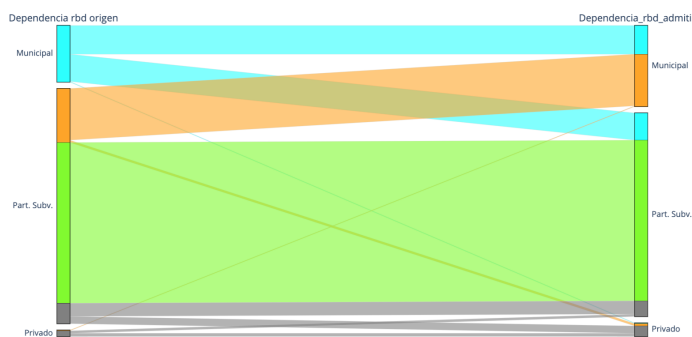
Tabla 5.8: Centroides de *clusters*.

La Tabla 5.8 y la Figura 5.12 (a) muestran los centroides de cada *cluster*. Se observa que el Copago aumenta en todos los *clusters* excepto en el B, siendo estudiantes del *cluster D* los

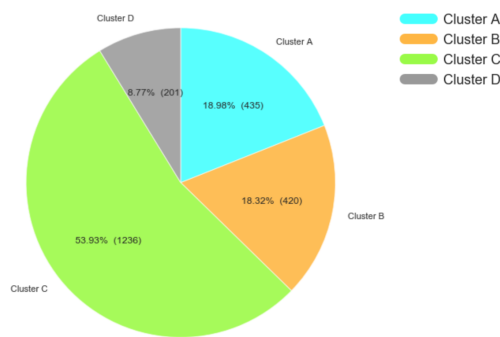
con mayor valor en este atributo (de \$55.000 a \$65.000). La vulnerabilidad (IVE) aumenta únicamente en el *cluster* B, disminuye en los *clusters* C y D, mientras que se mantiene constante en el *cluster* A. En términos de resultados SIMCE, en todos los *clusters* se observa una mejora en rendimiento, siendo en el primero (*cluster* A) el efecto de mayor magnitud.



(a) Diagrama de coordenadas paralelas de variables agrupación.



(b) Flujo de cambios entre dependencias administrativas.



(c) Gráfico circular del tamaño de cada cluster.

Figura 5.12: Variables a nivel de establecimiento.

Al comparar las variables a nivel de estudiante, en la Figura 5.13 se observa, para la variable de vulnerabilidad (No participa en la SEP), que el *cluster* A es el "menos no vulnerable" (en torno al 10% de los estudiantes no son Beneficiarios SEP), es decir, es el más vulnerable, seguido, en orden decreciente de vulnerabilidad, por los *clusters* B y C (40% de los estudiantes no son Beneficiarios SEP), mientras que el *cluster* D tiene una muy baja vulnerabilidad (90% de los estudiantes no son Beneficiarios SEP).

Con respecto al Ingreso del Hogar, los *cluster* A, B y C representan ingresos entre \$300.000 y \$500.000, mientras que familias del *cluster* D perciben un ingreso mayor, entre \$800.000 y \$1.000.000 en promedio. Por otro lado, para la variable Educación de los Padres, los *clusters* A, B y C representan en promedio educación media completa, en tanto que los padres de estudiantes del *cluster* D son en promedio titulados de CFT o IP.

Los estudiantes del *cluster* A obtienen los menores puntajes SIMCE (225 puntos en promedio), seguido por estudiantes de los *clusters* B y C, que obtienen 240 puntos en promedio; mientras que aquellos del *cluster* E promedian 260 puntos. Por otra parte, los estudiantes de

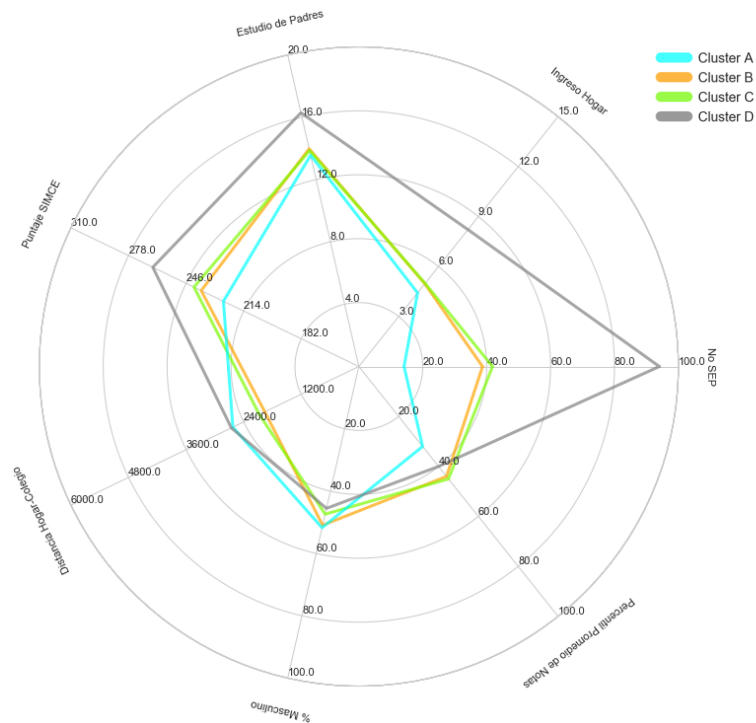


Figura 5.13: Variables a nivel de estudiante.

los *clusters* A y D recorren mayores distancias (Distancia Hogar Colegio), cercana a 3 km, en tanto que los estudiantes de los *clusters* B y C recorren en torno a los 2,2 km.

No se constatan diferencias significativas en la variable de género (Porcentaje de Hombres), teniendo todos los *clusters* una composición de hombres entre 45 % y 55 %. Por otro lado, se observan diferencias en rendimiento académico (Percentil Promedio de Notas): estudiantes del *cluster* A tienen menor rendimientos (percentil 35), mientras que aquellos de los *clusters* B, C y D tienen rendimiento promedio un poco mayor (percentil 45).

En este escenario se pueden destacar las siguientes características:

- Los *clusters* ya no están totalmente ordenados de acuerdo a la transición de dependencia administrativa, lo que sugiere un mayor desorden en la organización de la movilidad en el escenario no forzado.
- La mayoría de los cambios ocurren entre escuelas subvencionadas (*cluster* D) y la minoría entre municipales (*cluster* A).
- La vulnerabilidad del *cluster* D es menor: menor IVE, menos Beneficiarios SEP, mayor Copago, mayor rendimiento medido por el SIMCE, mayor Ingreso del Hogar y mayor Educación de Padres.
- Estudiantes que cambian desde EE municipales (*clusters* A) tienen bajo rendimiento académico.

Patrones de Cambio No Forzado en Iquique - Alto Hospicio Al analizar el escenario completo de cambios No Forzados en esta conurbación, se observa que en los escenarios SAE aparece un *cluster* que representa cambios de estudiantes de GSE alto (*cluster* D), mientras

que aquellos cambios desde establecimientos municipales forman una única agrupación. Esto responde a la similitud en las características de los cambios de escuela, y a como estas varían según escenario. Además, en los escenarios SAE se observa una mayor movilidad hacia establecimientos particulares pagados y una disminución de cambios hacia escuelas municipales.

En relación al rendimiento medido por el SIMCE, se destaca, que en el escenario Pre SAE, sólo *cluster* el A no sube en puntaje SIMCE en la transición hacia establecimiento de destino. En tanto que para los escenarios SAE, los *clusters* A, B y D demandan EE con mayores puntajes (SAE preferencias) que aquellos a los que son asignados. Es decir, hay una demanda por establecimientos de mejor rendimiento, medido por el SIMCE, que no está siendo cubierta. Sin embargo, todos los *clusters* suben en puntaje (SAE admisión) con respecto al escenario Pre SAE.

Finalmente, se observa que los *clusters* A y D presenta menor estructura (coeficiente de silueta) en casos SAE, lo que es consecuencia de la mezcla en dependencias administrativas de origen y destino, lo que marca una diferencia en dos de los ocho atributos utilizados en el proceso de agrupamiento, disminuyendo el valor del coeficiente. No obstante, estos perfiles de cambio siguen siendo lo suficientemente similares para ser considerados como una agrupación en sí.

Patrones de Cambio en Iquique - Alto Hospicio

Al contrastar, en esta aglomeración urbana, los tipos de cambio Forzado y No Forzado, se destaca, que en el segundo la proporción de cambios desde la dependencia municipal es mucho menor, esto pues hay muchos establecimientos municipales que entregan educación sólo hasta los niveles estudiados, y por ende, caen en la categoría de cambios Forzados.

Para los casos SAE, en ambos tipos de cambio hay una demanda por resultados SIMCE que no es completamente cubierta. Sin embargo, la Admisión SAE nuestra, que en promedio, los puntajes de establecimientos de destino son superiores a aquellos que obtenían los establecimientos que recibían estudiantes previo al SAE.

Además, la calidad de los *clusters* en cambios Forzados es significativamente mayor (Coeficiente de Silueta) que en el caso No Forzado, lo que indica que estos poseen patrones más definidos y estructurados.

5.2. Coquimbo - La Serena

5.2.1. Cambios Forzados

5.2.1.1. Escenario Pre SAE

En este escenario hay 4 *clusters*. A partir de la Figura 5.14 (b), se observa que estos quedan totalmente definidos por la relación entre dependencias administrativas de origen y destino: el *cluster* A representa la movilidad de estudiantes en establecimientos municipales, el *cluster* B aquella entre municipales y particulares subvencionados, el *cluster* C comprende los cambios entre particulares subvencionados y municipales, y el *cluster* D agrupa cambios entre particulares subvencionados, además de aquellos que involucran establecimientos particulares pagados (muy pocos). De las Figuras 5.14 (b) y (c), se observa que el *cluster* A tiene la mayor representación porcentual de estudiantes del escenario (35%), mientras que los *clusters* B, C y E comprenden el 21%, 13% y 31% de los cambios respectivamente.

Cluster	Copago rbd origen	Copago rbd destino	IVE rbd origen	IVE rbd destino	Promedio SIMCE rbd origen	Promedio SIMCE rbd destino	Dependencia rbd origen	Dependencia rbd destino
A	P60 (Gratuito)	P63	0,8	0,79	P22	P27	Municipal	Municipal
B	P60 (Gratuito)	P73	0,78	0,68	P25	P48	Municipal	Part. Subv.
C	P64	P64	0,73	0,77	P43	P39	Part. Subv.	Municipal
D	P66	P73	0,75	0,71	P47	P52	Part. Subv.	Part. Subv.

Tabla 5.9: Centroides de *clusters*.

La Tabla 5.9 y la Figura 5.14 (a) muestran los centroides de cada *cluster*. Se observa que el Copago para establecimientos de origen y destino es similar en los *clusters* A y C, en tanto, este aumenta en los *clusters* B y D. La vulnerabilidad (IVE) disminuye en los *clusters* B y D, aumenta en el *cluster* C y se mantiene constante en el *cluster* A. En términos de resultados SIMCE, en los *clusters* A, B y D hay una mejora en rendimiento, medido por SIMCE, siendo en el segundo (*cluster* B) muy pronunciado el aumento, en tanto, los cambios agrupados en el *cluster* C llegan a establecimientos con menor puntaje SIMCE que en su EE de origen.

Al comparar las variables a nivel de estudiante, en la Figura 5.15 se observa, para la variable de vulnerabilidad (No participa en la SEP), que el *cluster* A es el "menos no vulnerable" (en torno al 32% de los estudiantes no son Beneficiarios SEP), es decir, es el más vulnerable, seguido, en sentido decreciente en vulnerabilidad, por los *clusters* B, C y D (en torno al 45% de los estudiantes no son Beneficiarios SEP). Con respecto al Ingreso del Hogar, todos los *clusters* se caracterizan por ingresos promedio entre \$200.000 y \$400.000. Para la variable Educación de los Padres, los padres de estudiantes de todos los *clusters* presentan en promedio educación entre tercero medio y media completa.

Los estudiantes de todos los *clusters* obtienen puntajes SIMCE entre los 225 y los 246 puntos en promedio, siendo A, B, C y D su orden en orden creciente. Los estudiantes del *cluster* A recorren en promedio mayores distancias (Distancia Hogar Colegio), cercana a 3,3 km. Estudiantes de los *clusters* B, C, y D recorren en torno a los 2,3 km. No se constatan diferencias significativas en la variable de género (Porcentaje de Hombres), teniendo todos los *clusters* una composición de hombres entre 45% y 55%. Por otro lado, se observan pocas

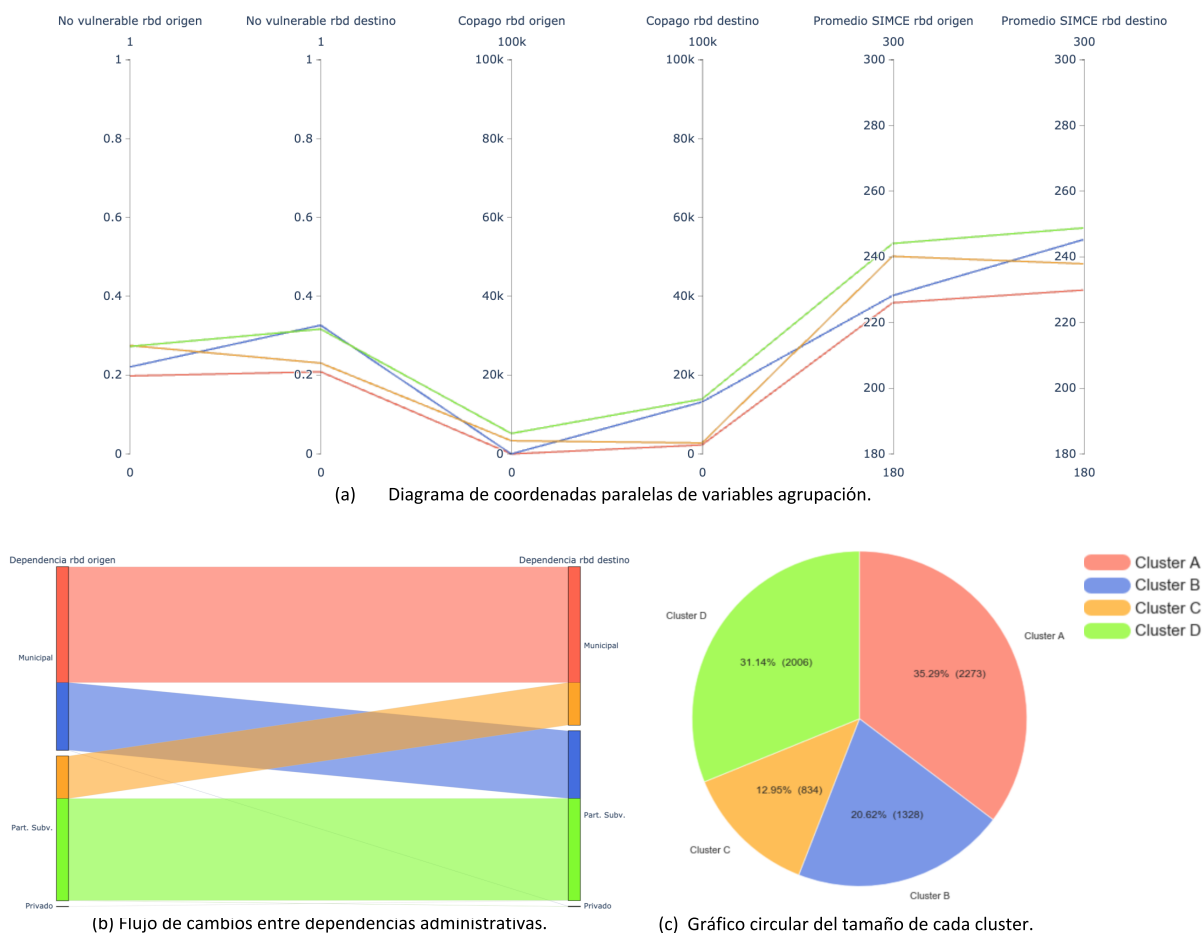


Figura 5.14: Variables a nivel de establecimiento.

diferencias en el rendimiento académico de los estudiantes (Percentil Promedio de Notas): aquellos de los *clusters* A y C tienen menor rendimiento (percentil 53), mientras que los de los *clusters* B y D tienen un mayor rendimiento en promedio (percentil 60).

En este escenario se pueden destacar las siguientes características:

- Ningún *cluster* sobresale fuertemente en términos de variables de establecimientos de origen y destino, o según características de los estudiantes que los componen.
- Hay un flujo similar de cambios entre EE municipales y entre EE particulares subvencionados.
- Estudiantes que se cambian entre establecimientos municipales (*cluster* A) recorren mayores distancias en promedio.
- Estudiantes que se cambian entre EE municipales tienen menor rendimiento académico, menor ingreso de hogar, padres con menos estudios, menores puntajes SIMCE y recorren mayores distancias. Aunque estas diferencias son leves.
- Hogares de estudiantes de todos los *cluster* perciben ingresos bajos.

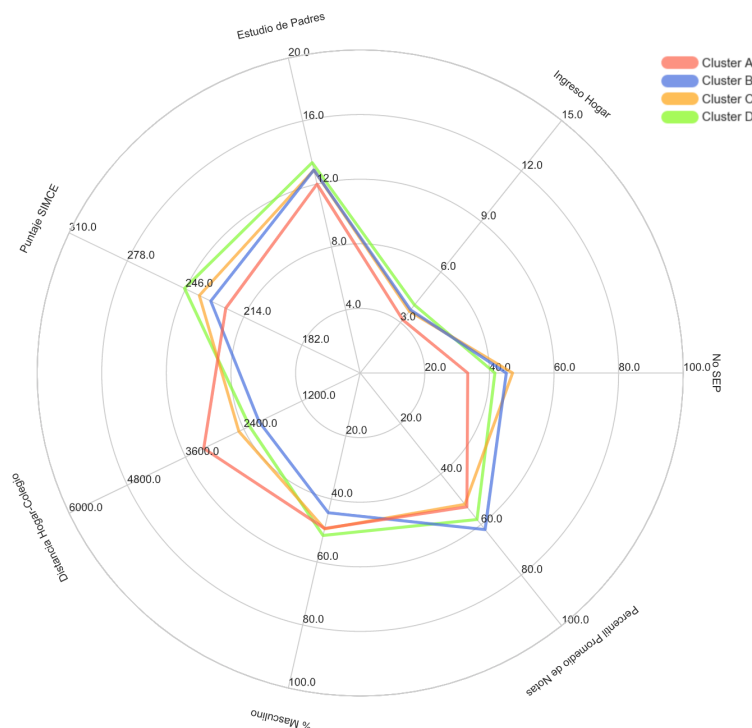


Figura 5.15: Variables a nivel de estudiante.

5.2.1.2. Escenario SAE

SAE Preferencias

En este escenario hay 4 *clusters*. A partir de la Figura 5.16 (b), se observa que estos quedan en su mayoría definidos por la relación entre dependencias administrativas de origen y destino: el *cluster* A representa la movilidad de estudiantes entre establecimientos municipales, el *cluster* B aquella entre municipales y particulares subvencionados, el *cluster* C comprende los cambios desde particulares subvencionados hacia municipales, y el *cluster* D agrupa cambios entre particulares subvencionados e incluye un pequeño flujo que considera establecimientos particulares pagados. De las Figuras 5.16 (b) y (c), se observa que el *cluster* A tiene la mayor representación porcentual de estudiantes del escenario (39%), mientras que los *cluster* B, C y D comprenden el 27%, 12% y 23% y de los cambios respectivamente.

Cluster	Copago rbd origen	Copago rbd preferencia	IVE rbd origen	IVE rbd preferencia	Promedio SIMCE rbd origen	Promedio SIMCE rbd preferencia	Dependencia rbd origen	Dependencia rbd preferencia
A	P61 (Gratuito)	P61 (Gratuito)	0,84	0,81	P22	P41	Municipal	Municipal
B	P61 (Gratuito)	P68	0,83	0,67	P26	P47	Municipal	Part. Subv.
C	P62	P61 (Gratuito)	0,76	0,78	P39	P52	Part. Subv.	Municipal
D	P62	P76	0,72	0,61	P47	P56	Part. Subv.	Part. Subv.

Tabla 5.10: Centroides de *clusters*.

La Tabla 5.10 y la Figura 5.16 (a) muestran los centroides de cada *cluster*. Se observa que el Copago es constante sólo en el *cluster* A (movilidad entre gratuitos), habiendo un aumento en este en los *clusters* B y D, y una disminución en el *cluster* C. Por otro lado, la vulnerabilidad (IVE) disminuye en los *clusters* A, B y D al pasar del establecimiento de origen al preferido,

mientras que el índice aumenta en el *cluster C*. En términos de resultados SIMCE, en todos los *clusters* se observa una mejora en rendimiento al pasar desde el establecimiento de origen al preferido, siendo todas estas diferencias bastante pronunciadas.

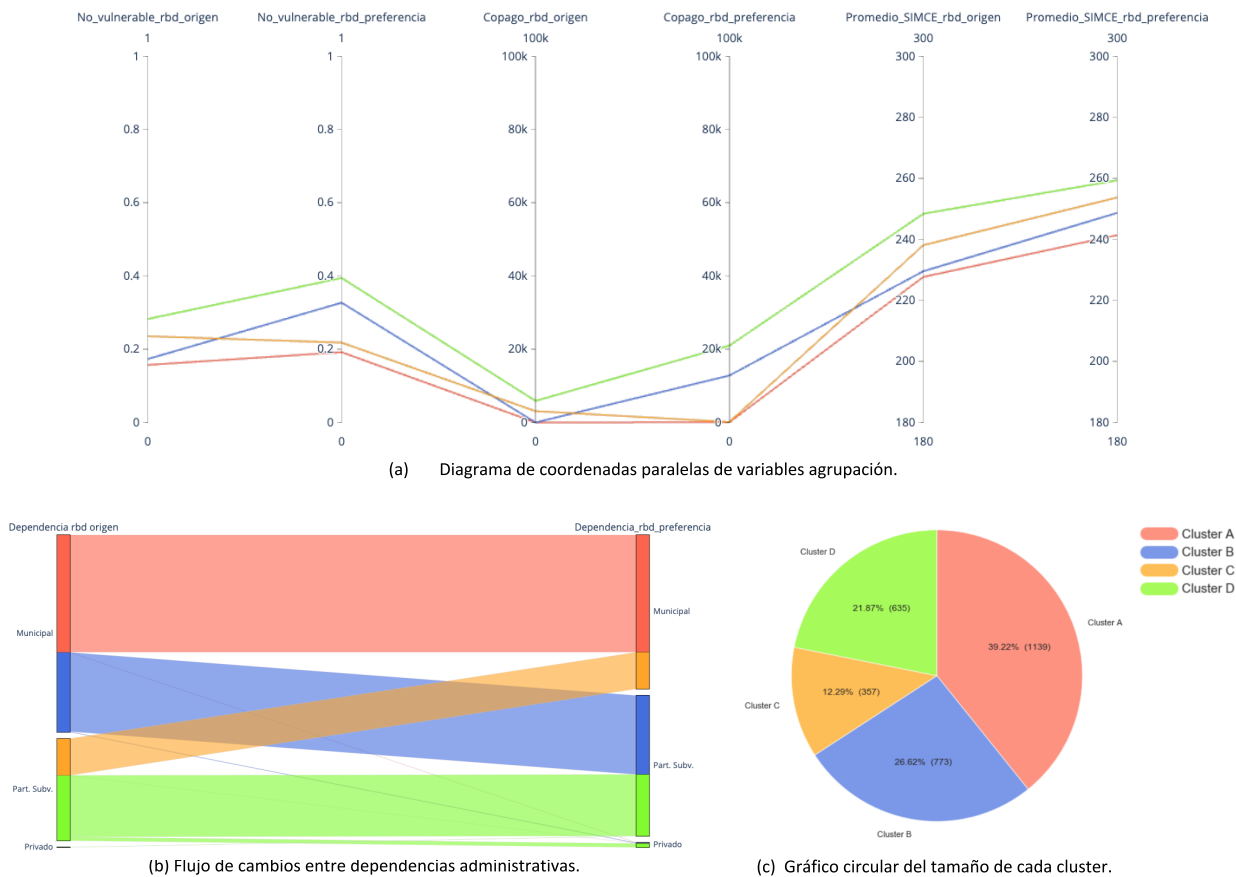


Figura 5.16: Variables a nivel de establecimiento.

Al comparar las variables a nivel de estudiante, en la Figura 5.17 se observa, para la variable de vulnerabilidad (No participa en la SEP), que los *clusters A* y *B* son los "menos no vulnerables" (entre 10 % y 15 % de los estudiantes no son Beneficiarios SEP), es decir, son los más vulnerable, seguidos, en dirección decreciente de vulnerabilidad, por el *clusters C* (18 % de los estudiantes no son Beneficiarios SEP), en tanto que el *cluster D* tienen la menor composición de estudiantes vulnerables (35 % de los estudiantes no son Beneficiarios SEP). Con respecto al Ingreso del Hogar, todos los *clusters* representan ingresos entre \$300.000 y \$400.000. Para la variable Educación de los Padres, los padres de estudiantes de todos los *clusters* presentan en promedio educación entre tercero medio y educación media completa.

Los estudiantes de los *clusters A* y *B* y *C* obtienen puntajes SIMCE promedio en torno a los 230 puntos, en tanto, el *cluster D* sobresale con puntajes promedio de 250 puntos. Por otra parte, los estudiantes de los *clusters A*, *B* y *D* recorren distancias similares (Distancia Hogar Colegio), cercanas a los 3 km, en tanto que los estudiantes del *cluster C* recorren distancias menores de 1,8 km en promedio. Asimismo, no se constatan diferencias significativas en la variable de género (Porcentaje de Hombre), teniendo todos los *clusters* una composición de hombres entre 45 % y 55 %. Por otro lado, se observan pocas diferencias en el rendimiento académico de los estudiantes (Percentil Promedio de Notas): estudiantes de los *clusters A*,

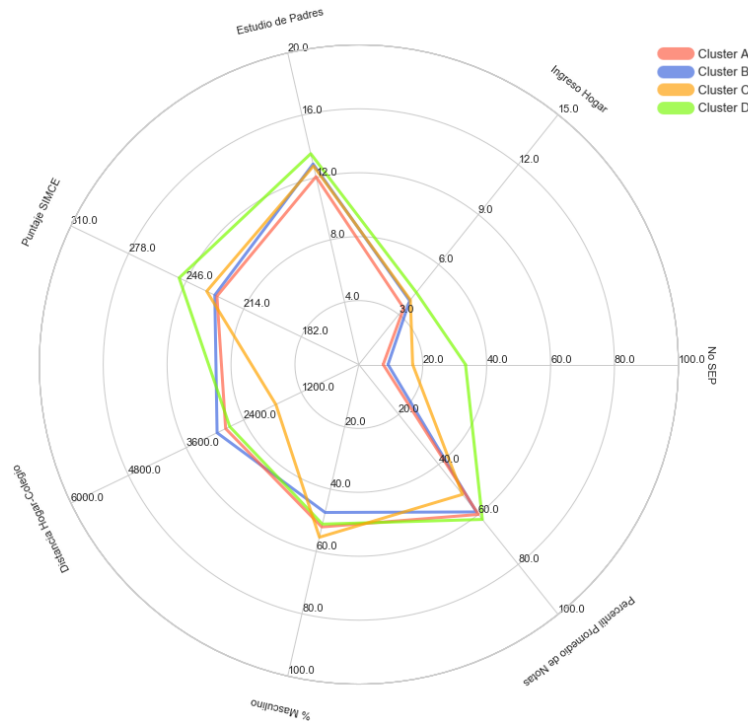


Figura 5.17: Variables a nivel de estudiante.

B y D tienen un rendimiento promedio cercano al percentil 60, mientras que los estudiantes del *cluster* C tienen un menor rendimiento (percentil 50).

En este escenario se pueden destacar las siguientes características:

- Hay una similar preferencia por establecimientos particulares subvencionados (*clusters* B y D) que por municipales (*cluster* A y C).
- Hay una preferencia marcada por establecimientos menos vulnerables (excepto *cluster* C) y con mejores resultados SIMCE.
- Estudiantes del *cluster* C recorren menores distancias y tienen menor rendimiento académico.
- Movilidad hacia particulares subvencionados aumenta en términos de copago, independiente de la dependencia del establecimiento de origen.

SAE Admisión

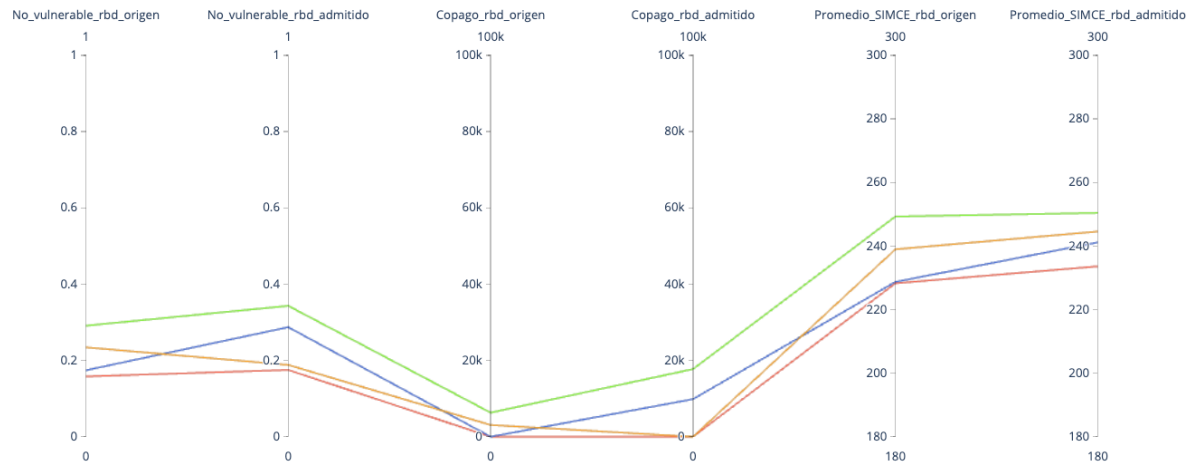
En este escenario hay 4 *clusters*. A partir de la Figura 5.18 (b), se observa que los *cluster* quedan en su mayoría definidos por la relación entre dependencias administrativas de origen y destino: el *cluster* A representa la movilidad de estudiantes entre establecimientos municipales, el *cluster* B aquella entre municipales y particulares subvencionados, el *cluster* C comprende los cambios desde particulares subvencionados hacia municipales, y el *cluster* D agrupa cambios entre particulares subvencionados (con un pequeño flujo desde y hacia particulares pagados). De las Figuras 5.18 (b) y (c), se observa que el *cluster* A tiene la mayor

representación porcentual de estudiantes del escenario (44%), mientras que los *cluster* B, C y D comprenden el 22%, 15% y 19% y de los cambios respectivamente.

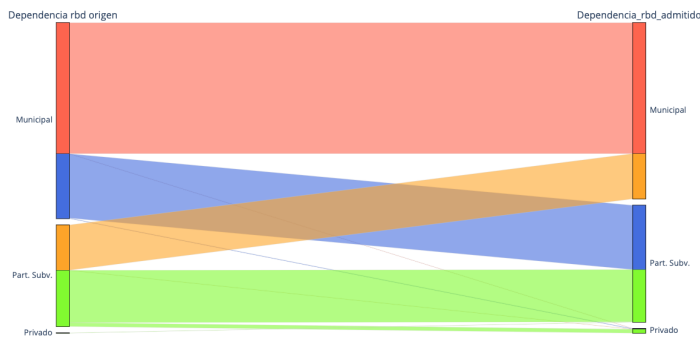
Cluster	Copago rbd origen	Copago rbd admitido	IVE rbd origen	IVE rbd admitido	Promedio SIMCE rbd origen	Promedio SIMCE rbd admitido	Dependencia rbd origen	Dependencia rbd admitido
A	P61 (Gratuito)	P61 (Gratuito)	0,84	0,82	P24	P32	Municipal	Municipal
B	P61 (Gratuito)	P65	0,83	0,71	P24	P41	Municipal	Part. Subv.
C	P62	P61 (Gratuito)	0,77	0,81	P39	P43	Part. Subv.	Municipal
D	P62	P73	0,71	0,66	P48	P48	Part. Subv.	Part. Subv.

Tabla 5.11: Centroides de *clusters*.

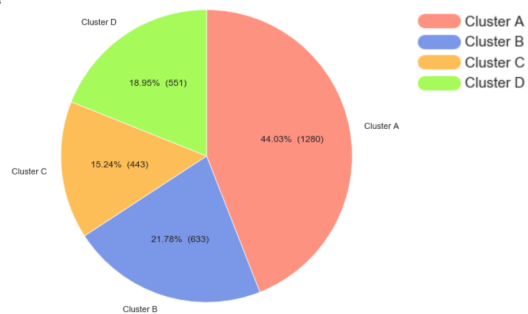
La Tabla 5.11 y la Figura 5.18 (a) muestran los centroides de cada *cluster*. Se observa que el Copago es constante sólo en el *cluster* A (movilidad entre municipales, gratuitos), habiendo un aumento en este atributo para los *clusters* B y D, y una disminución en el *cluster* C. La vulnerabilidad (IVE) disminuye en los *clusters* A, B y D al pasar desde el establecimiento de origen al admitido, y sólo aumenta en el *cluster* C. En términos de resultados SIMCE, en todos los *clusters* se observa una mejora en rendimiento, siendo particularmente elevado este efecto en el *cluster* B.



(a) Diagrama de coordenadas paralelas de variables agrupación.



(b) Flujo de cambios entre dependencias administrativas.



(c) Gráfico circular del tamaño de cada cluster.

Figura 5.18: Variables a nivel de establecimiento.

Al comparar las variables a nivel de estudiante, en la Figura 5.19 se observa, para la variable de vulnerabilidad (No participa en la SEP), que los *clusters* A y B son los "menos

no vulnerables" (menos del 10% de los estudiantes no son Beneficiarios SEP), es decir, son los más vulnerable, seguidos, en dirección decreciente de vulnerabilidad, por el *clusters* C (20% de los estudiantes no son Beneficiarios SEP) y finalmente por el *cluster* D (32% de los estudiantes no son Beneficiarios SEP). Con respecto al Ingreso del Hogar, todos los *clusters* representan ingresos entre \$300.000 y \$400.000. Para la variable Educación de los Padres, los padres de estudiantes de todos los *clusters* presentan en promedio educación entre tercero medio y enseñanza media completa.

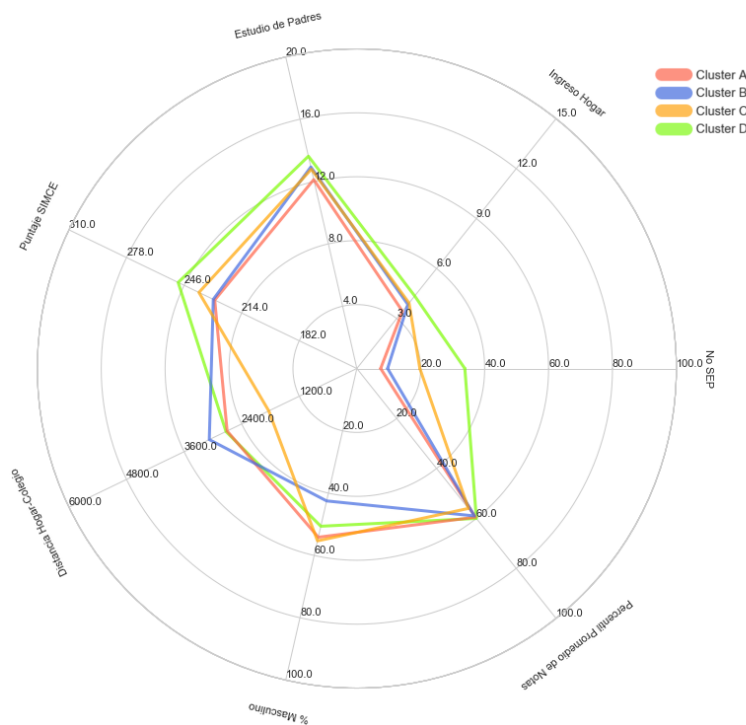


Figura 5.19: Variables a nivel de estudiante.

Los estudiantes de los *clusters* A y B obtienen puntajes SIMCE promedio en torno a los 230 puntos, seguidos, en orden creciente, por el *cluster* C (241 puntos) y finalmente por el *cluster* D (250 puntos). Los estudiantes de los *clusters* A, B y D recorren distancias similares (Distancia Hogar Colegio) cercanas a los 3 km, mientras que los estudiantes del *cluster* C recorren en promedio una distancia menor de 2 km. Por otro lado, no se constatan diferencias significativas en la variable de género (Porcentaje de Hombres) en los *clusters* A, C y D, teniendo estos una composición de hombres entre 55% y 60%, en tanto que se aprecia un sesgo hacia composición femenina en el *cluster* B (42% masculino). Además, se observan pocas diferencias en el rendimiento académico de los estudiantes (Percentil Promedio de Notas): todos los *clusters* tienen un rendimiento promedio cercano al percentil 60.

En este escenario se pueden destacar las siguientes características:

- El *cluster* D tiene una composición menos vulnerable: el Copago promedio de los establecimientos de origen y destino es mayor, su IVE es menor, sus puntajes SIMCE son mayores, sus padres tienen levemente mayor educación y perciben mayores ingresos.
- Hay una menor admisión en establecimientos particulares subvencionados (*clusters* B y D).

- Los cambios llevan a establecimientos menos vulnerables excepto en el *cluster* C.
- Todos los estudiantes que se cambian llegan a EE con mejor rendimiento, medido por el SIMCE (en promedio).
- Se observa un sesgo de género en el *cluster* B, compuesto por un 42 % de hombres.
- Todos los hogares de los estudiantes que cambian de establecimientos son en promedio de bajos ingresos.

Patrones de Cambios Forzado en Coquimbo - La Serena

Se observan diferencias en patrones entre escenarios en términos de representatividad de los *clusters*: la preferencia por establecimientos municipales crece levemente con respecto al escenario Pre SAE, aunque apunta hacia EE de mejores resultados SIMCE, resultando en una asignación aún mayor a esta dependencia, pero que en promedio obtienen menores puntajes a los demandados. No obstante, puntajes SIMCE a nivel de establecimiento en el escenario de admisión son mayores a aquellos previos al SAE. Más aún, las preferencias apuntan a establecimientos con mejor rendimiento, medido por el SIMCE que aquellos en los que son asignados, es decir, no existe una oferta educativa que satisfaga la demanda en términos de resultados SIMCE.

Además, la calidad de los perfiles de cambio (coeficiente de silueta) es buena en todos los casos, lo que indica que los patrones están bien definidos y estructurados. Al ser rígida la demanda de escuelas por motivo de fin del establecimiento de origen, el comportamiento de este tipo de movilidad está mejor categorizada, lo que se nota también porque todos los escenarios tienen igual número de *clusters*, y todos definidos por las mismas transiciones entre dependencias administrativas.

5.2.2. Cambios No Forzados

5.2.2.1. Escenario Pre SAE

En este escenario hay 4 *clusters*. A partir de la Figura 5.20 (b), se observa que estos quedan definidos por la relación entre dependencias administrativas de origen y destino: el *cluster* A representa la movilidad de estudiantes entre establecimientos municipales, el *cluster* B aquella desde municipales hacia particulares subvencionados, el *cluster* C comprende los cambios desde particulares subvencionados hacia municipales y el *cluster* D agrupa cambios entre particulares subvencionados además de aquellos que involucran establecimientos particulares pagados. De las Figuras 5.20 (b) y (c), se observa que el *cluster* D tiene la mayor representación porcentual de estudiantes del escenario (56 %), mientras que los *clusters* A, B y C comprenden el 12 %, 12 % y 20 % de los cambios, respectivamente.

La Tabla 5.12 y la Figura 5.20 (a) muestran los centroides de cada *cluster*. Se observa que el Copago varía poco en los *clusters* A y D, siendo el *cluster* D el con mayor valor (en torno a \$25.000). La misma variable aumenta en los casos de movilidad del *cluster* B y disminuye en el *cluster* C. La vulnerabilidad (IVE) aumenta en el *cluster* C y disminuye en el *cluster* B. El resto de los *clusters* (A y D) mantienen su nivel de vulnerabilidad. En términos de resultados

Cluster	Copago rbd origen	Copago rbd destino	IVE rbd origen	IVE rbd destino	Promedio SIMCE rbd origen	Promedio SIMCE rbd destino	Dependencia rbd origen	Dependencia rbd destino
A	P61	P63	0,79	0,77	P22	P40	Municipal	Municipal
B	P61	P75	0,77	0,66	P27	P46	Municipal	Part. Subv.
C	P73	P63	0,65	0,76	P52	P39	Part. Subv.	Municipal
D	P79	P81	0,59	0,58	P59	P58	Part. Subv./Priv.	Part. Subv./Priv.

Tabla 5.12: Centroides de *clusters*.

SIMCE, en los *clusters* A y B hay una mejora en rendimiento, siendo ambos efectos muy pronunciados. Para la misma variable, el *cluster* C disminuye en puntaje y el *cluster* D se mantiene casi invariante.

Figura 5.20: Variables a nivel de establecimiento.

Al comparar las variables a nivel de estudiante, en la Figura 5.21 se observa, para la variable de vulnerabilidad (No participa en la SEP), que el *cluster* A es el "menos vulnerable" (en torno al 30% de los estudiantes no son Beneficiarios SEP), es decir, es el más vulnerable, seguido, en orden decreciente de vulnerabilidad, por los *clusters* B (40% de los estudiantes no son Beneficiarios SEP), luego por los *clusters* C (60%) y D (70%). Con respecto al Ingreso del Hogar, los *cluster* A, B y C representan ingresos entre \$300.000 y \$500.000, mientras que familias del *cluster* D tienen un ingreso entre \$500.000 y \$600.000 en promedio. Por otro lado, para la variable Educación de los Padres, los padres de estudiantes en los *clusters* A, B y C presentan en promedio educación hasta tercero medio o enseñanza media completa, en tanto que los padres de estudiantes del *cluster* D tienen en promedio estudios incompletos en CFT o IP.

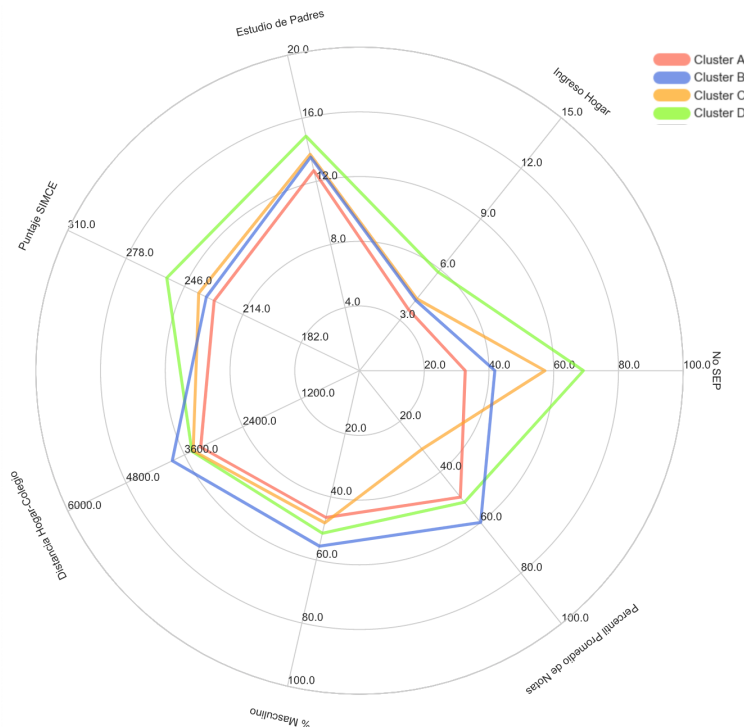


Figura 5.21: Variables a nivel de estudiante.

Los estudiantes de los *clusters* A, B y C obtienen en promedio puntajes SIMCE en torno a 235 puntos, en tanto que los estudiantes del *cluster* D obtienen en promedio puntajes superiores (260 puntos). Los estudiantes del *cluster* B recorren mayores distancias (Distancia Hogar Colegio), en promedio de 4 km, mientras que los estudiantes de los *clusters* A, C y D recorren en torno a los 3,5 km. No se constatan diferencias significativas en la variable de género (Porcentaje de Hombres), teniendo todos los *clusters* una composición masculina entre 45 % y 55 %. Por otro lado, se observan importantes diferencias en el rendimiento académico de los estudiantes (Percentil Promedio de Notas): estudiantes del *cluster* C tienen muy bajo rendimiento (percentil 30), estudiantes de los *clusters* A y D se encuentran en el percentil 50, y aquellos del *cluster* D tienen el mayor rendimiento promedio (percentil 60).

En este escenario se pueden destacar las siguientes características:

- La vulnerabilidad del *cluster* D es menor: menor IVE, menos Beneficiarios SEP, mayor Copago, mayor rendimiento, medido por el SIMCE del establecimiento, mayor Ingreso del Hogar, mayor Educación de los Padres y Puntajes SIMCE más altos. Esto concuerda con que además en este *cluster* se incluye la movilidad desde y hacia establecimientos particulares pagados.
- Estudiantes que cambian desde EE municipales a particulares subvencionados (*cluster* C) tienen un marcado bajo rendimiento académico.
- La mayoría de los cambios ocurre entre establecimientos particulares subvencionados (*cluster* D)

5.2.2.2. Escenario SAE

SAE Preferencia

En este escenario hay 5 *clusters*. A partir de la Figura 5.22 (b), se observa que estos quedan parcialmente definidos por la relación entre dependencias administrativas de origen y destino: el *cluster* A representa la movilidad de estudiantes entre municipales, el *cluster* B aquella desde municipales hacia particulares subvencionados, el *cluster* C comprende los cambios desde particulares subvencionados hacia municipales, el *cluster* D agrupa cambios entre establecimientos particulares subvencionados, y el *cluster* E representa algunos cambios entre particulares subvencionados, además de aquellos que involucran establecimientos particulares pagados. De las Figuras 5.22 (b) y (c), se observa que el *cluster* D tiene la mayor representación porcentual de estudiantes del escenario (46 %), mientras que los *clusters* A, B, C y E comprenden el 10 %, 14 %, 17 % y 13 % de los cambios, respectivamente.

Cluster	Copago rbd origen	Copago rbd preferencia	IVE rbd origen	IVE rbd preferencia	Promedio SIMCE rbd origen	Promedio SIMCE rbd preferencia	Dependencia rbd origen	Dependencia rbd preferencia
A	P61 (Gratuito)	P61 (Gratuito)	0,82	0,77	P26	P51	Municipal	Municipal
B	P61 (Gratuito)	P69	0,81	0,66	P27	P50	Municipal	Part. Subv.
C	P66	P61 (Gratuito)	0,69	0,77	P48	P53	Part. Subv.	Municipal
D	P71	P76	0,64	0,58	P53	P59	Part. Subv.	Part. Subv.
E	P86	P90	0,33	0,26	P70	P74	Subv./Priv.	Privado

Tabla 5.13: Centroides de *clusters*.

La Tabla 5.13 y la Figura 5.22 (a) muestran los centroides de cada *cluster*. Se observa

que el Copago se mantiene constante en el *cluster* A (cambio entre gratuitos), aumenta en los *clusters* B, D y E (siendo el *cluster* E el con valor más elevado), mientras que sólo disminuye en el *cluster* C. Por otro lado, la vulnerabilidad (IVE) aumenta sólo en los cambios del *cluster* C, disminuyendo en los cuatro casos restantes. Por otro lado, en términos de resultados SIMCE, en todos los *clusters* se observa que los establecimientos de preferencia tienen mejores puntajes que los de origen.

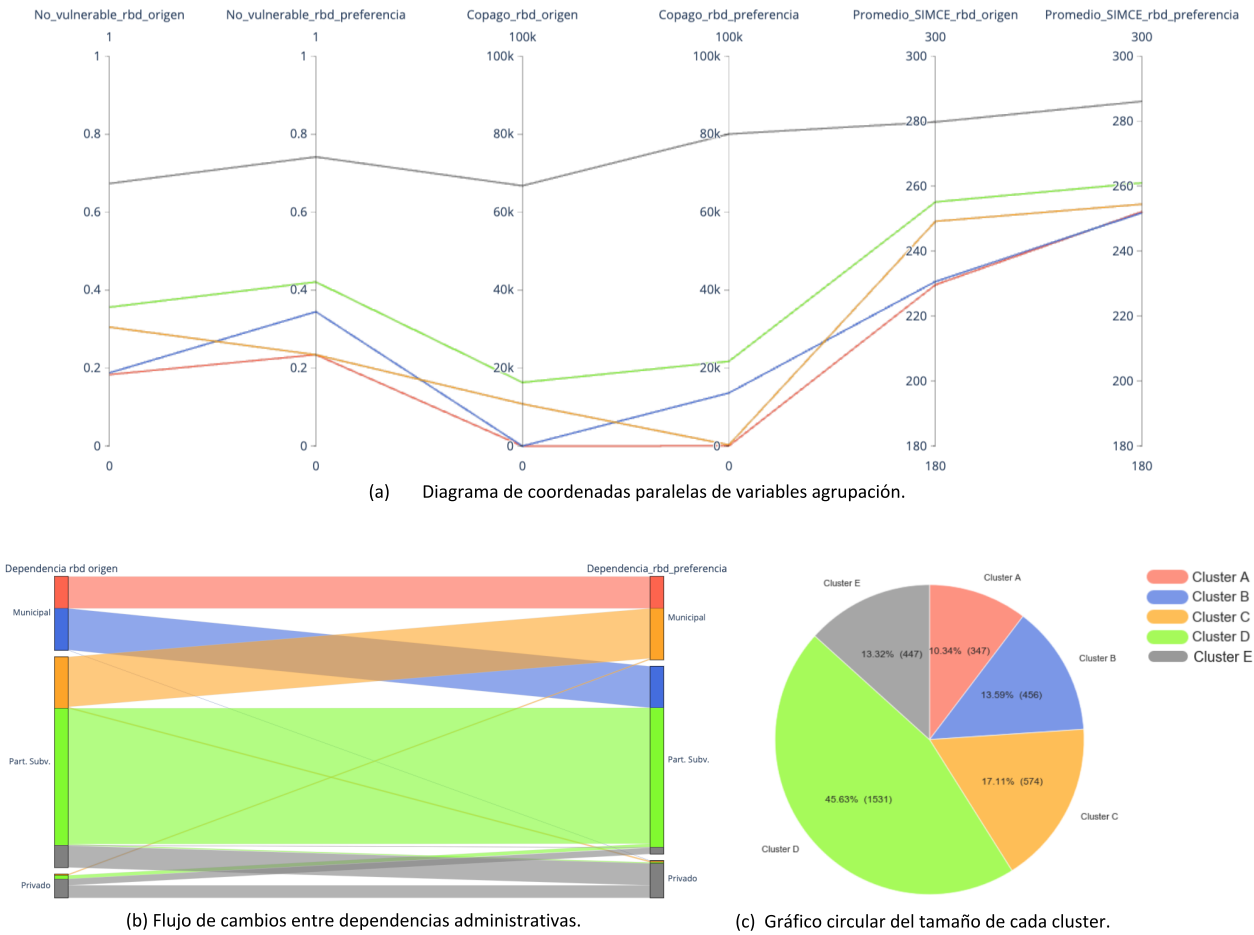


Figura 5.22: Variables a nivel de establecimiento.

Al comparar las variables a nivel de estudiante, en la Figura 5.23 se observa, para la variable de vulnerabilidad (No participa en la SEP), que los *clusters* A y B son los "menos vulnerables" (en torno al 15 % de los estudiantes no son Beneficiarios SEP), es decir, son los más vulnerables, seguidos, en orden decreciente de vulnerabilidad, por el *cluster* C (45 % de los estudiantes no son Beneficiarios SEP) y por el *cluster* D (55 % de los estudiantes no son Beneficiarios SEP); finalmente, el *cluster* D presenta muy baja vulnerabilidad (95 % de los estudiantes no son Beneficiarios SEP). Con respecto al Ingreso del Hogar, los *clusters* A, B y C representan ingresos promedio entre \$300.000 y \$500.000, mientras que familias del *cluster* D tienen un ingreso entre \$500.000 y \$600.000, en tanto que familias del *cluster* E perciben ingresos promedio entre \$1.200.000 y \$1.400.000. Por otro lado, para la variable Educación de los Padres, los *clusters* A, B y C representan en promedio educación media completa; el *cluster* D corresponde a educación incompleta en IP o CFP; y finalmente, padres del *cluster* E tienen en promedio estudios universitarios incompletos.

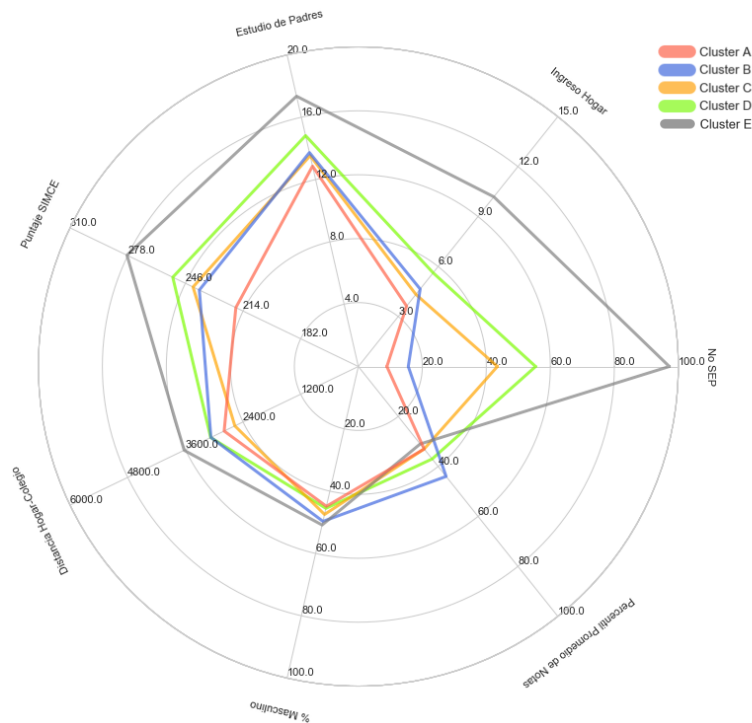


Figura 5.23: Variables a nivel de estudiante.

Los estudiantes del *cluster* A obtienen puntajes SIMCE en promedio de 215 puntos, en tanto que aquellos de los *clusters* B, C y D obtienen puntajes en torno a los 245 puntos en promedio, mientras que los estudiantes del *cluster* E obtienen los mayores puntajes, de 280 en promedio. Asimismo, los estudiantes del *cluster* E recorren mayores distancias (Distancia Hogar Colegio), en promedio de 3,6 km. Estudiantes de los *clusters* B y D recorren en torno a los 3 km. y aquellos de los *clusters* A y C recorren una menor distancia de 2,5 km en promedio. Por otra parte, no se constatan diferencias en la variable de género (Porcentaje de Hombre), teniendo todos los *clusters* una composición de hombres entre 50% y 60%. Por otro lado, se observan pocas diferencias en el rendimiento académico de los estudiantes (Percentil Promedio de Notas), pues el rendimiento promedio de los estudiantes de todos los *clusters* se ubican entre los percentiles 32 y 43, siendo E, A, C, D y B su orden en sentido creciente.

En este escenario se pueden destacar las siguientes características:

- Los *clusters* ya no están totalmente ordenados de acuerdo a la transición de dependencia administrativa, lo que sugiere un mayor desorden en la organización de la movilidad en el escenario no forzado.
- El *cluster* E resalta como grupo de estudiantes de GSE alto: tienen nula vulnerabilidad, altos ingresos y estudios de padres, altos puntajes SIMCE, copago alto y asisten a establecimientos de alto SIMCE.
- La mayor movilidad en este escenario ocurre entre establecimientos particulares subvencionados.
- Se observa mayor presencia de movilidad que involucra a establecimientos particulares pagados.

SAE Admisión

Este escenario hay 5 *clusters*. A partir de la Figura 5.24 (b), se observa que estos quedan parcialmente definidos por la relación entre dependencias administrativas de origen y destino: el *cluster* A representa la movilidad de estudiantes entre municipales, el *cluster* B aquella desde municipales hacia particulares subvencionados, el *cluster* C comprende los cambios desde particulares subvencionados hacia municipales, el *cluster* D agrupa cambios entre establecimientos particulares subvencionados, y el *cluster* E representa algunos cambios entre particulares subvencionados, además de aquellos que involucran establecimientos particulares pagados. De las Figuras 5.24 (b) y (c), se observa que el *cluster* D tiene la mayor representación porcentual de estudiantes del escenario (44%), mientras que los *clusters* A, B, C y E comprenden el 13%, 11%, 19% y 12% de los cambios, respectivamente.

Cluster	Copago rbd origen	Copago rbd admitido	IVE rbd origen	IVE rbd admitido	Promedio SIMCE rbd origen	Promedio SIMCE rbd admitido	Dependencia rbd origen	Dependencia rbd admitido
A	P61 (Gratuito)	P61 (Gratuito)	0,81	0,78	P27	P45	Municipal	Municipal
B	P61 (Gratuito)	P67	0,82	0,69	P27	P43	Municipal	Part. Subv.
C	P69	P61 (Gratuito)	0,68	0,78	P50	P45	Part. Subv.	Municipal
D	P72	P71	0,64	0,65	P53	P50	Part. Subv.	Part. Subv.
E	P86	P90	0,33	0,25	P70	P74	Subv./Priv.	Privado

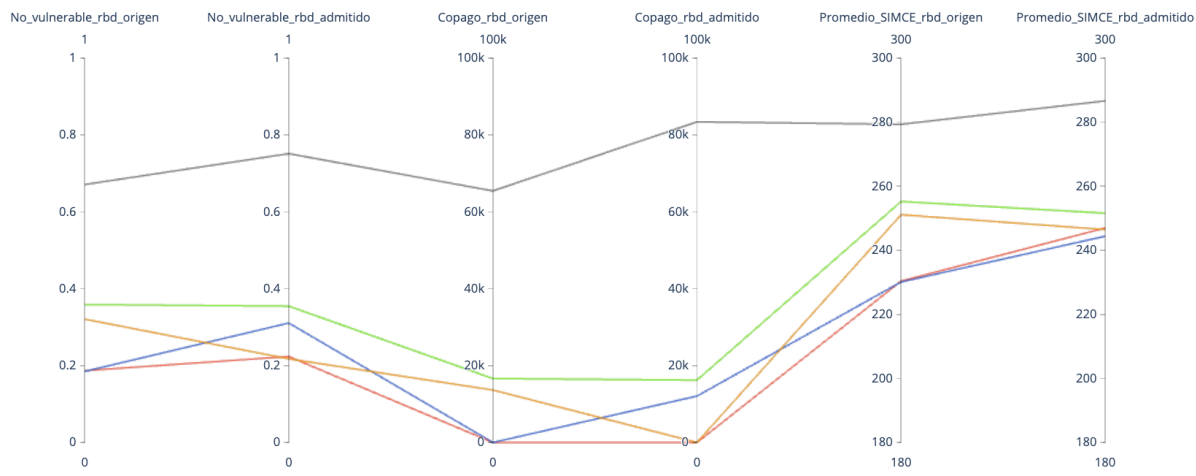
Tabla 5.14: Centroides de *clusters*.

La Tabla 5.14 y la Figura 5.24 (a) muestran los centroides de cada *cluster*. Se observa que el Copago disminuye sólo en el *cluster* C, es constante en los *clusters* A y D, y aumenta en los *clusters* B y E, en el cual la variable llega al valor promedio de \$80.000. Por otra parte, la vulnerabilidad (IVE) disminuye en los *clusters* A, B y E, aumenta en el *cluster* C, en tanto que permanece constante en el *cluster* D. En términos de resultados SIMCE, en los *clusters* A, B y E hay una mejora en rendimiento, mientras que en los *clusters* C y D los puntajes disminuyen.

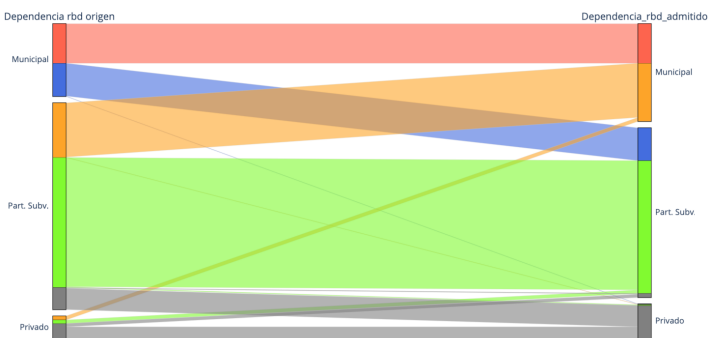
Al comparar las variables a nivel de estudiante, en la Figura 5.25 se observa, para la variable de vulnerabilidad (No participa en la SEP), que los *clusters* A y B son los "menos vulnerables" (en torno al 12% de los estudiantes no son Beneficiarios SEP), es decir, son los más vulnerable, seguidos, en orden decreciente de vulnerabilidad, por los *clusters* C (45% de los estudiantes no son Beneficiarios SEP); y finalmente, el *cluster* E tiene una muy baja vulnerabilidad (95% de los estudiantes no son Beneficiarios SEP).

Con respecto al Ingreso del Hogar, los *clusters* A, B y C representan ingresos entre \$300.000 y \$500.000, mientras que familias del *cluster* D perciben ingreso entre \$500.000 y \$600.000; en tanto que el *cluster* E representa los mayores ingresos, que figuran entre \$1.200.000 y \$1.400.000 en promedio. Por otro lado, para la variable Educación de los Padres, los *clusters* A, B, C y D presentan en promedio educación media completa, en tanto padres del *cluster* E poseen, en general, educación universitaria completa.

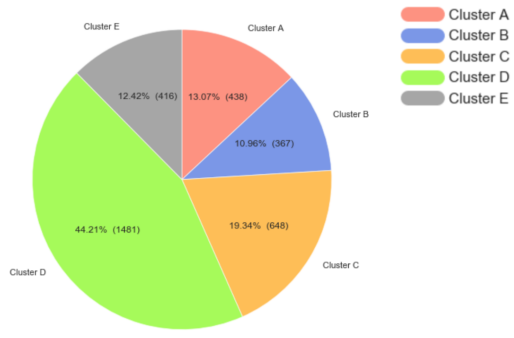
Los estudiantes de los *clusters* B, C y D obtienen en promedio puntajes SIMCE en torno a 246 puntos, aquellos del *cluster* A obtiene en promedio 220 puntos, y los del *cluster* E superan ampliamente al resto, con un rendimiento promedio de 280 puntos. Asimismo, los estudiantes de todos los *clusters* recorren distancias (Distancia Hogar Colegio) entre 3 y 3,6 km en promedio, siendo C, A, B, D, E su orden en sentido creciente. Con respecto a la variable de



(a) Diagrama de coordenadas paralelas de variables agrupación.



(b) Flujo de cambios entre dependencias administrativas.



(c) Gráfico circular del tamaño de cada cluster.

Figura 5.24: Variables a nivel de establecimiento.

género (Porcentaje de Hombres), todos los *clusters* tienen una composición homogénea (entre 45 % y 50 %). Por otro lado, se observan pequeñas diferencias en el rendimiento académico de los estudiantes (Percentil Promedio de Notas), los que en promedio están entre los percentiles 32 y 40, valores que indican un bajo desempeño de estos.

En este escenario se pueden destacar las siguientes características:

- Los *clusters* ya no están totalmente ordenados de acuerdo a la transición de dependencia administrativa, lo que sugiere un mayor desorden en la organización de la movilidad en el escenario no forzado.
- La mayoría de los cambios ocurren entre escuelas subvencionadas (*cluster D*) y la minoría desde municipales hacia particulares subvencionados (*cluster B*).
- La vulnerabilidad del *cluster E* es menor: menor IVE, menos beneficiarios SEP, mayor Copago, mayor rendimiento, medido por el SIMCE del establecimiento, mayor Ingreso del Hogar y mayor Educación de los Padres.
- Los establecimientos particulares pagados tienen una mayor participación que en otros escenarios.

Discusión de patrones de cambio No Forzado en Rancagua - Machalí

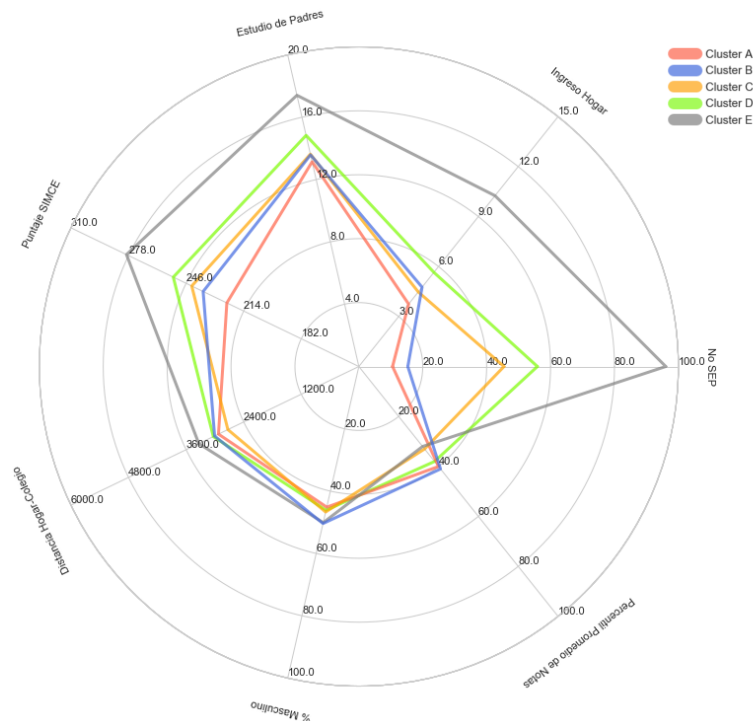


Figura 5.25: Variables a nivel de estudiante.

En todos los escenarios se observa que los cambios entre municipales tienen baja representatividad. Además, en los escenarios SAE aparece un quinto *cluster* caracterizado por estudiantes de GSE alto, junto con el aumento significativo de cambios desde y hacia establecimientos particulares pagados.

Además, gran parte de la movilidad desde particulares subvencionados (*clusters* C y D) tiene preferencias (SAE) por mayores puntajes SIMCE que aquellos que obtienen los establecimientos a los que son finalmente admitidos. Es más, estudiantes de estos *cluster* terminan, en promedio, en establecimientos con menor rendimiento, medido por el SIMCE, que sus EE de origen. Es decir, nuevamente se observa una demanda por establecimiento con mejores resultados SIMCE, para la cual no hay oferta suficiente.

Por otro lado, el *cluster* D en el escenario Pre SAE tiene una estructura más débil, esto pues debe incluir también los cambios de estudiantes de GSE alto, lo que reduce su puntaje de calidad, al ser estos, en parte, de dependencia privada. En la misma línea, en ambos casos SAE, el *cluster* E presenta menor estructura dado que involucra diversos flujos entre dependencias administrativas

Finalmente, se observa que el *cluster* C, originalmente (Pre SAE) compuesto, en promedio, por estudiantes con bajo rendimiento académico, deja de estarlo (en ambos escenarios SAE). Por otra parte, tanto vulnerabilidad como copago tienen comportamientos similares en todos los escenarios.

Patrones de Cambio en Coquimbo - La Serena

En los cambios Forzados priman aquellos que involucran a establecimientos municipales,

en tanto que los cambios No Forzados están compuestos principalmente por transiciones que involucran a particulares subvencionados, los cuales son de mayor copago que en el caso complementario.

La calidad de agrupación de los *clusters* es inferior en el caso de cambios No Forzados, lo que se explica por el carácter voluntario de estos cambios. Además en estos mismos aparece un quinto *cluster* en los escenarios SAE.

En términos de resultados SIMCE, al comparar los escenarios (Pre SAE, SAE Admisión, SAE Preferencia) entre sí para ambos casos (Forzado, No Forzado), se observan tendencias similares. En este sentido, se constata que bajo el nuevo sistema, existe una demanda por resultados SIMCE para la cual no hay oferta suficiente.

5.3. Rancagua - Machalí

5.3.1. Cambios Forzados

5.3.1.1. Escenario Pre SAE

En este caso hay cuatro *clusters*. A partir de la Figura 5.26 (b), se observa que estos quedan muy bien definidos por la relación entre dependencias administrativas de origen y destino: el *cluster* A representa la movilidad de estudiantes entre establecimientos municipales, el *cluster* B aquella desde municipales hacia particulares subvencionados, el *cluster* C comprende los cambios desde particulares subvencionados hacia municipales, y finalmente, en *cluster* D, incluye los cambios entre establecimientos particulares subvencionados. De las Figuras 5.26 (b) y (c), se observa que el *cluster* A tiene la mayor representación porcentual de estudiantes del escenario (32%), mientras que los *clusters* B, C y D comprenden el 27%, 15% y 26% de los cambios respectivamente.

La Tabla 5.15 y la Figura 5.26 (a) muestra los centroides de cada *cluster*. Se observa que el Copago no cambia en el *cluster* A (movilidad entre gratuitos), aumenta en los *clusters* B y D, en tanto que disminuye en el *cluster* C. La vulnerabilidad (IVE) permanece constante en los *clusters* A y E, en tanto que disminuye en el *cluster* B y aumenta en el *cluster* C. En términos de resultados SIMCE de establecimientos de origen y destino, cada uno de los *clusters* C, D y E representan cambios que aumentan en puntajes, mientras que sólo en el *cluster* A hay una disminución de este.

Cluster	Copago rbd origen	Copago rbd destino	IVE rbd origen	IVE rbd destino	Promedio SIMCE rbd origen	Promedio SIMCE rbd destino	Dependencia rbd origen	Dependencia rbd destino
A	P60 (Gratuito)	P60 (Gratuito)	0,76	0,74	P43	P40	Municipal	Municipal
B	P60 (Gratuito)	P72	0,76	0,67	P43	P59	Municipal	Part. Subv.
C	P71	P60 (Gratuito)	0,60	0,67	P53	P59	Part. Subv.	Municipal
D	P73	P78	0,57	0,58	P56	P60	Part. Subv.	Part. Subv.

Tabla 5.15: Centroides de *clusters*.

Al comparar las variables a nivel de estudiante, en la Figura 5.27 se observa, para la

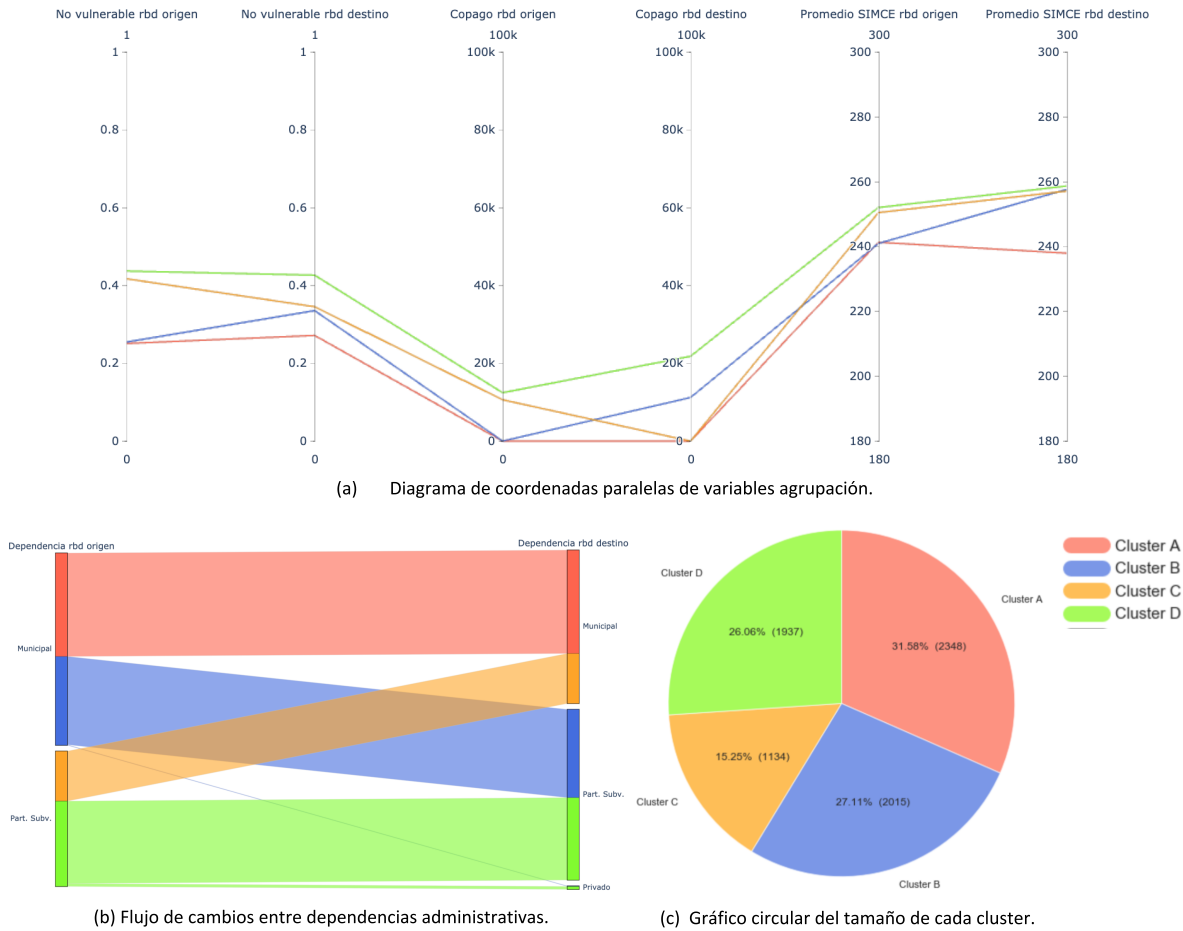


Figura 5.26: Variables a nivel de establecimiento.

variable de vulnerabilidad (No participa en la SEP), que los *clusters* A y B son los "menos no vulnerables" (cerca del 35% de los estudiantes no son Beneficiarios SEP), es decir, son los más vulnerables, seguidos, en orden decreciente de vulnerabilidad, por los *clusters* C y D (65% de los estudiantes no son Beneficiarios SEP). Por otro lado, familias de estudiantes cuyos cambios son caracterizados por todos los *clusters* tienen un Ingreso del Hogar similar entre \$300.000 y \$500.000, siendo el de los *clusters* A y B levemente menor. Por otro lado, todos los *clusters* son caracterizados por Educación de Padre en promedio hasta tercero medio o enseñanza media completa, siendo nuevamente la de los *clusters* A y B levemente menor.

Los mayores puntajes SIMCE los obtienen los estudiantes del *cluster* C (cerca de 253 puntos en promedio), seguidos de los *clusters* A y D (cerca de 245 puntos), y terminando con los estudiantes del *cluster* B (alrededor de 235 puntos). Por otro lado, los estudiantes de los todos los *clusters* recorren distancias promedio similares de 3 km (Distancia Hogar Colegio). En tanto, se observan diferencias en la variable de género (Porcentaje de Hombres): el *cluster* A lo forman 40% de hombres, los *cluster* B y C un 50%, y el *cluster* D lo componen en un 60% el género masculino. Asimismo, estudiantes del *cluster* C tienen mejor rendimiento académico, del percentil 65 en promedio (Percentil Promedio de Notas), en tanto que el resto de los *clusters* se ubican en promedio en torno al percentil 53 de rendimiento en notas.

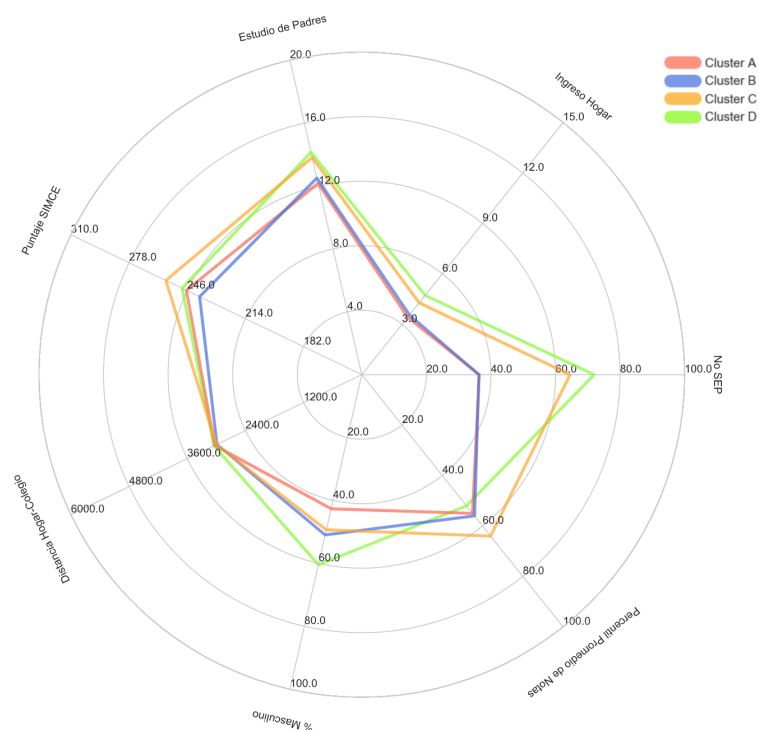


Figura 5.27: Variables a nivel de estudiante.

En este escenario se pueden destacar las siguientes características:

- Los *clusters* C y D sobresalen en las variables Puntaje SIMCE (estudiante), Educación de los Padres, Ingreso del Hogar y No Vulnerabilidad. Además, estos mismos *clusters* representan movilidad desde EE particulares subvencionados, con copago mayor en EE de origen, menor vulnerabilidad y mejores resultados SIMCE (EE).
- Los *clusters* A y B son los de mayor vulnerabilidad: menor Puntaje SIMCE (estudiante) e Ingreso del Hogar, con más Beneficiarios SEP; asisten a establecimientos de bajo Copago, con alta Vulnerabilidad y con resultados SIMCE bajos. Estos son además los *clusters* con establecimientos de origen de dependencia municipal.
- Hay sesgos importantes en la variable de género.
- Se recorren distancias similares en promedio.
- La mayoría de los cambios ocurre desde la dependencia municipal.

5.3.1.2. Escenario SAE

SAE Preferencias

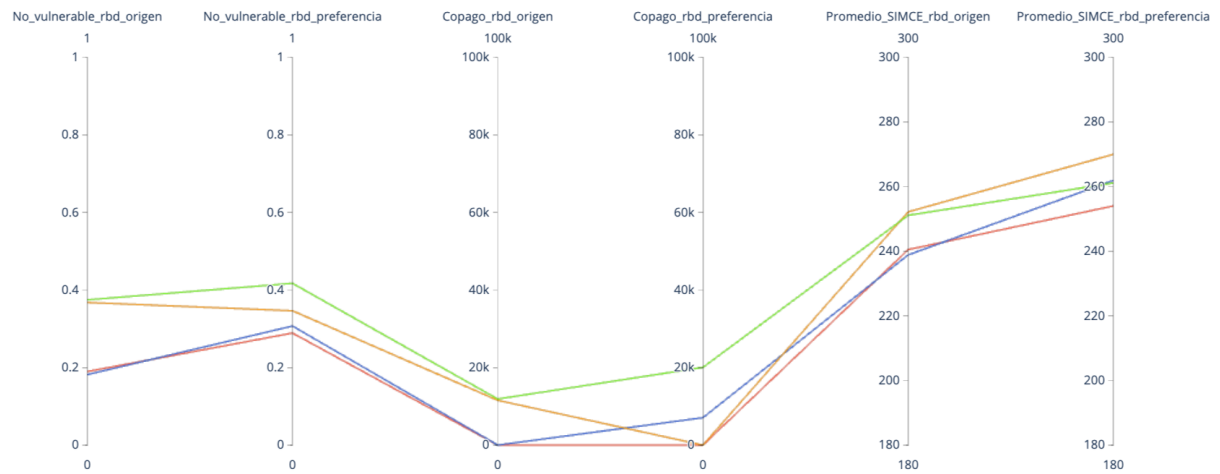
En este caso hay cuatro *clusters*. A partir de la Figura 5.28 (b), se observa que estos quedan definidos por la relación entre dependencias administrativas de origen y destino: el *cluster* A representa la movilidad de estudiantes entre establecimientos municipales, el *cluster* B aquella desde municipales hacia particulares subvencionados o particulares pagados, el *cluster* C comprende los cambios desde particulares subvencionados hacia municipales, y finalmente,

en *cluster* D, incluye los cambios desde establecimientos particulares subvencionados hacia aquellos de la misma dependencia o particulares pagados. De las Figuras 5.28 (b) y (c), se observa que el *cluster* A tiene la mayor representación porcentual de estudiantes del escenario (32%), mientras que los *clusters* B, C y D comprenden el 30%, 15% y 23% de los cambios respectivamente.

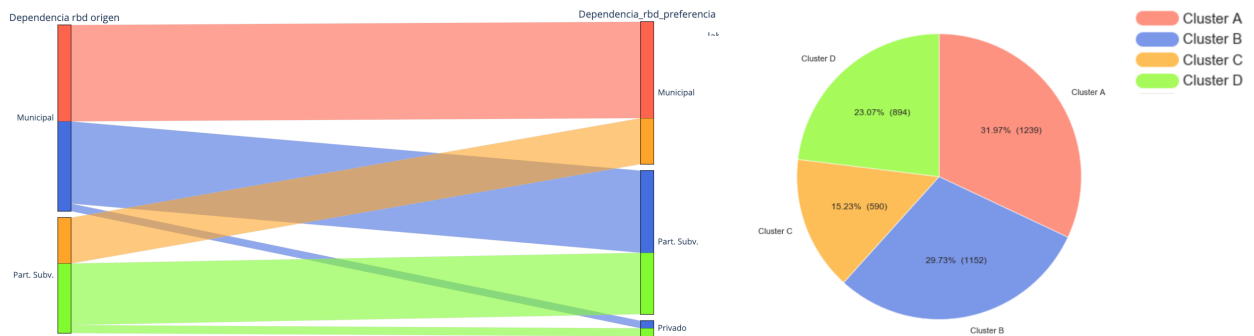
Cluster	Copago rbd origen	Copago rbd preferencia	IVE rbd origen	IVE rbd preferencia	Promedio SIMCE rbd origen	Promedio SIMCE rbd preferencia	Dependencia rbd origen	Dependencia rbd preferencia
A	P63 (Gratuito)	P63 (Gratuito)	0,81	0,71	P40	P52	Municipal	Municipal
B	P63 (Gratuito)	P64	0,82	0,69	P37	P61	Municipal	Part. Subv.
C	P66	P63 (Gratuito)	0,63	0,65	P50	P66	Part. Subv.	Municipal
D	P66	P71	0,63	0,58	P49	P61	Part. Subv.	Part. Subv.

Tabla 5.16: Centroides de *clusters*.

La Tabla 5.16 y la Figura 5.28 (a) muestra los centroides de cada *cluster*. Se observa que el Copago no cambia en el *cluster* A (movilidad entre gratuitos), aumenta en los *clusters* B y D, en tanto que disminuye en el *cluster* C. La vulnerabilidad (IVE) disminuye en los *clusters* A, B y D, mientras que aumenta en el *cluster* C. En términos de resultados SIMCE de establecimientos de origen y destino, todos los *clusters* representan cambios que aumentan en puntajes.



(a) Diagrama de coordenadas paralelas de variables agrupación.



(b) Flujo de cambios entre dependencias administrativas.

(c) Gráfico circular del tamaño de cada cluster.

Figura 5.28: Variables a nivel de establecimiento.

Al comparar las variables a nivel de estudiante, en la Figura 5.29 se observa, para la variable de vulnerabilidad (No participa en la SEP), que los *clusters* A y B son los "menos no

vulnerables" (8% de los estudiantes no son Beneficiarios SEP), es decir, son los más vulnerables, seguidos, en orden decreciente de vulnerabilidad, por los *clusters* C y D (50% de los estudiantes no son Beneficiarios SEP). Por otro lado, familias de estudiantes cuyos cambios son caracterizados por todos los *clusters* tienen un Ingreso del Hogar similar entre \$300.000 y \$500.000, siendo el de los *clusters* A y B levemente menor. Por otro lado, todos los *clusters* son caracterizados en promedio por Educación de Padres hasta tercer medio o enseñanza media completa, siendo nuevamente la de los *clusters* A y B levemente menor.

Los mayores puntajes SIMCE lo obtienen los estudiantes del *cluster* C (253 puntos en promedio), seguidos de cerca por los *clusters* A y D (246 puntos), y terminando con los estudiantes del *cluster* B (alrededor de 235 puntos). Por otro lado, los estudiantes de los todos los *clusters* recorren distancias promedio similares de 2,3 km (Distancia Hogar Colegio). En tanto, se observan diferencias en la variable de género (Porcentaje de Hombres): los *clusters* A y C lo forman 48% de hombres, mientras que los *cluster* B y D lo hacen un 60%. Asimismo, estudiantes del *cluster* A tienen mejor rendimiento académico, del percentil 65 en promedio (Percentil Promedio de Notas), en tanto que los *clusters* B y C se ubican en promedio en torno al percentil 53, y estudiantes del *cluster* D tienen el menor rendimiento (percentil 46 en promedio).

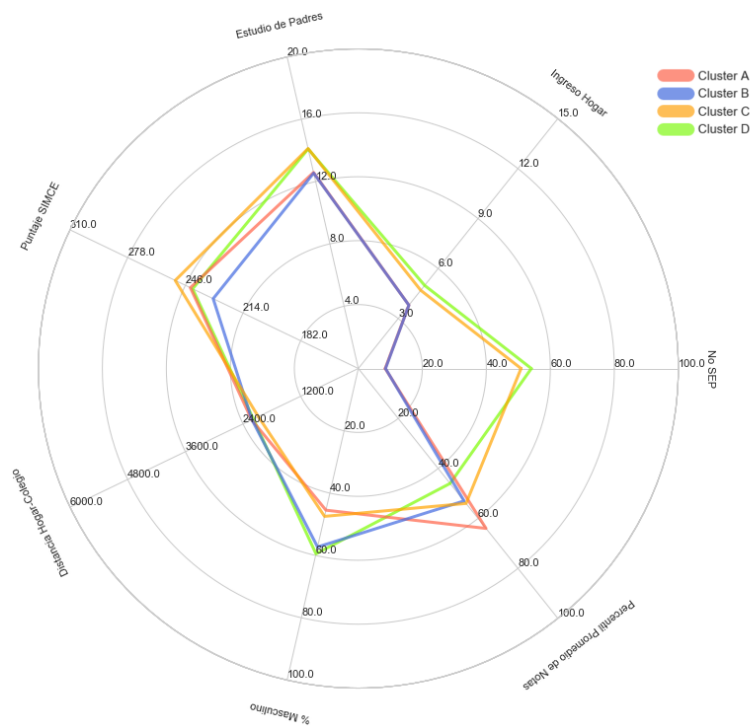


Figura 5.29: Variables a nivel de estudiante.

En este escenario se pueden destacar las siguientes características:

- Los *clusters* C y D sobresalen en las variables Educación de los Padres, Ingreso del Hogar y No Vulnerabilidad. Además, estos mismos *clusters* representan movilidad desde EE particulares subvencionados, con copago mayor en EE de origen, menor vulnerabilidad y mejores resultados SIMCE.
- Los *clusters* A y B son los de mayor vulnerabilidad: menor Ingreso del Hogar, con más

Beneficiarios SEP; asisten a establecimientos de menor Copago, con mayor Vulnerabilidad y con resultados SIMCE menores. Estos son además los *clusters* con establecimientos de origen de dependencia municipal.

- Hay diferencias significativas en la variable de género.
- Se recorren distancias similares en promedio.
- La mayoría de los cambios ocurre desde la dependencia municipal.

SAE Admisión

En este caso hay cuatro *clusters*. A partir de la Figura 5.30 (b), se observa que estos quedan definidos parcialmente por la relación entre dependencias administrativas de origen y destino: el *cluster* A representa la movilidad de estudiantes entre establecimientos municipales (y un pequeño flujo hacia particulares pagados), el *cluster* B aquella desde municipales hacia particulares subvencionados (también un pequeño flujo hacia particulares pagados), el *cluster* C comprende los cambios desde particulares subvencionados hacia municipales, y finalmente, en *cluster* D, incluye los cambios desde establecimientos particulares subvencionados hacia aquellos de la misma dependencia o particulares pagados. De las Figuras 5.30 (b) y (c), se observa que el *cluster* A tiene la mayor representación porcentual de estudiantes del escenario (31%), mientras que los *clusters* B, C y D comprenden el 31%, 13% y 25% de los cambios respectivamente.

Cluster	Copago rbd origen	Copago rbd admitido	IVE rbd origen	IVE rbd admitido	Promedio SIMCE rbd origen	Promedio SIMCE rbd admitido	Dependencia rbd origen	Dependencia rbd admitido
A	P63 (Gratuito)	P63 (Gratuito)	0,81	0,79	P40	P25	Municipal	Municipal
B	P63 (Gratuito)	P64	0,81	0,71	P40	P56	Municipal	Part. Subv.
C		P66 P63 (Gratuito)	0,63	0,75	P49	P41	Part. Subv.	Municipal
D		P66 P69	0,63	0,63	P49	P52	Part. Subv.	Part. Subv.

Tabla 5.17: Centroides de *clusters*.

La Tabla 5.17 y la Figura 5.30 (a) muestra los centroides de cada *cluster*. Se observa que el Copago no cambia en el *cluster* A (movilidad entre gratuitos), aumenta en los *clusters* B y D, en tanto que disminuye en el *cluster* C. Asimismo, la vulnerabilidad (IVE) disminuye en los *clusters* A (levemente) y B, mientras que aumenta en el *cluster* C y se mantiene constante en el *cluster* D. En términos de resultados SIMCE de establecimientos de origen y destino, sólo los *clusters* B y D representan cambios que aumentan en puntajes (cambios hacia particulares subvencionados), en tanto que el rendimiento, medido por el SIMCE, disminuye en los cambios de los *clusters* A y C (cambios hacia municipales).

Al comparar las variables a nivel de estudiante, en la Figura 5.31 se observa, para la variable de vulnerabilidad (No participa en la SEP), que los *clusters* A y B son los "menos no vulnerables" (7% de los estudiantes no son Beneficiarios SEP), es decir, son los más vulnerables, seguidos, en orden decreciente de vulnerabilidad, por los *clusters* C y D (en torno al 50% de los estudiantes no son Beneficiarios SEP). Por otro lado, familias de estudiantes cuyos cambios son caracterizados por todos los *clusters* tienen un Ingreso del Hogar similar entre \$300.000 y \$500.000, siendo el de los *clusters* A y B levemente menor. Por otro lado, los padres de estudiantes de todos los *clusters* presentan en promedio educación hasta tercero

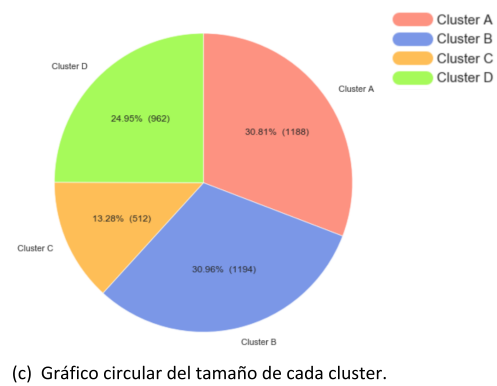
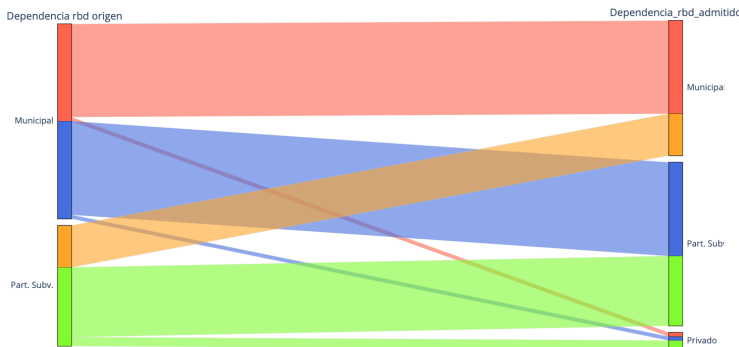
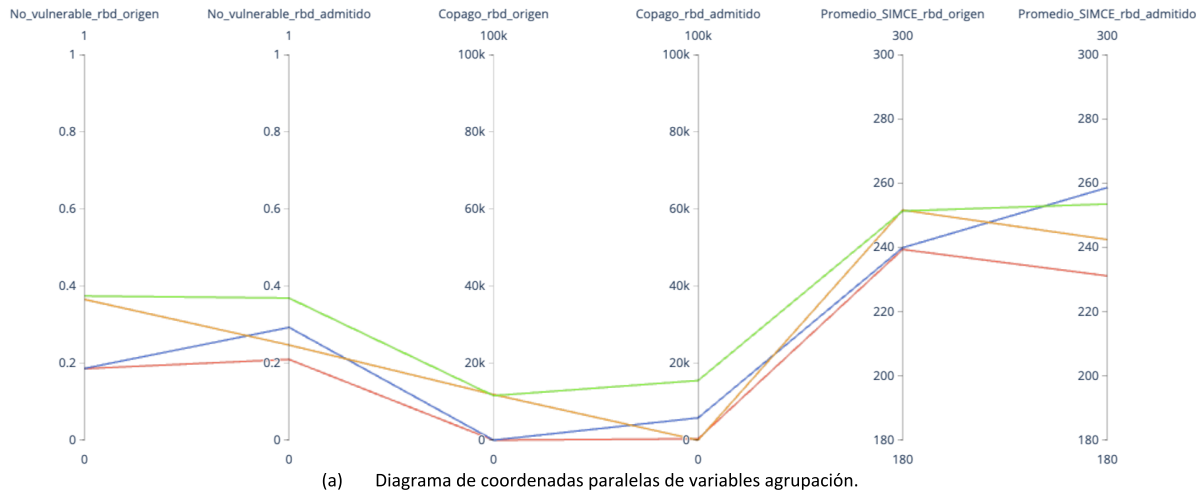


Figura 5.30: Variables a nivel de establecimiento.

medio o enseñanza media completa, siendo nuevamente la de los *clusters* A y B levemente menor.

Los mayores puntajes SIMCE lo obtienen los estudiantes del *cluster* C (253 puntos en promedio), seguidos de cerca por los *clusters* A y D (246 puntos), y terminando con los estudiantes del *cluster* B (alrededor de 235 puntos). Por otro lado, los estudiantes de los todos los *clusters* recorren distancias promedio similares de 2,3 km (Distancia Hogar Colegio). En tanto, se observan diferencias en la variable de género (Porcentaje de Hombres): los *clusters* A y C lo forman 48% de hombres, mientras que los *cluster* B y D lo hacen un 60%. Asimismo, estudiantes del *cluster* A tienen mejor rendimiento académico, del percentil 65 en promedio (Percentil Promedio de Notas), en tanto que los *clusters* B y C se ubican en promedio en torno al percentil 55, y estudiantes del *cluster* D tienen el menor rendimiento (percentil 46 en promedio).

En este escenario se pueden destacar las siguientes características:

- Los *clusters* C y D sobresalen en las variables Educación de los Padres, Ingreso del Hogar y No Vulnerabilidad. Además, estos mismos *clusters* representan movilidad desde EE particulares subvencionados, con copago mayor en EE de origen, menor vulnerabilidad y mejores resultados SIMCE.

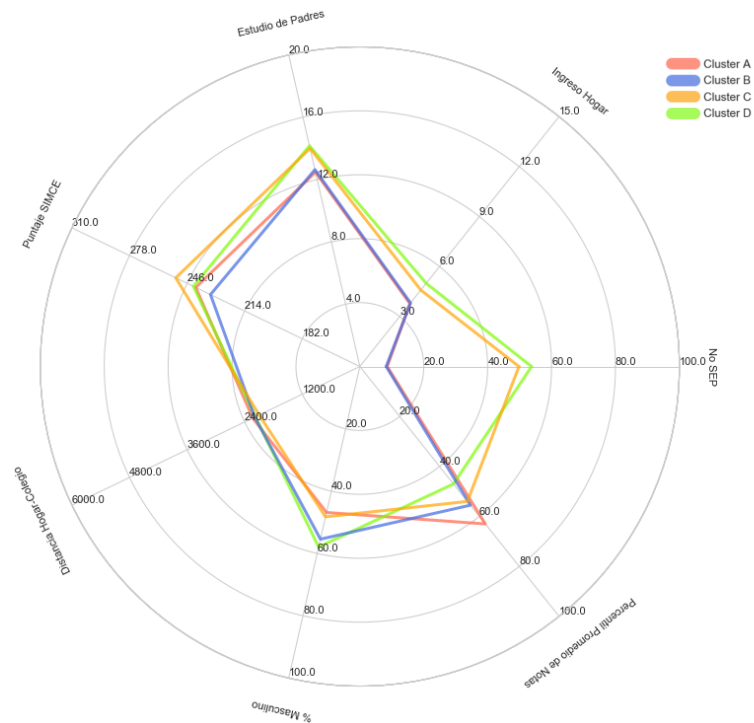


Figura 5.31: Variables a nivel de estudiante.

- Los *clusters* A y B son los de mayor vulnerabilidad: menor Ingreso del Hogar, con más Beneficiarios SEP; asisten a establecimientos de menor Copago, con mayor Vulnerabilidad y con resultados SIMCE bajos. Estos son además los *clusters* con establecimientos de origen de dependencia municipal.
- Hay diferencias significativas la variable de género.
- Se recorren distancias similares en promedio.
- La mayoría de los cambios ocurre desde la dependencia municipal.

Patrones de Cambio Forzado en Rancagua - Machalí

En este caso, se observa que en todos los escenarios las transiciones entre dependencias administrativas definen los *clusters*. Además, se destaca, que todos lo escenarios presentan similar flujo de cambios en cada agrupación, es decir, la representatividad de cada *cluster* es similar en todos los escenarios.

Al observar las variables de establecimientos de origen y destino, se destaca, que en el escenario Pre SAE el copago de estos es mayor. En cuanto a vulnerabilidad (IVE), el escenario SAE Preferencias sobresale con una demanda por establecimientos menos vulnerables, la cual no es cubierta, lo que se observa en el escenario SAE Admisión, en el cual la movilidad según IVE retoma valores similares al escenario Pre SAE.

En términos de resultados SIMCE, establecimiento de preferencia (SAE) tienen mayores puntajes que aquellos tanto del escenario Pre SAE como a los del escenario de Admisión SAE. Es decir, nuevamente se observa una demanda por establecimientos con mayores resultados SIMCE no cubierta.

En lo que respecta a las características de los estudiantes que componen los *clusters*, se observa que los escenarios SAE Preferencia y SAE Admisión muy similares, en tanto que hay pequeñas diferencias con respecto al escenario Pre SAE, en el cual los *clusters* difieren, en promedio, en rendimiento académico.

5.3.2. Cambios No Forzados

5.3.2.1. Escenario Pre SAE

Este escenario hay 5 *clusters*. A partir de la Figura 5.32 (b), se observa que estos quedan parcialmente definidos por la relación entre dependencias administrativas de origen y destino: el *cluster* A representa la movilidad de estudiantes entre municipales, el *cluster* B aquella desde municipales hacia particulares subvencionados, el *cluster* C comprende los cambios desde particulares subvencionados hacia municipales, el *cluster* D agrupa cambios entre establecimientos particulares subvencionados, y el *cluster* E representa algunos cambios entre particulares subvencionados, además de aquellos que involucran establecimientos particulares pagados. De las Figuras 5.32 (b) y (c), se observa que los *clusters* C y D tienen la mayor representación porcentual de estudiantes del escenario (33% y 34% respectivamente), mientras que los *clusters* A, B y E comprenden el 16%, 8% y 9% de los cambios, respectivamente.

La Tabla 5.18 y la Figura 5.32 (a) muestran los centroides de cada *cluster*. Se observa que el Copago disminuye sólo en el *cluster* C, es constante en los *clusters* A, D y E, y aumenta en el *cluster* B. Por otra parte, la vulnerabilidad (IVE) disminuye en los *clusters* A, B y E, aumenta en el *cluster* C, en tanto que permanece constante en el *cluster* D. Por otro lado, en términos de resultados SIMCE de EE de origen y destino, en los *clusters* A, B, C y D hay una mejora en rendimiento medido por el SIMCE, mientras que en el *cluster* E los puntajes de establecimientos de origen y destino son similares.

Cluster	Copago rbd origen	Copago rbd destino	IVE rbd origen	IVE rbd destino	Promedio SIMCE rbd origen	Promedio SIMCE rbd destino	Dependencia rbd origen	Dependencia rbd destino
A	P60 (Gratuito)	P61 (Gratuito)	0,74	0,64	P43	P67	Municipal	Municipal
B	P60 (Gratuito)	P73	0,76	0,63	P34	P53	Municipal	Part. Subv.
C	P77	P60 (Gratuito)	0,57	0,61	P53	P70	Part. Subv.	Municipal
D	P78	P79	0,58	0,58	P52	P59	Part. Subv.	Part. Subv.
E	P93	P93	0,17	0,13	P74	P74	Subv./Priv.	Privado

Tabla 5.18: Centroides de *clusters*.

Al comparar las variables a nivel de estudiante, en la Figura 5.33 se observa, para la variable de vulnerabilidad (No participa en la SEP), que los *clusters* A y B son los "menos no vulnerables" (en torno al 40% de los estudiantes no son Beneficiarios SEP), es decir, son los más vulnerable, seguidos, en orden decreciente de vulnerabilidad, por los *clusters* C y D (70% de los estudiantes no son Beneficiarios SEP); y finalmente, el *cluster* E tiene una muy baja vulnerabilidad (99% de los estudiantes no son Beneficiarios SEP).

Con respecto al Ingreso del Hogar, los *clusters* A, B, C y D representan ingresos entre \$300.000 y \$500.000, en tanto que el *cluster* E representa los mayores ingresos, los que figuran entre \$1.200.000 y \$1.400.000 en promedio. Por otro lado, para la variable Educación de los

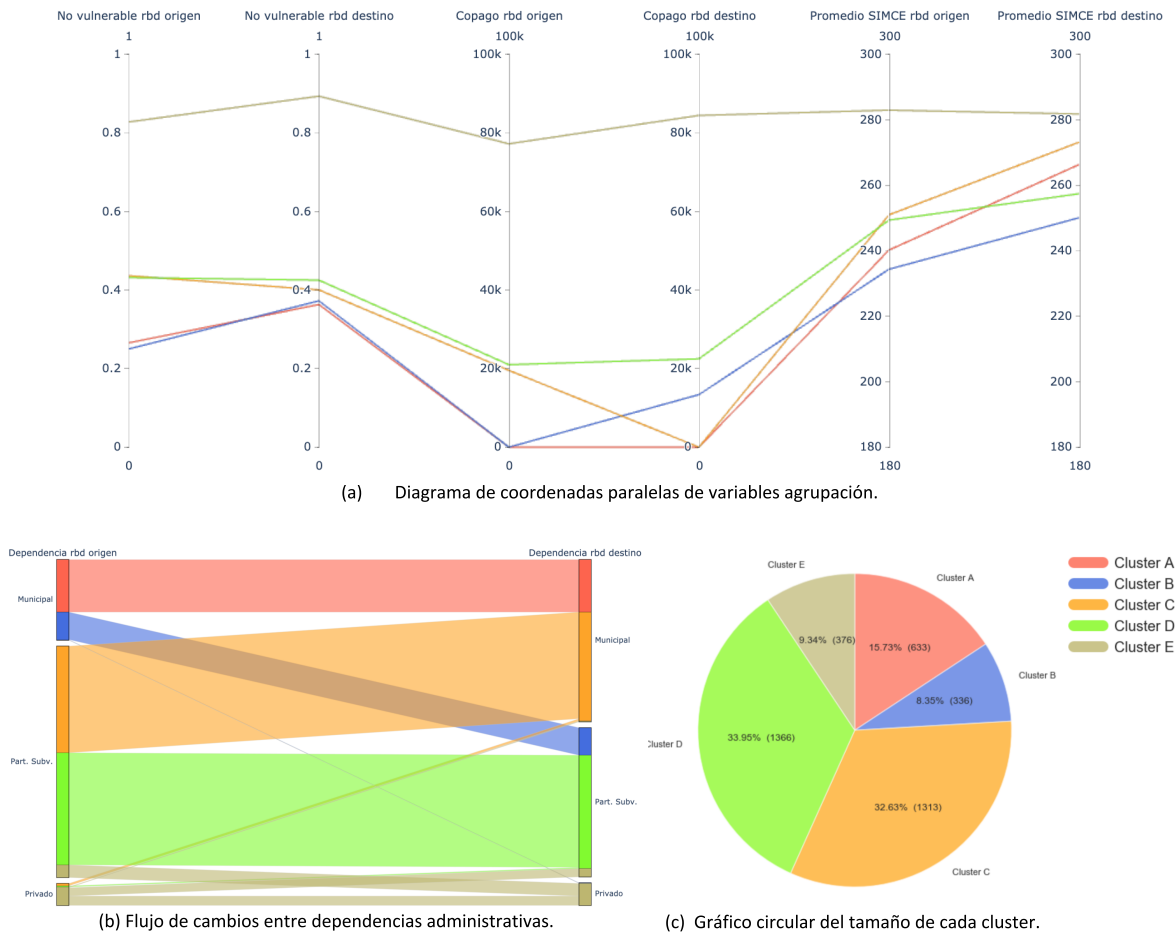


Figura 5.32: Variables a nivel de establecimiento.

Padres, los *clusters* A, B, C y D presentan en promedio educación media completa, en tanto que padres del *cluster* E poseen, en general, educación universitaria incompleta.

Los estudiantes de los *clusters* B y D obtienen en promedio puntajes SIMCE en torno a 240 puntos, aquellos de los *clusters* A y C obtiene en promedio 270 puntos, y los del *cluster* E superan al resto, con un rendimiento promedio de 280 puntos. Asimismo, los estudiantes de los *clusters* B, C y D recorren distancias (Distancia Hogar Colegio) entre 2,6 y 3,2 km en promedio, seguidos por el *cluster* A (4 km en promedio) y finalmente por estudiantes del *cluster* E (recorren 4,8 km en promedio).

Con respecto a la variable de género (Porcentaje de Hombres), los *clusters* presentan una composición disímil: el *cluster* A está compuesto por un 37% de hombres, el *cluster* E por un 45%, los *clusters* B y C por un 50%, y finalmente el *cluster* D lo componen un 55% de hombres. Por otro lado, también se observan diferencias en el rendimiento académico de los estudiantes (Percentil Promedio de Notas): estudiantes de los *clusters* B, D y E se ubican entre los percentiles 40 y 50, mientras que los *clusters* A y C tienen un alto desempeño académico, ubicándose en el percentil 78 en promedio.

En este escenario se pueden destacar las siguientes características:

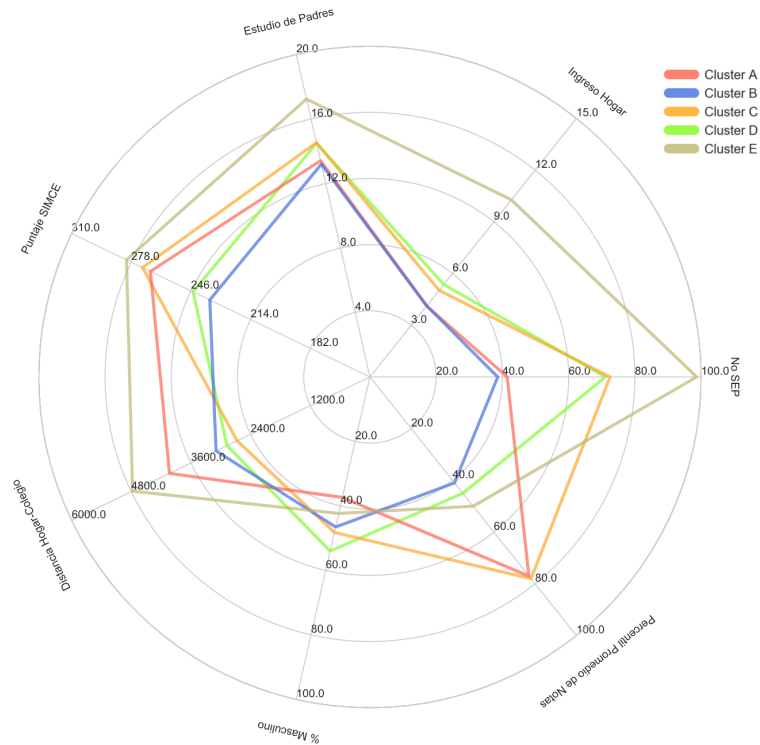


Figura 5.33: Variables a nivel de estudiante.

- Los *clusters* ya no están totalmente ordenados de acuerdo a la transición entre dependencia administrativa, lo que sugiere un mayor desorden en la organización de la movilidad en el escenario no forzado.
- La mayoría de los cambios ocurren desde escuelas subvencionadas (*clusters* C, D y E) y la minoría desde particulares pagados (*cluster* E).
- La vulnerabilidad del *cluster* E es menor: menor IVE, menos beneficiarios SEP, mayor Copago, mayor rendimiento, medido por el SIMCE del establecimiento, mayor Ingreso del Hogar, mayor Educación de los Padres y mejores puntajes SIMCE (estudiantes).
- Los establecimientos particulares pagados tienen una mayor participación que en otros escenarios.
- Se observa un sesgo de género importante en el *cluster* A
- Cambios hacia municipales (*clusters* A y C) sobresalen fuertemente en desempeño académico (Promedio de Notas).

5.3.2.2. Escenario SAE

SAE Preferencias

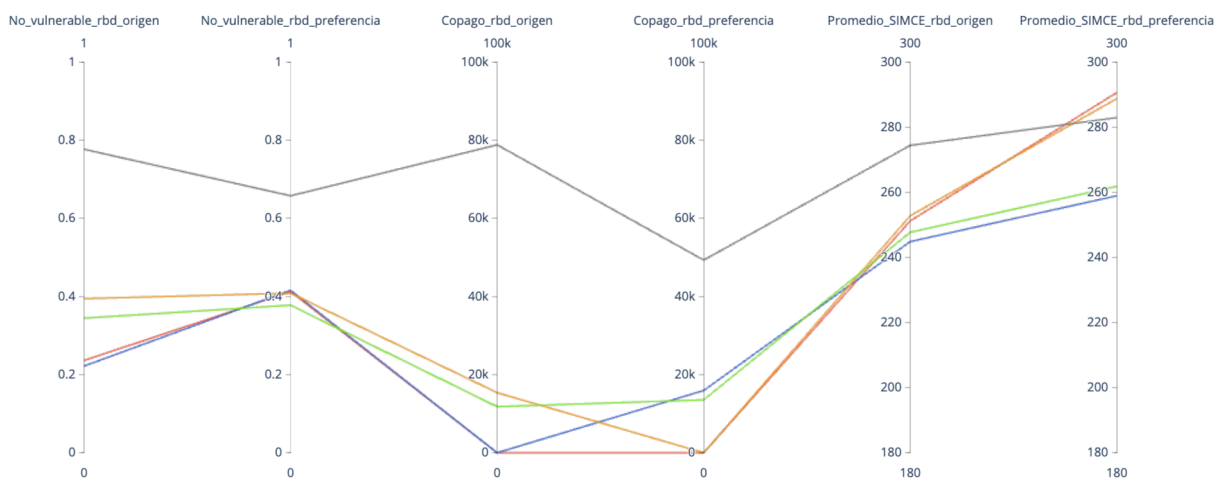
Este escenario hay 5 *clusters*. A partir de la Figura 5.34 (b), se observa que estos quedan parcialmente definidos por la relación entre dependencias administrativas de origen y destino: el *cluster* A representa la movilidad de estudiantes entre municipales, el *cluster* B aquella desde municipales hacia particulares subvencionados, el *cluster* C comprende los cambios

desde particulares subvencionados hacia municipales, el *cluster* D agrupa cambios entre establecimientos particulares subvencionados, y el *cluster* E representa cambios desde y hacia establecimientos particulares pagados. De las Figuras 5.34 (b) y (c), se observa que el *cluster* C tiene la mayor representación porcentual de estudiantes del escenario (37%), mientras que los *clusters* A, B, D y E comprenden el 22 %, 7 %, 27 % y 8 % de los cambios, respectivamente.

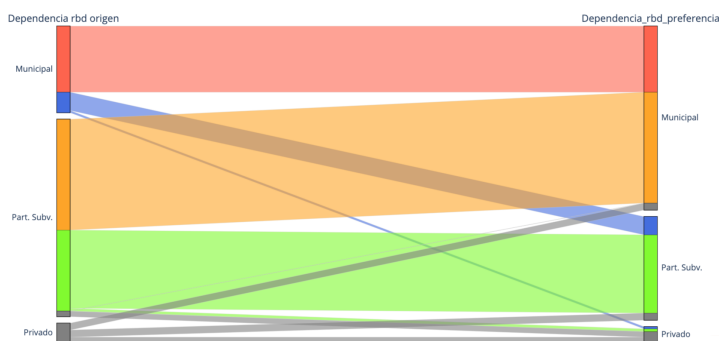
Cluster	Copago rbd origen	Copago rbd preferencia	IVE rbd origen	IVE rbd preferencia	Promedio SIMCE rbd origen	Promedio SIMCE rbd preferencia	Dependencia rbd origen	Dependencia rbd preferencia
A	P63 (Gratis)	P63 (Gratis)	0,76	0,59	P49	P75	Municipal	Municipal
B	P63 (Gratis)	P69	0,78	0,58	P42	P56	Municipal	Part. Subv.
C	P69	P63 (Gratis)	0,61	0,59	P50	P74	Part. Subv.	Municipal
D	P66	P67	0,66	0,62	P46	P61	Part. Subv.	Part. Subv.
E	P90	P81	0,22	0,34	P68	P74	Privado	Todos

Tabla 5.19: Centroides de *clusters*.

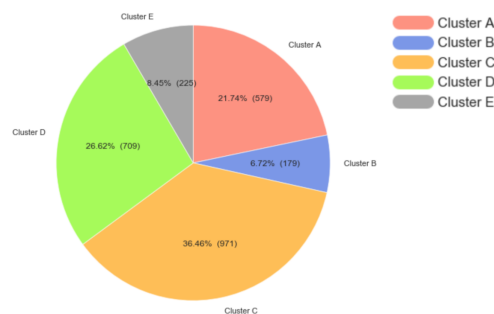
La Tabla 5.19 y la Figura 5.34 (a) muestran los centroides de cada *cluster*. Se observa que el Copago disminuye sólo en los *clusters* C y E, es constante en los *clusters* A y D, y aumenta sólo en el *cluster* B. Por otra parte, la vulnerabilidad (IVE) disminuye en los *clusters* A, B, C y D, en tanto que aumenta en el *cluster* E. Por otro lado, en términos de resultados SIMCE de EE de origen y destino, en todos los *clusters* hay una mejora en rendimiento, medido por el SIMCE,.



(a) Diagrama de coordenadas paralelas de variables agrupación.



(b) Flujo de cambios entre dependencias administrativas.



(c) Gráfico circular del tamaño de cada cluster.

Figura 5.34: Variables a nivel de establecimiento.

Al comparar las variables a nivel de estudiante, en la Figura 5.35 se observa, para la variable de vulnerabilidad (No participa en la SEP), que los *clusters* A y B son los "menos no vulnerables" (en torno al 15% de los estudiantes no son Beneficiarios SEP), es decir, son los más vulnerable, seguidos, en orden decreciente de vulnerabilidad, por los *clusters* C (40% de los estudiantes no son Beneficiarios SEP) y D (52% de los estudiantes no son Beneficiarios SEP); y finalmente, el *cluster* E tiene una muy baja vulnerabilidad (99% de los estudiantes no son Beneficiarios SEP).

Con respecto al Ingreso del Hogar, los *clusters* A, B, C y D representan ingresos entre \$300.000 y \$500.000, en tanto que el *cluster* E representa los mayores ingresos, los que figuran entre \$1.200.000 y \$1.400.000 en promedio. Por otro lado, para la variable Educación de los Padres, los *clusters* A, B, C y D presentan en promedio educación media completa, en tanto que padres del *cluster* E poseen, en general, educación universitaria incompleta.

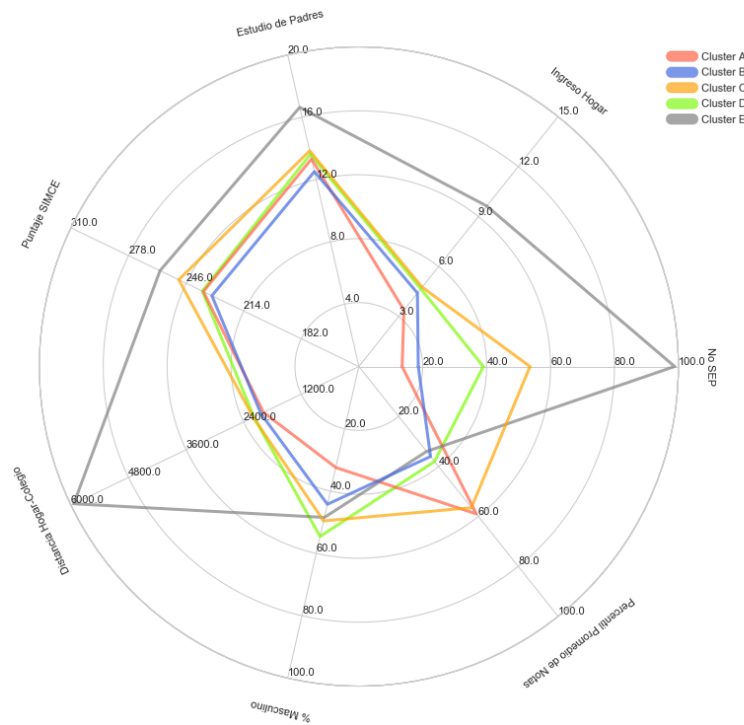


Figura 5.35: Variables a nivel de estudiante.

Los estudiantes de los *clusters* A, B y D obtienen en promedio puntajes SIMCE en torno a 235 puntos, aquellos de los *clusters* C obtiene en promedio 250 puntos, y los del *cluster* E superan al resto, con un rendimiento promedio de 260 puntos. Asimismo, los estudiantes de los *clusters* A, B, C y D recorren distancias (Distancia Hogar Colegio) entre 2 y 2,4 km en promedio. Datos de distancia para el *cluster* E poseen una única observación (6 km) , por lo que se considera inválido el resultado promedio que entrega.

Con respecto a la variable de género (Porcentaje de Hombres), los *clusters* presentan una composición disímil: el *cluster* A está compuesto por un 35% de hombres, en tanto que los *clusters* B, C, D y E lo forman entre 45% y 55% de hombres. Por otro lado, se observan importantes diferencias en el rendimiento académico de los estudiantes (Percentil Promedio de Notas): estudiantes de los *clusters* B, D y E se ubican entre los percentiles 35 y 40, mientras

que aquellos de los *clusters* A y C tienen un mayor desempeño académico, ubicándose en el percentil 60 en promedio.

En este escenario se pueden destacar las siguientes características:

- Los *clusters* ya no están totalmente ordenados de acuerdo a la transición entre dependencia administrativa, lo que sugiere un mayor desorden en la organización de la movilidad en el escenario no forzado.
- La mayoría de los cambios ocurren desde escuelas subvencionadas (*clusters* C, D y E) y la minoría desde particulares pagados (*cluster* E).
- La vulnerabilidad del *cluster* E es menor: menor IVE, menos beneficiarios SEP, mayor Copago, mayor rendimiento, medido por el SIMCE de establecimiento, mayor Ingreso del Hogar, mayor Educación de los Padres y mejores puntajes SIMCE (estudiantes), aunque este presente un flujo hacia establecimientos municipales.
- Los establecimientos particulares pagados tienen una mayor participación que en otros escenarios.
- Se observa un sesgo de género importante en el *cluster* A.
- Cambios hacia municipales (*clusters* A y C) sobresalen en desempeño académico (Promedio de Notas).

SAE Admisión

Este escenario hay 4 *clusters*. A partir de la Figura 5.36 (b), se observa que estos quedan parcialmente definidos por la relación entre dependencias administrativas de origen y destino: el *cluster* A representa la movilidad de estudiantes desde municipales o particulares pagados hacia particulares subvencionados, el *cluster* B aquella entre municipales, el *cluster* C comprende los cambios desde particulares subvencionados o particulares pagados hacia municipales, y el *cluster* D agrupa cambios desde establecimientos particulares subvencionados, hacia particulares subvencionados o particulares pagados. De las Figuras 5.36 (b) y (c), se observa que el *cluster* D tiene la mayor representación porcentual de estudiantes del escenario (40 %), mientras que los *clusters* A, B y C comprenden el 12 %, 21 % y 40 % de los cambios, respectivamente.

Cluster	Copago rbd origen	Copago rbd admitido	IVE rbd origen	IVE rbd admitido	Promedio SIMCE rbd origen	Promedio SIMCE rbd admitido	Dependencia rbd origen	Dependencia rbd admitido
A	P75	P75	0,54	0,5	P55	P62	Muni./Priv.	Subv./Priv.
B	P63 (Gratuito)	P63 (Gratuito)	0,76	0,67	P49	P66	Municipal	Municipal
C	P71	P63 (Gratuito)	0,59	0,65	P52	P68	Part. Subv.	Municipal
D	P69	P70	0,62	0,61	P48	P53	Part. Subv.	Part. Subv.

Tabla 5.20: Centroides de *clusters*.

La Tabla 5.20 y la Figura 5.36 (a) muestran los centroides de cada *cluster*. Se observa que el Copago disminuye sólo en el *cluster* C, en tanto que es constante en los otros tres. Por otra parte, la vulnerabilidad (IVE) disminuye en los *clusters* A, B y D, en tanto que aumenta en el *cluster* C. Asimismo, en términos de resultados SIMCE de EE de origen y destino, en todos los *clusters* hay una mejora en rendimiento.

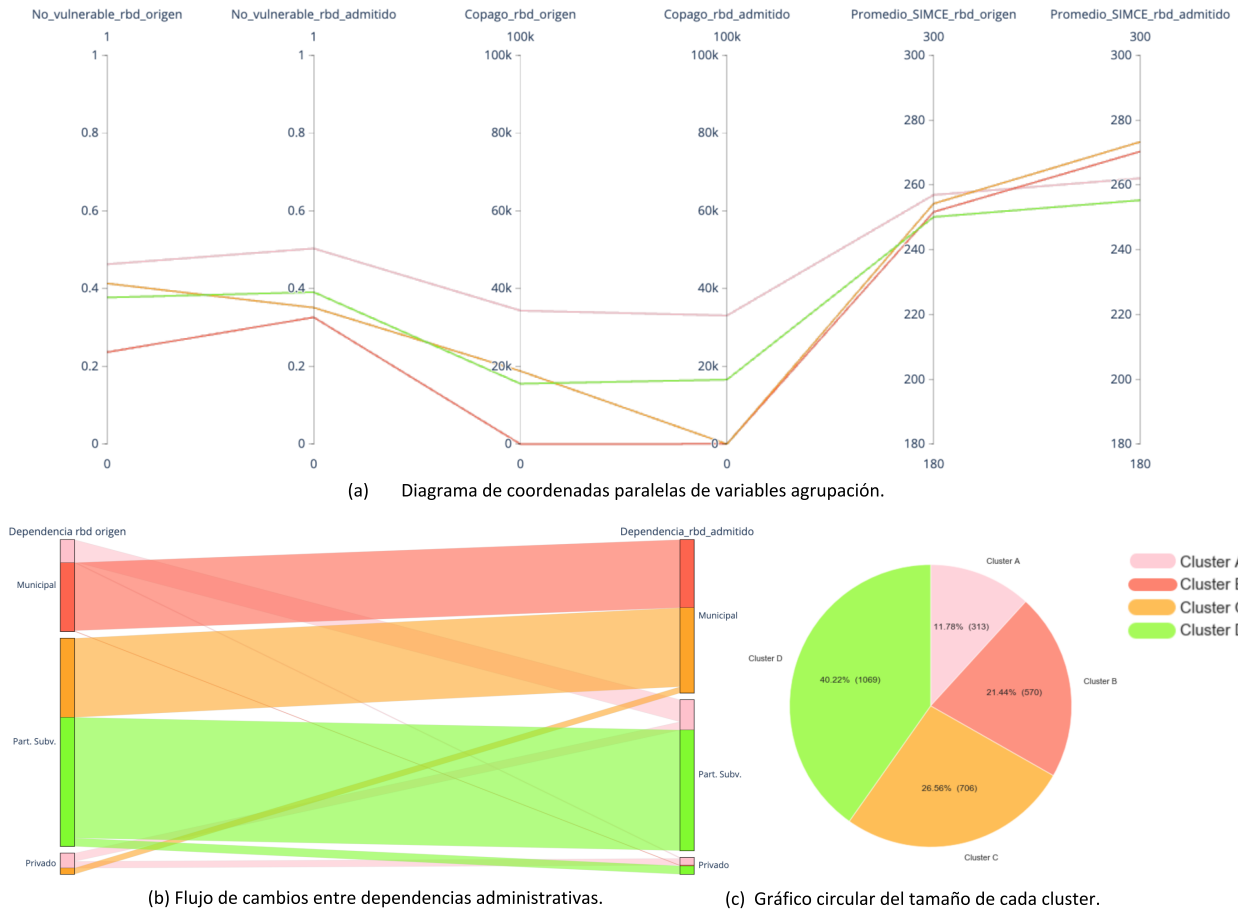


Figura 5.36: Variables a nivel de establecimiento.

Al comparar las variables a nivel de estudiante, en la Figura 5.37 se observa, para la variable de vulnerabilidad (No participa en la SEP), que el *clusters* A es el "menos no vulnerables" (13% de los estudiantes no son Beneficiarios SEP), es decir, son los más vulnerable, seguidos, en orden decreciente de vulnerabilidad, por los *clusters* D, A y C (entre 45% y 55% de los estudiantes no son Beneficiarios SEP).

Con respecto al Ingreso del Hogar, el *cluster* B representa los menores ingresos entre \$300.000 y \$400.000, seguido por los *clusters* C y D, con un ingreso promedio entre \$400.000 y \$500.000, en tanto que el *cluster* A representa los mayores ingresos, que figuran entre \$800.000 y \$1.000.000 en promedio. Por otro lado, para la variable Educación de los Padres, los *clusters* A, B y C presentan en promedio educación media completa, en tanto que padres del *cluster* A poseen, en general, educación incompleta en CFT o IP.

Los estudiantes de todos los *clusters* obtienen en promedio puntajes SIMCE en torno a 245 puntos. Asimismo, todos los estudiantes recorren distancias (Distancia Hogar Colegio) entorno a los 2,2 km en promedio.

Con respecto a la variable de género (Porcentaje de Hombres), los *clusters* presentan una composición disímil: el *cluster* B está compuesto por un 32% de hombres, en tanto que los *clusters* A, C y D lo forman entre 45% y 55% de hombres. Por otro lado, se observan

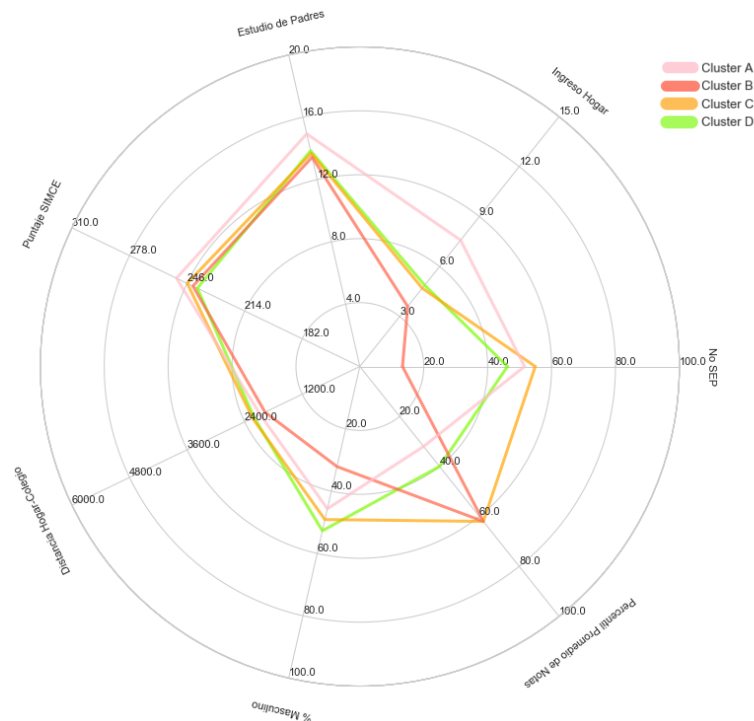


Figura 5.37: Variables a nivel de estudiante.

importantes diferencias en el rendimiento académico de los estudiantes (Percentil Promedio de Notas): estudiantes de los *clusters* A y D se ubican entre los percentiles 35 y 40, mientras que aquellos de los *clusters* B y C tienen un mayor desempeño académico, ubicándose en el percentil 60 en promedio.

En este escenario se pueden destacar las siguientes características:

- Los *clusters* ya no están totalmente ordenados de acuerdo a la transición entre dependencia administrativa, lo que sugiere un mayor desorden en la organización de la movilidad en el escenario no forzado.
- La mayoría de los cambios ocurren desde escuelas subvencionadas (*clusters* C y D) y la minoría desde particulares pagados.
- Los establecimientos particulares pagados tienen una mayor participación que en otros escenarios.
- Se observa un sesgo de género importante en el *cluster* B.
- Cambios hacia municipales (*clusters* B y C) sobresalen en desempeño académico (Promedio de Notas).
- El *cluster* B resalta levemente en no vulnerabilidad, aunque representa cambios entre establecimiento de puntajes SIMCE no mayores que el resto de las agrupaciones.

Patrones de Cambio No Forzado en Rancagua - Machalí

Al igual que en otros casos No Forzados, los *clusters* ya no están definidos completamente por las transiciones entre dependencias administrativas. Además, cada escenario tiene un *cluster* poco estructurado en términos de su coeficiente de silueta. Asimismo, establecimientos

particulares pagados aumentan su representatividad en cada escenario.

En los tres escenarios sobresale un *cluster* que excede al resto en diversos atributos: menor IVE, mayor copago, mejor rendimiento, medido por el SIMCE, etc. Estos son los *cluster* de estudiantes de GSE alto, cuya representatividad y composición es similar en los escenarios Pre SAE y SAE Preferencias. Mientras en el escenario restante, este perfil se combina con cambios desde municipales a particulares subvencionados, lo que genera una disminución en los atributos señalados, aunque continúan siendo mayores al resto de los *clusters*.

Se observa también que los resultados SIMCE de establecimiento de preferencias (SAE) son mayores a que en los escenarios Pre SAE y SAE Admisión. Demanda por de establecimientos de mayor SIMCE no cubierta, sobre todo de EE municipales. Esto último llama la atención, pues los *clusters* A y C en escenario SAE Preferencias, demandan establecimientos municipales con elevados resultados SIMCE, lo que el sistema no es capaz de proveer, según se observa en SAE Admisión.

Se destaca, además, que el *cluster* de cambio de estudiantes de GSE alto presente en los escenarios Pre SAE y SAE Preferencias, deja de estar presente en SAE Admisión, lo que indica que este perfil de cambio ya no está lo suficientemente estructurado para formar su propia agrupación. En este sentido, se plantea que el SAE fuerza, en un cierto sentido, una mayor homogeneidad en los perfiles de cambio de escuela.

En términos de características de estudiantes, en el caso Pre SAE resaltan estudiantes de los *clusters* A y C (movilidad hacia municipales), quienes obtienen mejores resultados SIMCE y mejor rendimiento académico, este último (*cluster* C) resaltando aún más en los escenarios SAE. Por otro lado, en el escenario SAE Admisión se observa que los estudiantes del *cluster* A son tienen un bajo rendimiento académico y sobresalen en variables socioeconómicas (Estudio de Padre e Ingreso del Hogar).

Patrones de Cambio en Rancagua - Machalí

Al comparar todos los escenarios y casos correspondiente a esta aglomeración urbana, se observa claramente una mayor estructura en el caso Forzado: tanto en el número de *clusters* como en su caracterización según transiciones de dependencia administrativa. En este mismo sentido, sólo en el caso No Forzado aparecen *clusters* de estudiantes de GSE alto.

Al contrastar los tipos de cambio Forzado y No Forzado, se destaca, que en el segundo la proporción de cambios desde la dependencia municipal es mucho menor, esto pues hay muchos establecimientos municipales que entregan educación sólo hasta los niveles estudiados, y por ende, caen en la categoría de cambios Forzados.

Para los casos SAE, en ambos tipos de cambio hay una demanda por resultados SIMCE que no es completamente cubierta. Sin embargo, la Admisión SAE nuestra, que en promedio, los puntajes de establecimientos de destino son superiores a aquellos que obtenían los establecimientos que recibían estudiantes previo al SAE. La excepción a esto la marcan los *clusters* A y C en el caso Forzado (movilidad hacia municipales), en que los puntajes de establecimientos de destino en el escenario SAE Admisión son inferiores a aquellos de los mismos *cluster* del escenario Pre SAE.

5.4. Puerto Montt - Puerto Varas

5.4.1. Cambios Forzados

5.4.1.1. Escenario Pre SAE

Este escenario hay 4 *clusters*. A partir de la Figura 5.38 (b), se observa que estos quedan bien definidos por la relación entre dependencias administrativas de origen y destino: el *cluster A* representa la movilidad de estudiantes entre municipales, el *cluster B* aquella desde municipales hacia particulares subvencionados, el *cluster C* comprende los cambios desde particulares subvencionados hacia municipales, y el *cluster D* agrupa cambios entre establecimientos particulares subvencionados. De las Figuras 5.38 (b) y (c), se observa que el *cluster A* tiene la mayor representación porcentual de estudiantes del escenario (58%), seguido por el *cluster B* (26%), mientras que los *clusters C* y *D* comprenden el 5% y 11% de los cambios, respectivamente.

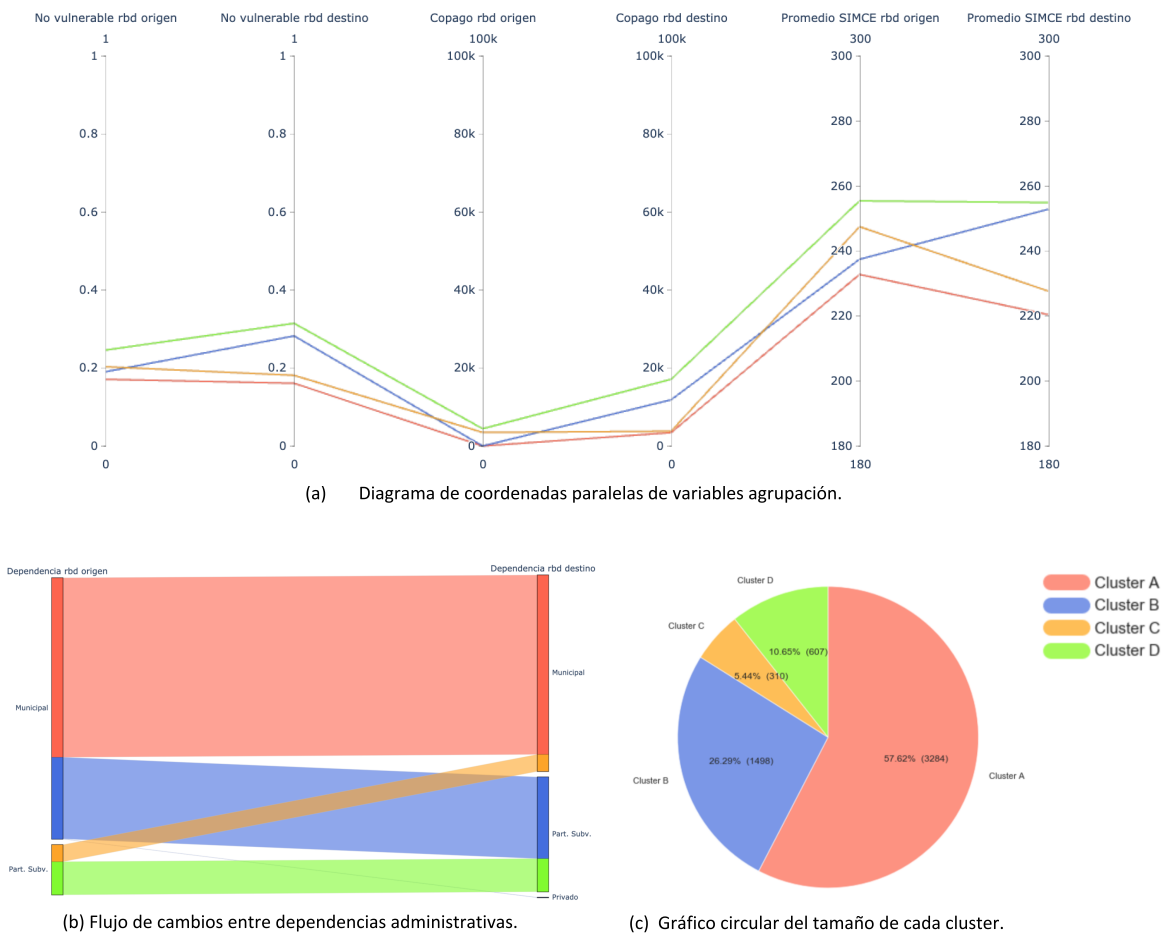


Figura 5.38: Variables a nivel de establecimiento.

La Tabla 5.21 y la Figura 5.38 (a) muestran los centroides de cada *cluster*. Se observa que el Copago disminuye sólo en el *cluster C*, es constante en el *cluster A*, y aumenta en los

clusters B y D. Por otra parte, la vulnerabilidad (IVE) disminuye en los clusters B y D, en tanto que aumenta en los clusters A y C. Por otro lado, en términos de resultados SIMCE de EE de origen y destino, en el cluster B hay una mejora en rendimiento, mientras que en los clusters A, C y D los puntajes disminuyen.

Cluster	Copago rbd origen	Copago rbd destino	IVE rbd origen	IVE rbd destino	Promedio SIMCE rbd origen	Promedio SIMCE rbd destino	Dependencia rbd origen	Dependencia rbd destino
A	P60 (Gratis)	P64	0,83	0,84	P31	P14	Municipal	Municipal
B	P60 (Gratis)	P72	0,81	0,71	P39	P56	Municipal	Part. Subv.
C	P64	P65	0,8	0,82	P51	P23	Part. Subv.	Municipal
D	P65	P76	0,76	0,69	P58	P57	Part. Subv.	Part. Subv.

Tabla 5.21: Centroides de clusters.

Al comparar las variables a nivel de estudiante, en la Figura 5.39 se observa, para la variable de vulnerabilidad (No participa en la SEP), que los clusters A y C son los "menos vulnerables" (25% de los estudiantes no son Beneficiarios SEP), es decir, son los más vulnerables, seguidos, en orden decreciente de vulnerabilidad, por los clusters B y D (33% y 36% de los estudiantes no son Beneficiarios SEP, respectivamente).

Con respecto al Ingreso del Hogar, todos los clusters representan ingresos entre \$200.000 y \$400.000, siendo el de los clusters A y C levemente menores. Por otro lado, para la variable Educación de los Padres, todos los clusters representan en promedio entre segundo medio y educación media completa.

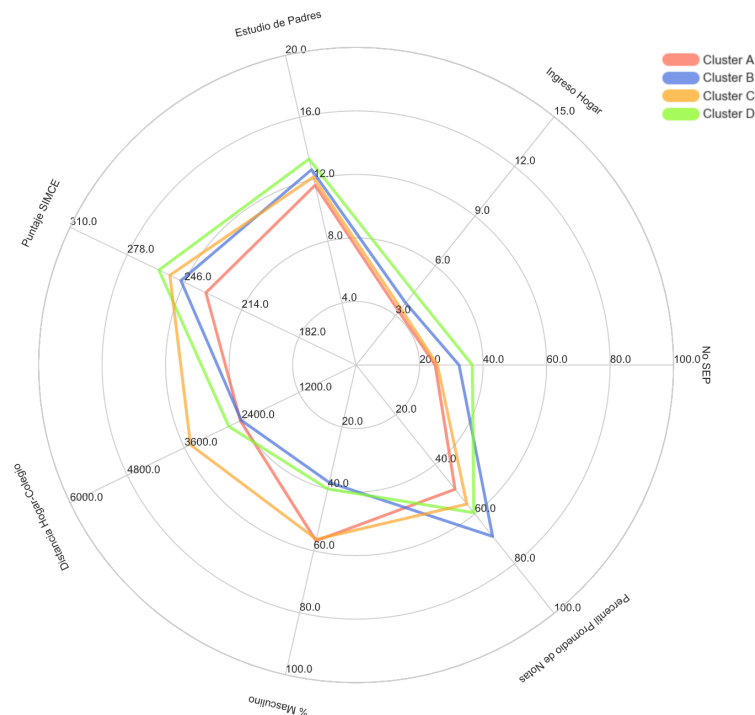


Figura 5.39: Variables a nivel de estudiante.

Los estudiantes del cluster D obtienen los mayores puntajes, de 260 puntos en promedio, seguidos por el cluster C (252 puntos en promedio), luego por el cluster B (246 puntos en promedio), y finalmente por el cluster A (232 puntos en promedio). Asimismo, los estudiantes

de los *clusters* A, B y D recorren distancias (Distancia Hogar Colegio) entre 2,4 y 2,6 km en promedio, mientras que los estudiantes del *cluster* C recorren mayores distancias de 3,5 km en promedio.

Con respecto a la variable de género (Porcentaje de Hombres), los *clusters* presentan una composición disímil: los *clusters* A y C están compuesto por un 55 % de hombres, mientras que los *clusters* B y D por un 39 %. Por otro lado, también se observan diferencias en rendimiento académico (Percentil Promedio de Notas): estudiantes del *cluster* A tienen el rendimiento menor (percentil 48), seguidos del *cluster* C (percentil 55), luego del *cluster* D (percentil 60), finalmente por el *cluster* B (percentil 68), que demuestra el mayor rendimiento en promedio.

En este escenario se pueden destacar las siguientes características:

- Los *clusters* están totalmente ordenados de acuerdo a la transición entre dependencia administrativa.
- La mayoría de los cambios ocurren desde escuelas municipales, mientras que la minoría lo hacen desde particular subvencionado hacia municipal.
- No se observan elevadas diferencias en IVE, beneficiarios SEP, Ingreso del Hogar y Educación de los Padres.
- Se observa un sesgo de género importante en los *clusters* B y D (cambios hacia particulares subvencionados).
- Cambios hacia municipales (*clusters* A y C) sobresalen fuertemente en desempeño académico (Promedio de Notas).
- En general el copago de los establecimientos es bajo.
- Movilidad hacia particulares subvencionados mejora en rendimiento, medido por el SIMCE, mientras que movilidad hacia municipales empeora en este valor.

5.4.1.2. Escenario SAE

SAE Preferencias

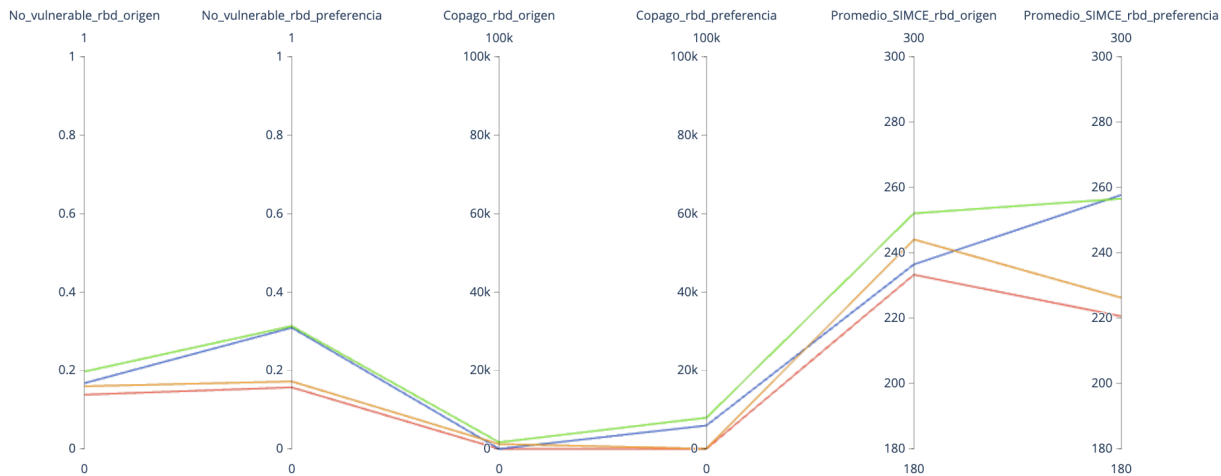
Este escenario hay 4 *clusters*. A partir de la Figura 5.40 (b), se observa que estos quedan bien definidos por la relación entre dependencias administrativas de origen y destino: el *cluster* A representa la movilidad de estudiantes entre municipales, el *cluster* B aquella desde municipales hacia particulares subvencionados, el *cluster* C comprende los cambios desde particulares subvencionados hacia municipales, y el *cluster* D agrupa cambios entre establecimientos particulares subvencionados. Sin embargo, también hay pequeños flujos desde todos los *clusters* que llegan a establecimientos particulares pagados. De las Figuras 5.40 (b) y (c), se observa que el *cluster* B tiene la mayor representación porcentual de estudiantes del escenario (45 %), seguido por el *cluster* A (34 %), mientras que los *clusters* C y D comprenden el 6 % y 15 % de los cambios, respectivamente.

La Tabla 5.22 y la Figura 5.40 (a) muestran los centroides de cada *cluster*. Se observa que el Copago es constante en los *clusters* A y C, y aumenta en los *clusters* B y D. Por otra parte, la vulnerabilidad (IVE) disminuye en los *clusters* B y D, en tanto que se mantiene

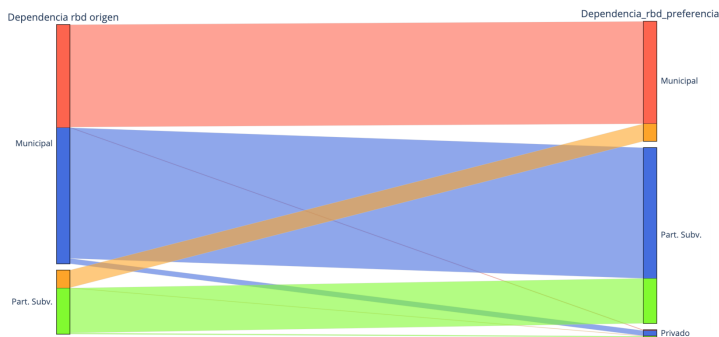
Cluster	Copago		IVE		Promedio SIMCE		Dependencia	
	rbd origen	rbd preferencia	rbd origen	rbd preferencia	rbd origen	rbd preferencia	rbd origen	rbd preferencia
A	P83 (Gratuito)	P83 (Gratuito)	0,86	0,84	P35	P14	Municipal	Municipal
B	P83 (Gratuito)	P84	0,83	0,69	P39	P67	Municipal	Part. Subv.
C	P83 (Gratuito)	P83 (Gratuito)	0,84	0,83	P52	P26	Part. Subv.	Municipal
D	P83 (Gratuito)	P85	0,8	0,69	P64	P67	Part. Subv.	Part. Subv.

Tabla 5.22: Centroides de *clusters*.

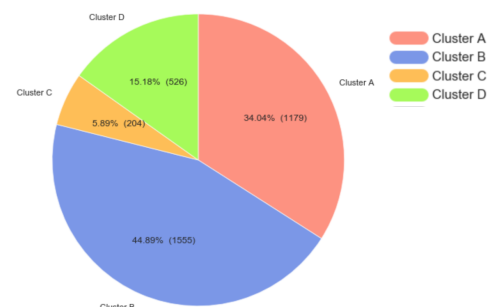
constante en los *clusters* A y C. Por otro lado, en términos de resultados SIMCE de EE de origen y destino, en los *clusters* B y D hay una mejora en rendimiento, medido por el SIMCE, mientras que en los *clusters* A y C los puntajes disminuyen.



(a) Diagrama de coordenadas paralelas de variables agrupación.



(b) Flujo de cambios entre dependencias administrativas.



(c) Gráfico circular del tamaño de cada cluster.

Figura 5.40: Variables a nivel de establecimiento.

Al comparar las variables a nivel de estudiante, en la Figura 5.41 se observa, para la variable de vulnerabilidad (No participa en la SEP), que todos los *clusters* presentan bajos valores de "no vulnerabilidad" (menos del 13% de los estudiantes no son Beneficiarios SEP), es decir, todos tienen una alta composición de beneficiarios SEP. Asimismo, con respecto al Ingreso del Hogar, todos los *clusters* representan ingresos entre \$300.000 y \$400.000. Por otro lado, para la variable Educación de los Padres, todos los *clusters* representan en promedio educación hasta segundo o tercero medio.

Los estudiantes del *cluster* D obtienen los mayores puntajes, de 250 puntos en promedio, seguidos por los *clusters* C y B (237 puntos en promedio), y finalmente por el *cluster* A (228

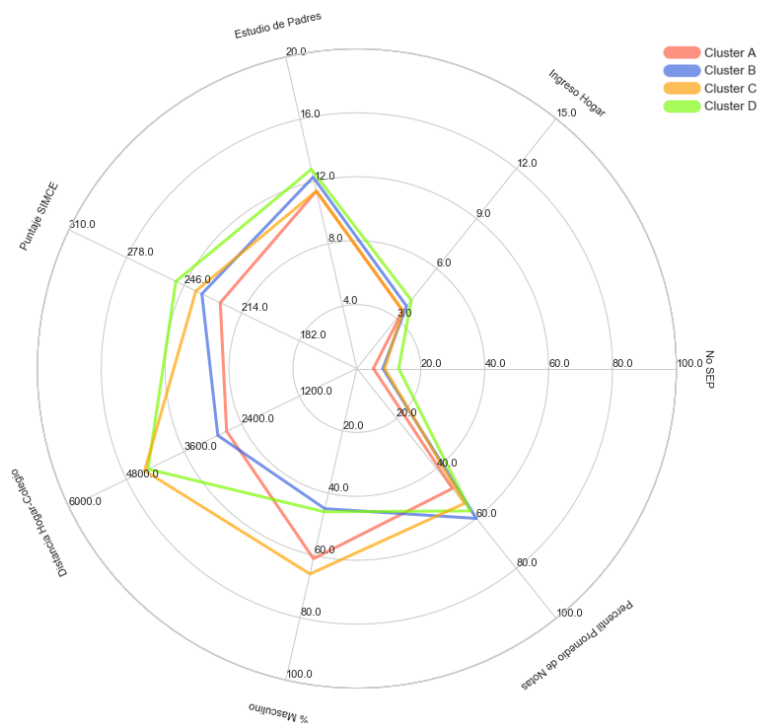


Figura 5.41: Variables a nivel de estudiante.

puntos en promedio). Asimismo, los estudiantes de los *clusters* A y B recorren distancias (Distancia Hogar Colegio) en torno a los 2,9 km en promedio, mientras que los estudiantes de los *clusters* C y D recorren mayores distancias de 4,5 km en promedio.

Con respecto a la variable de género (Porcentaje de Hombres), los *clusters* presentan una composición disímil: los *clusters* A y C están compuesto por un 60% y 65% de hombres respectivamente, mientras que los *clusters* B y D por un 47%. Por otro lado, también se observan algunas diferencias en rendimiento académico (Percentil Promedio de Notas): estudiantes del *cluster* A tienen el rendimiento menor (percentil 48), seguidos del *cluster* C (percentil 53), luego del *cluster* D (percentil 58), y finalmente por el *cluster* B (percentil 60), que demuestra un rendimiento en promedio levemente mayor.

En este escenario se pueden destacar las siguientes características:

- Los *clusters* están ordenados de acuerdo a la transición entre dependencia administrativa.
- La mayoría de los cambios ocurren desde escuelas municipales, mientras que la minoría lo hacen desde particular subvencionado hacia municipal.
- No se observan elevadas diferencias en IVE, beneficiarios SEP, Ingreso del Hogar y Educación de los Padres.
- Se observa un sesgo de género importante en los *clusters* A y C (cambios hacia municipales).
- En general el copago de los establecimientos es bajo.
- Todos los *clusters* tienen alta composición de beneficiarios SEP.

- Movilidad hacia particulares subvencionados mejora en rendimiento, medido por el SIMCE, mientras que movilidad hacia municipales empeora en este valor.
- Se identifica una mayor preferencia por establecimientos particulares subvencionados.

SAE Admisión

Este escenario hay 4 *clusters*. A partir de la Figura 5.42 (b), se observa que estos quedan bien definidos por la relación entre dependencias administrativas de origen y destino: el *cluster* A representa la movilidad de estudiantes entre municipales, el *cluster* B aquella desde municipales hacia particulares subvencionados, el *cluster* C comprende los cambios desde particulares subvencionados hacia municipales, y el *cluster* D agrupa cambios entre establecimientos particulares subvencionados. Sin embargo, también hay pequeños flujos desde todos los *clusters* que llegan a establecimientos particulares pagados. De las Figuras 5.42 (b) y (c), se observa que el *cluster* A tiene la mayor representación porcentual de estudiantes del escenario (49%), seguido por el *cluster* B (30%), mientras que los *clusters* C y D comprenden el 10% y 11% de los cambios, respectivamente.

Cluster	Copago		IVE		Promedio SIMCE		Dependencia	
	rbd origen	rbd admitido	rbd origen	rbd admitido	rbd origen	rbd admitido	rbd origen	rbd admitido
A	P83 (Gratuito)	P83 (Gratuito)	0,85	0,85	P35	P13	Municipal	Municipal
B	P83 (Gratuito)	P84	0,83	0,74	P39	P55	Municipal	Part. Subv.
C	P83 (Gratuito)	P83 (Gratuito)	0,82	0,84	P55	P22	Part. Subv.	Municipal
D	P83 (Gratuito)	P84	0,8	0,74	P64	P52	Part. Subv.	Part. Subv.

Tabla 5.23: Centroides de *clusters*.

La Tabla 5.23 y la Figura 5.42 (a) muestran los centroides de cada *cluster*. Se observa que el Copago es constante en los *clusters* A y C, y aumenta en los *clusters* B y D. Por otra parte, la vulnerabilidad (IVE) disminuye en los *clusters* B y D, en tanto que se mantiene casi constante en los *clusters* A y C. Por otro lado, en términos de resultados SIMCE de EE de origen y destino, sólo en el *cluster* B hay una mejora en rendimiento, mientras que en los *clusters* A, C y D los puntajes disminuyen.

Al comparar las variables a nivel de estudiante, en la Figura 5.43 se observa, para la variable de vulnerabilidad (No participa en la SEP), que todos los *clusters* presentan bajos valores de "no vulnerabilidad" (menos del 13% de los estudiantes no son Beneficiarios SEP), es decir, todos tienen una alta composición de beneficiarios SEP. Asimismo, con respecto al Ingreso del Hogar, todos los *clusters* representan ingresos entre \$300.000 y \$400.000. Por otro lado, para la variable Educación de los Padres, cada *cluster* representa en promedio educación hasta segundo o tercer medio.

Los estudiantes de los *clusters* C y D obtienen los mayores puntajes, de 246 puntos en promedio, seguidos por el *cluster* B (235 puntos en promedio), y finalmente por el *cluster* A (230 puntos en promedio). Asimismo, los estudiantes de los *clusters* A y B recorren distancias (Distancia Hogar Colegio) en torno a los 2,9 km en promedio, mientras que los estudiantes de los *clusters* C y D recorren mayores distancias de 4 y 4,7 km en promedio, respectivamente.

Con respecto a la variable de género (Porcentaje de Hombres), los *clusters* presentan una composición disímil: los *clusters* A y c están compuesto por un 60% de hombres, mientras

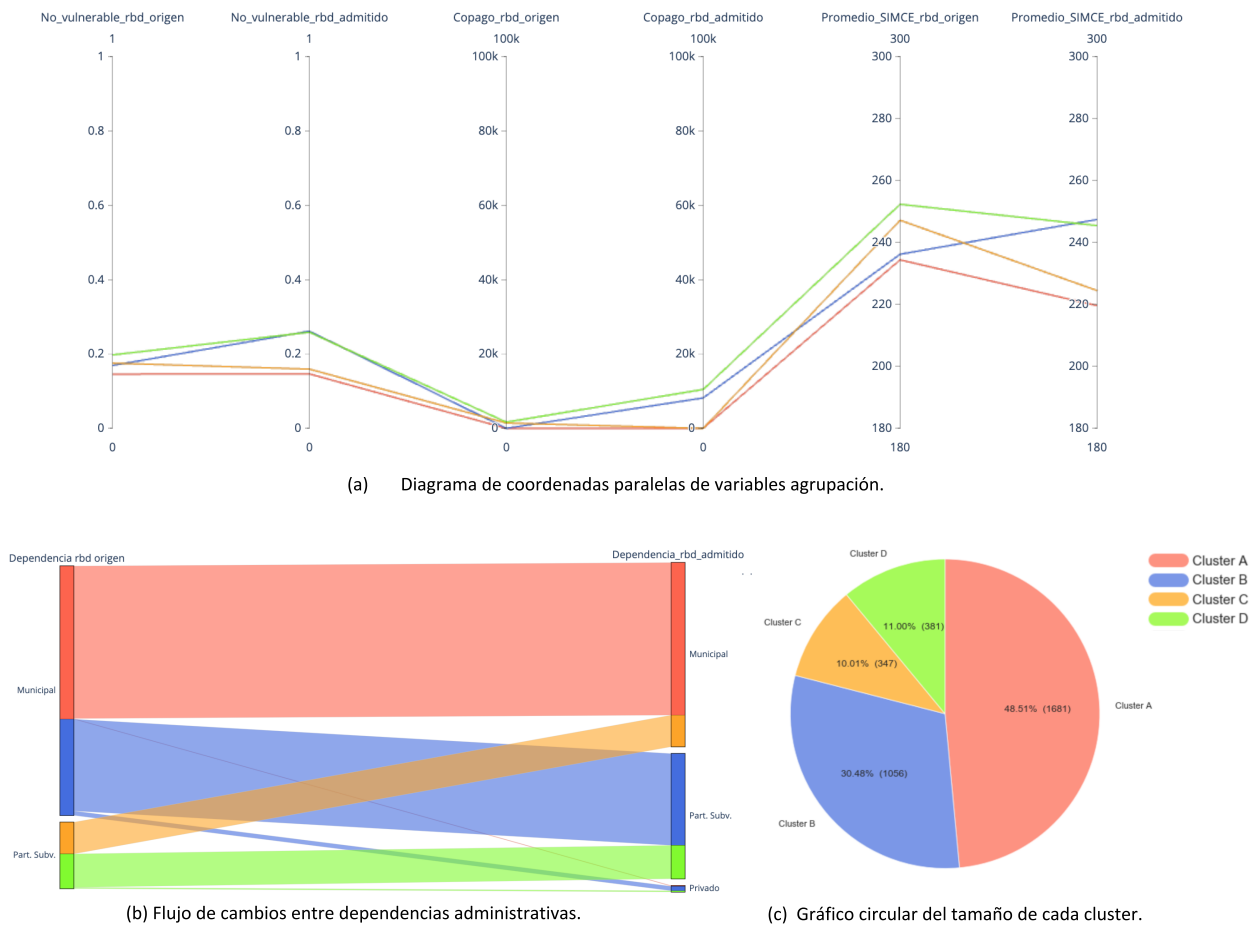


Figura 5.42: Variables a nivel de establecimiento.

que los *clusters* B y D por un 42%. Por otro lado, también se observan algunas diferencias en rendimiento académico (Percentil Promedio de Notas): estudiantes del *cluster* A tienen el rendimiento menor (percentil 48 en promedio), seguidos del *cluster* C (percentil 53), y finalmente por los *clusters* B y D (percentil 60 en promedio).

En este escenario se pueden destacar las siguientes características:

- Los *clusters* están ordenados de acuerdo a la transición entre dependencia administrativa.
- La mayoría de los cambios ocurren desde escuelas municipales, mientras que la minoría lo hacen desde particular subvencionado hacia municipal.
- No se observan elevadas diferencias en IVE, beneficiarios SEP, Ingreso del Hogar y Educación de los Padres.
- Se observa un sesgo de género importante en todos los *clusters*.
- En general el copago de los establecimientos es bajo.
- Todos los *clusters* tienen alta composición de beneficiarios SEP.
- Únicamente la movilidad desde municipales hacia particulares subvencionados mejora en rendimiento, medido por el SIMCE, mientras que el resto de los perfiles de movilidad empeoran en este valor.

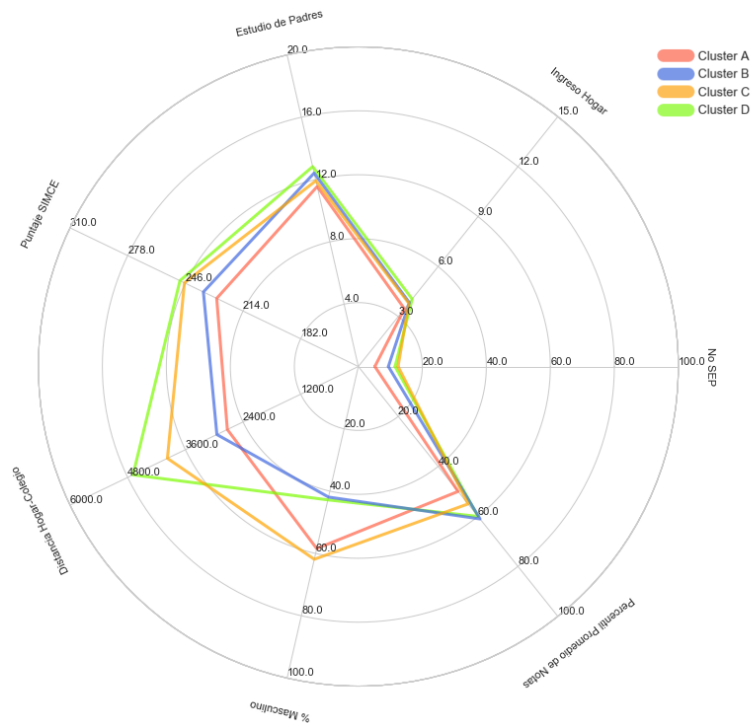


Figura 5.43: Variables a nivel de estudiante.

- Se identifica una mayor admisión en establecimientos municipales.

Patrones de Cambio Forzado en Puerto Montt - Puerto Varas

Al comparar todos los escenarios del caso Forzado, vemos que las transiciones entre dependencias administrativas definen *clusters*. Además, se observa que las transiciones según IVE son similares: cambios hacia municipales (*clusters* A y C) no varían mucho en el indicador, mientras que cambios hacia particulares subvencionados disminuyen en vulnerabilidad del establecimiento. Para la variable copago, se observa que su valor absoluto disminuye en los escenarios SAE, aunque en términos de percentiles este aumenta, lo cual se explica por la gratuidad progresiva en establecimientos particulares subvencionados impulsada por la ley de inclusión escolar.

Los patrones de resultados SIMCE a nivel de establecimiento son similares en los escenarios Pre SAE y SAE Preferencias, aunque en el caso de cambios hacia particulares subvencionados (*clusters* B y D) se demandan establecimientos con mejores resultados que aquellos que son asignados a los estudiantes (SAE Admisión).

En el escenario SAE Preferencias, se destaca, una disminución en la demanda por EE municipales y un aumento en la demanda por establecimientos particulares subvencionados, la cual no se satisface, lo que queda claro al observar la representatividad de los *clusters* en el escenario SAE Admisión. Sin embargo, los cambios hacia particulares subvencionados, según asignación, aumentan con respecto al escenario Pre SAE.

Al contrastar las características de los estudiantes pertenecientes a cada *cluster*, se observa en todos los escenarios que se cambian más mujeres hacia establecimientos particular subven-

cionado (*clusters* B y D), mientras más hombres se cambian hacia municipales (*clusters* A y C), todo en términos proporcionales. Por otro lado, también se remarca que los estudiantes que cambian desde municipal a particular subvencionado tienen mejor desempeño académico en todos los escenarios.

Finalmente, al igual que en otros casos de cambio Forzado, los *cluster* presentan una buena calidad en término de definición y estructura, lo que es indicado por el coeficiente de silueta en cada agrupación.

5.4.2. Cambios No Forzados

5.4.2.1. Escenario Pre SAE

Este escenario hay 5 *clusters*. A partir de la Figura 5.44 (b), se observa que estos quedan parcialmente definidos por la relación entre dependencias administrativas de origen y destino: el *cluster* A representa la movilidad de estudiantes entre municipales, el *cluster* B aquella desde municipales hacia particulares subvencionados, el *cluster* C comprende los cambios desde particulares subvencionados hacia municipales, el *cluster* D agrupa cambios entre establecimientos particulares subvencionados, y el *cluster* E representa algunos cambios entre particulares subvencionados, además de aquellos que involucran establecimientos particulares pagados. De las Figuras 5.44 (b) y (c), se observa que los *clusters* A, C y D tienen la mayor representación porcentual de estudiantes del escenario (27 %, 25 % y 30 % respectivamente), mientras que los *clusters* B y E comprenden el 12 % y 7 % de los cambios, respectivamente.

Cluster	Copago rbd origen	Copago rbd destino	IVE rbd origen	IVE rbd destino	Promedio SIMCE rbd origen	Promedio SIMCE rbd destino	Dependencia rbd origen	Dependencia rbd destino
A	P61 (Gratuito)	P62	0,83	0,79	P31	P39	Municipal	Municipal
B	P61 (Gratuito)	P75	0,81	0,68	P36	P59	Municipal	Part. Subv.
C	P77	P62	0,67	0,79	P61	P39	Part. Subv.	Municipal
D	P80	P81	0,60	0,60	P66	P64	Part. Subv.	Part. Subv.
E	P90	P94	0,27	0,07	P76	P80	Subv./Priv.	Subv./Priv.

Tabla 5.24: Centroides de *clusters*.

La Tabla 5.24 y la Figura 5.44 (a) muestran los centroides de cada *cluster*. Se observa que el Copago disminuye sólo en el *cluster* C, es constante en los *clusters* A y D, y aumenta en los *clusters* B y E. Por otra parte, la vulnerabilidad (IVE) disminuye en los *clusters* A, B y E, aumenta en el *cluster* C, en tanto que permanece constante en el *cluster* D. Por otro lado, en términos de resultados SIMCE de EE de origen y destino, en los *clusters* A, B y E hay una mejora en rendimiento, mientras que en los *clusters* C y D los puntajes disminuyen.

Al comparar las variables a nivel de estudiante, en la Figura 5.45 se observa, para la variable de vulnerabilidad (No participa en la SEP), que los *clusters* A y B son los "menos no vulnerables" (en promedio 30 % y 38 % de los estudiantes no son Beneficiarios SEP, respectivamente), es decir, son los más vulnerable, seguidos, en orden decreciente de vulnerabilidad, por los *clusters* C y D (en promedio el 50 % y 67 % de los estudiantes no son Beneficiario SEP, respectivamente); y finalmente, el *cluster* E tiene una muy baja vulnerabilidad (94 % de los estudiantes no son Beneficiarios SEP).

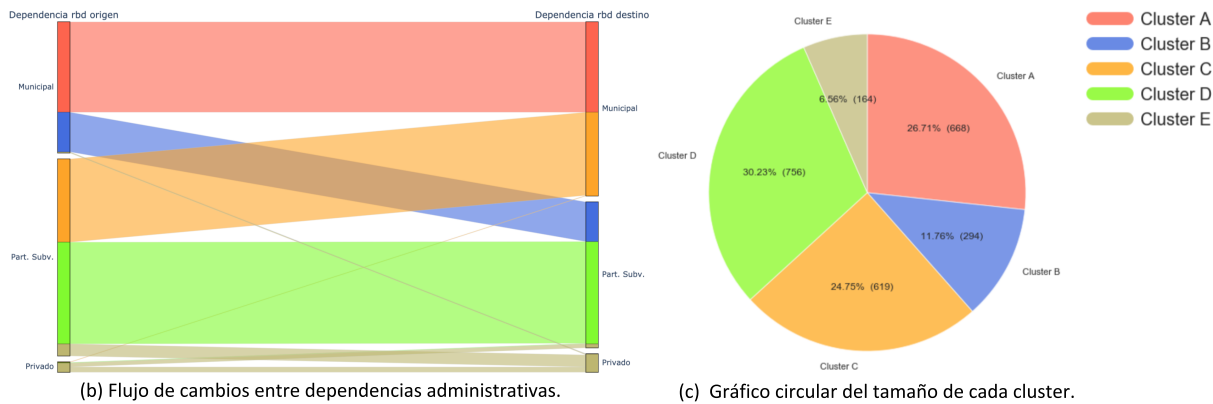
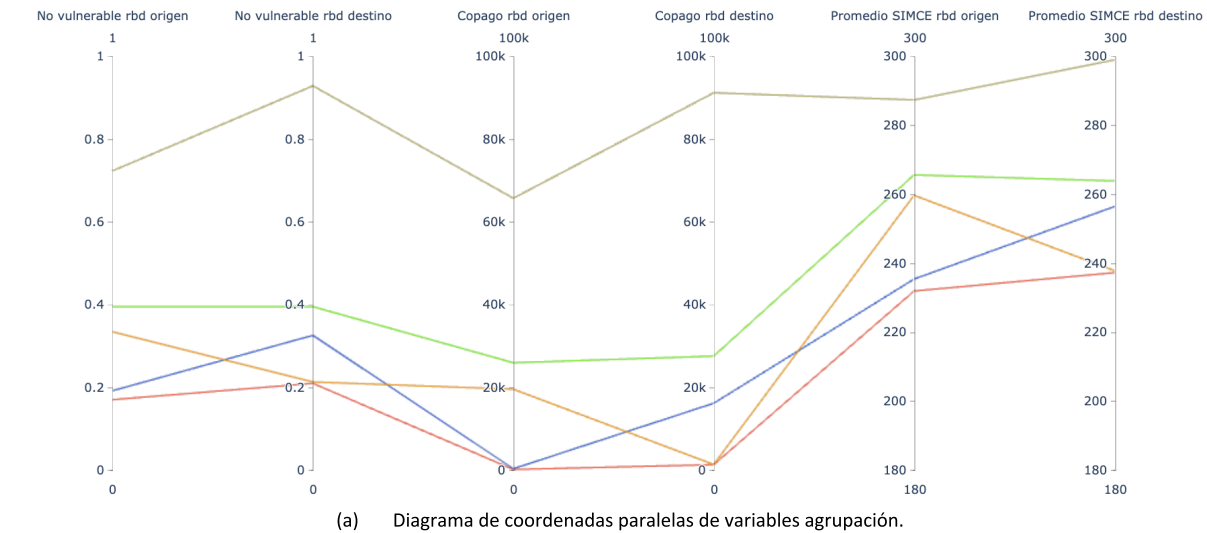


Figura 5.44: Variables a nivel de establecimiento.

Con respecto al Ingreso del Hogar, el *cluster A* representa ingresos entre \$300.000 y \$400.000, seguido por los *clusters B* y *C* (entre \$400.000 y \$500.000), luego por el *cluster D* (entre \$500.000 y \$600.000, valor 6), en tanto que el *cluster E* representa los mayores ingresos, los que figuran entre \$1.400.000 y \$1.600.000 en promedio. Por otro lado, para la variable Educación de los Padres, los *clusters B* y *C* presentan en promedio educación media completa, los padres del *cluster A* poseen educación, en promedio, hasta tercero medio, y aquellos del *cluster D* promedian educación incompleta en IP o CFT. En tanto, padres del *cluster E* poseen, en general, educación universitaria completa.

Los estudiantes de los *clusters A, B, C* y *D* obtienen en promedio puntajes SIMCE entre 240 y 260 puntos (en orden creciente de puntaje), mientras que los del *cluster E* superan al resto, con un rendimiento promedio de 295 puntos. Asimismo, los estudiantes de los *clusters B, C, D* y *E* recorren distancias (Distancia Hogar Colegio) de 3; 3,6; 4 y 2,6 km en promedio, respectivamente; en tanto que los estudiantes del *cluster A* recorren en promedio las menores distancias (1,8 km).

Con respecto a la variable de género (Porcentaje de Hombres), los *clusters* presentan una composición disímil: los *clusters A* y *C* está compuesto por un 60% de hombres, el *cluster E*

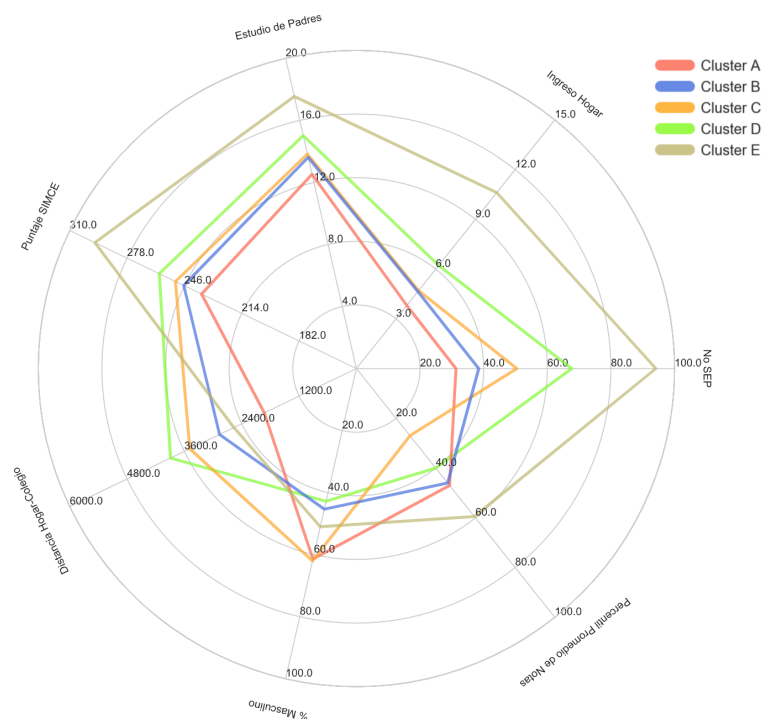


Figura 5.45: Variables a nivel de estudiante.

por un 50 %, y los *clusters* B y D por un 45 %. Por otro lado, se observan elevadas diferencias en el rendimiento académico de los estudiantes (Percentil Promedio de Notas): estudiantes de los *clusters* A, B y D se ubican entre los percentiles 40 y 45, mientras que aquellos del *cluster* C tienen el menor rendimiento (percentil 27 en promedio), y los del *cluster* E presentan el más alto desempeño académico, ubicándose en el percentil 60 en promedio.

En este escenario se pueden destacar las siguientes características:

- Los *clusters* están totalmente ordenados de acuerdo a la transición entre dependencia administrativa, exceptuando el *cluster* E.
- La mayoría de los cambios ocurren desde escuelas subvencionadas (*clusters* C, D y E) y la minoría desde particulares pagados (*cluster* E).
- La vulnerabilidad del *cluster* E es menor: menor IVE, menos beneficiarios SEP, mayor Copago, mayor rendimiento, medido por el SIMCE del establecimiento, mayor Ingreso del Hogar, mayor Educación de los Padres y mejores puntajes SIMCE (estudiantes).
- Los establecimientos particulares pagados tienen una mayor participación que en otros escenarios.
- Se observa un sesgo de género importante en los *clusters* A y C.
- Cambios del *cluster* E sobresalen fuertemente en desempeño académico (Promedio de Notas).
- El *cluster* A es el más vulnerable.

5.4.2.2. Escenario SAE

SAE Preferencias

Este escenario hay 5 *clusters*. A partir de la Figura 5.46 (b), se observa que estos quedan parcialmente definidos por la relación entre dependencias administrativas de origen y destino: el *cluster* A representa la movilidad de estudiantes entre municipales, el *cluster* B aquella desde municipales hacia particulares subvencionados, el *cluster* C comprende los cambios desde particulares subvencionados hacia municipales, el *cluster* D agrupa cambios entre establecimientos particulares subvencionados, y el *cluster* E representa algunos cambios entre particulares subvencionados, además de aquellos que involucran establecimientos particulares pagados (aunque todos los *clusters* tienen muy pequeños flujos hacia particulares pagados también). De las Figuras 5.46 (b) y (c), se observa que el *cluster* D tiene la mayor representación porcentual de estudiantes del escenario (37%), mientras que los *clusters* A, B, C y E comprenden el 15%, 26%, 15% y 8% de los cambios, respectivamente.

Cluster	Copago rbd origen	Copago rbd preferencia	IVE rbd origen	IVE rbd preferencia	Promedio SIMCE rbd origen	Promedio SIMCE rbd preferencia	Dependencia rbd origen	Dependencia rbd preferencia
A	P83 (Gratuito)	P83 (Gratuito)	0,85	0,76	P32	P46	Municipal	Municipal
B	P83 (Gratuito)	P84	0,83	0,62	P35	P73	Municipal	Part. Subv.
C	P84	P83 (Gratuito)	0,69	0,73	P65	P57	Part. Subv.	Municipal
D	P84	P85	0,67	0,6	P67	P74	Part. Subv.	Part. Subv.
E	P91	P94	0,3	0,15	P86	P87	Subv./Priv.	Privado

Tabla 5.25: Centroides de *clusters*.

La Tabla 5.25 y la Figura 5.46 (a) muestran los centroides de cada *cluster*. Se observa que el Copago disminuye sólo en el *cluster* C, es constante en los *clusters* A y D, y aumenta en los *clusters* B y E. Por otra parte, la vulnerabilidad (IVE) disminuye en los *clusters* A, B, D y E, en tanto que aumenta en el *cluster* C. Por otro lado, en términos de resultados SIMCE de EE de origen y destino, en los *clusters* A, B, D y E hay una mejora en rendimiento, medido por el SIMCE, mientras que en el *cluster* C los puntajes disminuyen.

Al comparar las variables a nivel de estudiante, en la Figura 5.47 se observa, para la variable de vulnerabilidad (No participa en la SEP), que los *clusters* A y B son los "menos vulnerables" (en promedio el 8% y 12% de los estudiantes no son Beneficiarios SEP, respectivamente), es decir, son los más vulnerables, seguidos, en orden decreciente de vulnerabilidad, por los *clusters* C y D (en promedio el 26% y 36% de los estudiantes no son Beneficiarios SEP, respectivamente); y finalmente, el *cluster* E tiene una muy baja vulnerabilidad (93% de los estudiantes no son Beneficiarios SEP).

Con respecto al Ingreso del Hogar, el *cluster* A representa ingresos entre \$200.000 y \$300.000, seguido por los *clusters* B y C (entre \$400.000 y \$500.000), luego por el *cluster* D (entre \$500.000 y \$600.000), en tanto que el *cluster* E representa los mayores ingresos, los que figuran entre \$1.600.000 y \$1.800.000 en promedio. Por otro lado, para la variable Educación de los Padres, los *clusters* B y C presentan en promedio educación media completa, los padres del *cluster* A poseen educación, en promedio, hasta segundo medio, y aquellos del *cluster* D promedian educación incompleta en IP o CFT. En tanto, padres del *cluster* E poseen, en general, educación universitaria incompleta.

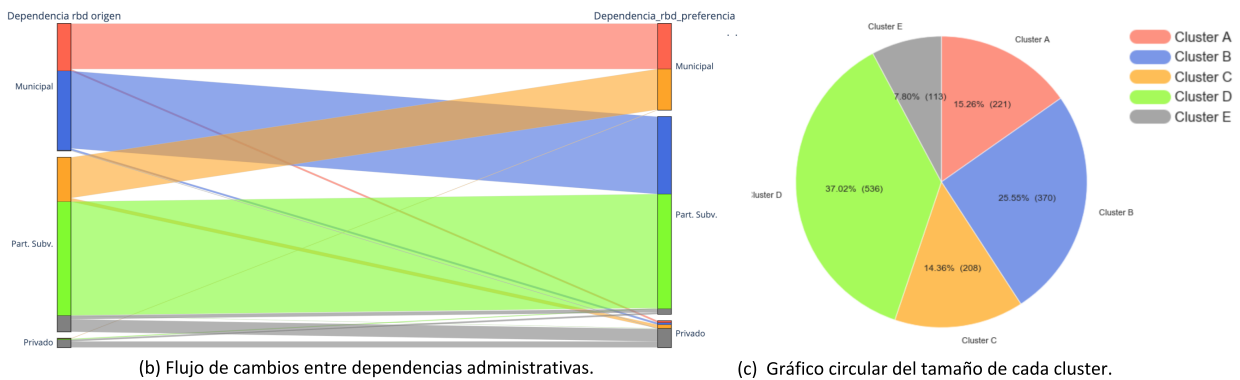
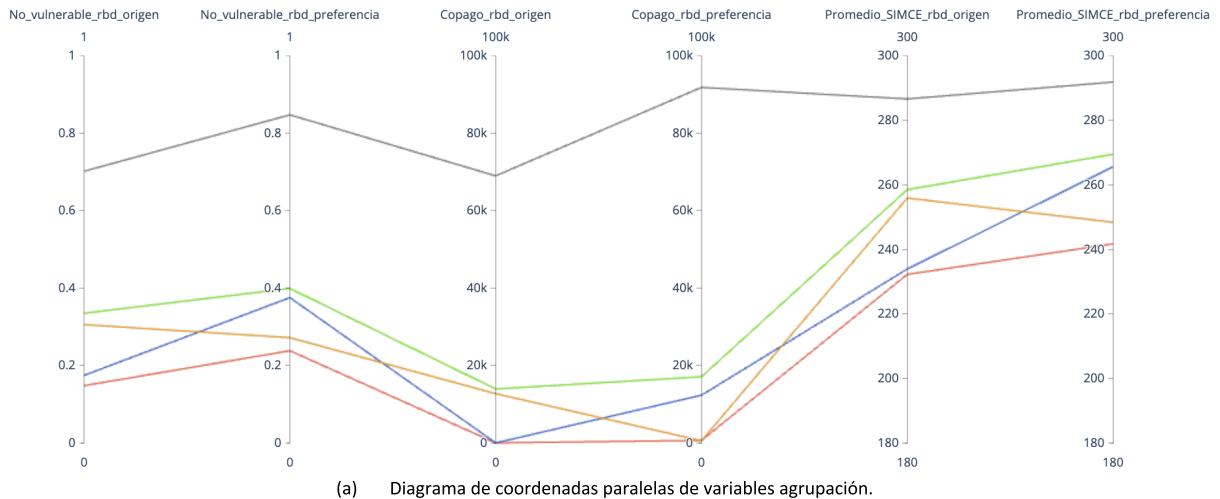


Figura 5.46: Variables a nivel de establecimiento.

Los estudiantes de los *clusters* B, C y D obtienen en promedio puntajes SIMCE entre 240 y 258 puntos (en orden creciente de puntaje), aquellos del *cluster* A obtienen 214 puntos en promedio, mientras que los del *cluster* E superan al resto, con un rendimiento promedio de 293 puntos. Asimismo, los estudiantes de los *clusters* B, C y D recorren distancias promedio (Distancia Hogar Colegio) entorno a las 3,4 km, en tanto que los estudiantes del *cluster* A recorren en promedio 2,4 km y aquellos del *cluster* E recorren las menores distancias (1,8 km).

Con respecto a la variable de género (Porcentaje de Hombres), los *clusters* presentan una composición disímil: los *clusters* A y C está compuesto por un 72% y un 66% de hombres respectivamente, los *clusters* B y E por un 50%, y el *cluster* D por un 37%. Por otro lado, se observan diferencias en el rendimiento académico de los estudiantes (Percentil Promedio de Notas): estudiantes de los *clusters* A y C se ubican entre los percentiles 32 y 38 (menor rendimiento), seguidos por aquellos del *cluster* D (percentil 43), y por los *clusters* B y E, que presentan el más alto desempeño académico, ubicándose en el percentil 53 en promedio.

En este escenario se pueden destacar las siguientes características:

- A partir de todos los *clusters* hay un flujo de cambio hacia particulares pagados, lo que sugiere un mayor desorden en la organización de la movilidad en el escenario no

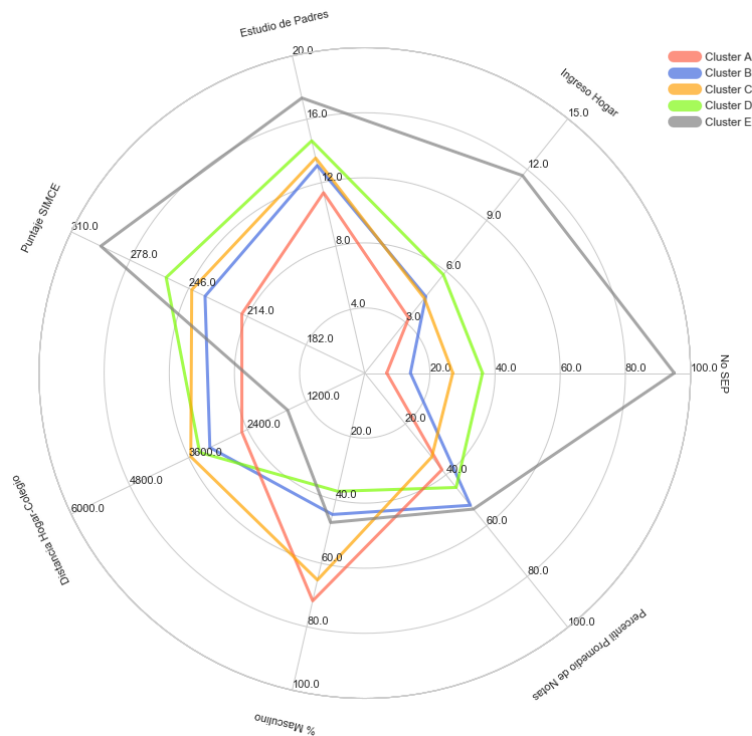


Figura 5.47: Variables a nivel de estudiante.

forzado.

- La mayoría de los cambios ocurren desde escuelas subvencionadas (*clusters* C, D y E) y la minoría desde particulares pagados (*cluster* E).
- La vulnerabilidad del *cluster* E es mucho menor: menor IVE, menos beneficiarios SEP, mayor Copago, mayor rendimiento, medido por el SIMCE del establecimiento, mayor Ingreso del Hogar, mayor Educación de los Padres y mejores puntajes SIMCE (estudiantes).
- Los establecimientos particulares pagados tienen una mayor participación que en otros escenarios.
- Se observa un sesgo de género importante en los *clusters* A, C y D.
- El *cluster* A es el más vulnerable y el que tiene mayor composición de hombres.
- El copago en todos los *cluster* es bajo, exceptuando en el *cluster* E.

SAE Admisión

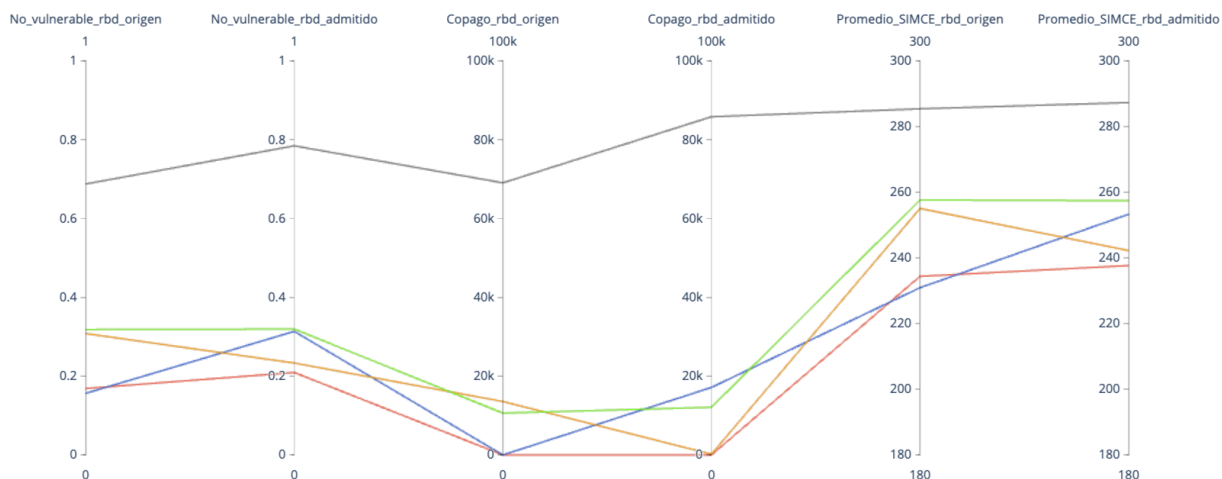
Este escenario hay 5 *clusters*. A partir de la Figura 5.48 (b), se observa que estos quedan parcialmente definidos por la relación entre dependencias administrativas de origen y destino: el *cluster* A representa la movilidad de estudiantes entre municipales, el *cluster* B aquella desde municipales hacia particulares subvencionados, el *cluster* C comprende los cambios desde particulares subvencionados hacia municipales, el *cluster* D agrupa cambios entre establecimientos particulares subvencionados, y el *cluster* E representa algunos cambios entre particulares subvencionados, además de aquellos que involucran establecimientos particulares

pagados (aunque todos los *clusters* tienen muy pequeños flujos hacia particulares pagados también). De las Figuras 5.48 (b) y (c), se observa que el *cluster* D tiene la mayor representación porcentual de estudiantes del escenario (33%), mientras que los *clusters* A, B, C y E comprenden el 28%, 13%, 16% y 10% de los cambios, respectivamente.

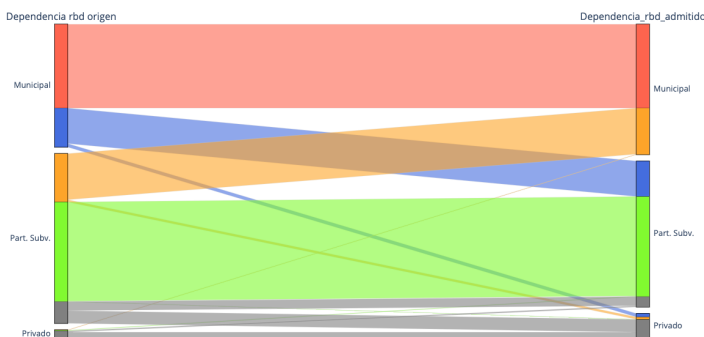
Cluster	Copago rbd origen	Copago rbd admitido	IVE rbd origen	IVE rbd admitido	Promedio SIMCE rbd origen	Promedio SIMCE rbd admitido	Dependencia rbd origen	Dependencia rbd admitido
A	P83 (Gratuito)	P83 (Gratuito)	0,83	0,79	P35	P41	Municipal	Municipal
B	P83 (Gratuito)	P85	0,84	0,69	P30	P64	Municipal	Part. Subv.
C	P84	P83 (Gratuito)	0,69	0,77	P65	P49	Part. Subv.	Municipal
D	P84	P84	0,68	0,68	P67	P67	Part. Subv.	Part. Subv.
E	P91	P93	0,31	0,22	P86	P86	Part. Subv.	Subv./Priv.

Tabla 5.26: Centroides de *clusters*.

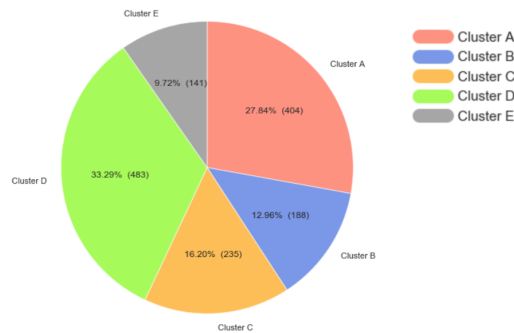
La Tabla 5.26 y la Figura 5.48 (a) muestran los centroides de cada *cluster*. Se observa que el Copago disminuye sólo en el *cluster* C, es constante en los *clusters* A y D, y aumenta en los *clusters* B y E. Por otra parte, la vulnerabilidad (IVE) disminuye en los *clusters* A, B, D y E, en tanto que aumenta en el *cluster* C. Por otro lado, en términos de resultados SIMCE de EE de origen y destino, en los *clusters* A y B hay una mejora en rendimiento, medido por el SIMCE, este permanece constante en los *clusters* D y E, mientras que en el *cluster* C los puntajes disminuyen.



(a) Diagrama de coordenadas paralelas de variables agrupación.



(b) Flujo de cambios entre dependencias administrativas.



(c) Gráfico circular del tamaño de cada cluster.

Figura 5.48: Variables a nivel de establecimiento.

Al comparar las variables a nivel de estudiante, en la Figura 5.49 se observa, para la

variable de vulnerabilidad (No participa en la SEP), que los *clusters* A y B son los "menos vulnerables" (en torno al 10 % de los estudiantes no son Beneficiarios SEP), es decir, son los más vulnerables, seguidos, en orden decreciente de vulnerabilidad, por los *clusters* C y D (en promedio el 26 % y 32 % de los estudiantes no son Beneficiarios SEP, respectivamente); y finalmente, el *cluster* E tiene una muy baja vulnerabilidad (95 % de los estudiantes no son Beneficiarios SEP).

Con respecto al Ingreso del Hogar, los *clusters* A y B representan ingresos entre \$300.000 y \$400.000, seguido por el *cluster* C (entre \$400.000 y \$500.000), luego por los *clusters* D y E (entre \$500.000 y \$600.000), en tanto que el *cluster* E representa los mayores ingresos, los que figuran entre \$1.400.000 y \$1.600.000 en promedio. Por otro lado, para la variable Educación de los Padres, los *clusters* B, C y D presentan en promedio educación media completa, los padres del *cluster* A poseen educación, en promedio, hasta tercer medio, en tanto, padres del *cluster* E poseen, en general, educación universitaria incompleta.

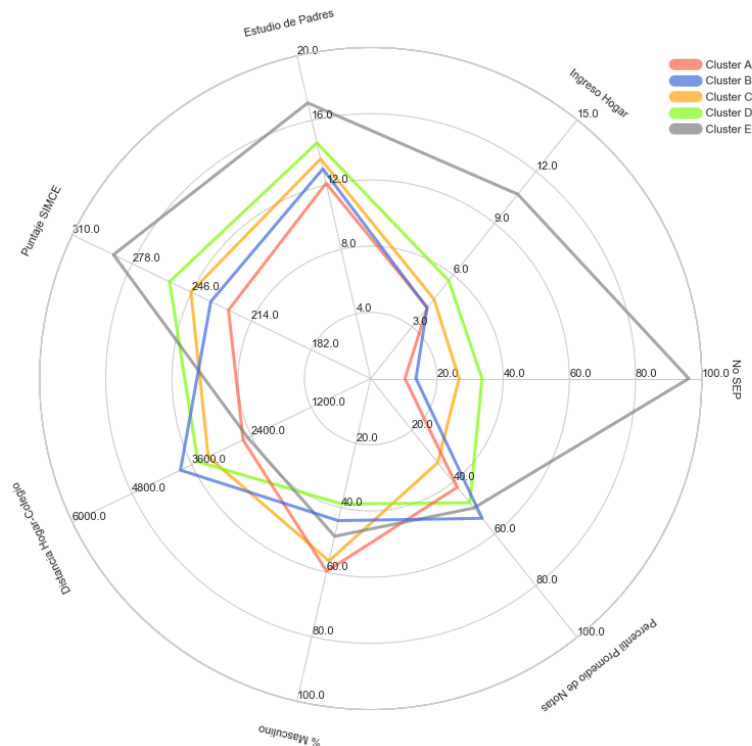


Figura 5.49: Variables a nivel de estudiante.

Los estudiantes de los *clusters* A, B, C y D obtienen en promedio puntajes SIMCE entre 225 y 258 puntos (en orden creciente de puntaje), mientras que los del *cluster* E superan al resto, con un rendimiento promedio de 285 puntos. Asimismo, los estudiantes de los *clusters* B, C y D recorren distancias promedio (Distancia Hogar Colegio) entorno a las 3,6 km, en tanto que los estudiantes de los *clusters* A y E recorren en promedio 2,4 km.

Con respecto a la variable de género (Porcentaje de Hombres), los *clusters* presentan una composición disímil: los *clusters* A y C está compuesto por un 58 % de hombres, los *clusters* B y E por un 48 %, y el *cluster* D por un 39 %. Por otro lado, se observan diferencias en el rendimiento académico de los estudiantes (Percentil Promedio de Notas): estudiantes de los

clusters C y A se ubican en los percentiles 32 y 42 respectivamente (menor rendimiento), seguidos por aquellos del *cluster* D (percentil 45), y por los *clusters* B y E, que presentan el más alto desempeño académico, ubicándose en el percentil 53 en promedio.

En este escenario se pueden destacar las siguientes características:

- A partir de todos los *clusters* hay un flujo de cambio hacia particulares pagados, lo que sugiere un mayor desorden en la organización de la movilidad en el escenario no forzado.
- La mayoría de los cambios ocurren desde escuelas subvencionadas (*clusters* C, D y E) y la minoría desde particulares pagados (*cluster* E).
- La vulnerabilidad del *cluster* E es mucho menor: menor IVE, menos beneficiarios SEP, mayor Copago, mayor rendimiento, medido por el SIMCE de establecimiento, mayor Ingreso del Hogar, mayor Educación de los Padres y mejores puntajes SIMCE (estudiantes).
- Los establecimientos particulares pagados tienen una mayor participación que en otros escenarios.
- Se observa un sesgo de género importante en los *clusters* A, C y D.
- El *cluster* A es el más vulnerable y el que tiene mayor composición de hombres.
- El copago en todos los *cluster* es bajo, exceptuando en el *cluster* E.

Patrones de Cambio No Forzado en Puerto Montt - Puerto Varas

Se destaca, que el *cluster* E aparece en todos los escenarios como aquel que identifica transiciones de estudiantes de GSE alto, los que resaltan con mayor copago, mejores resultados SIMCE (por establecimiento y por estudiantes), mayores Ingreso del Hogar y Educación de los Padres.

Al comparar patrones de cambios según IVE de establecimientos de origen y destino, se observa que estos son similares en todos los escenarios, siendo un poco menor su variación en el escenario SAE Admisión, aunque sigue manteniendo la estructura. Este mismo efecto se observa en la variable copago.

En cuanto a la variable de rendimiento, medido por el SIMCE del establecimiento, esta sigue patrones similares en todos los escenarios, remarcando que para los *clusters* B y D (cambios hacia particulares subvencionados), la preferencia por puntajes más altos no es completamente cubierta, y en consecuencia los resultados de SAE Admisión son similares a los del escenario Pre SAE.

Por otro lado, del escenario SAE Preferencias se observa una disminución en la preferencia por establecimientos municipales y un aumento por particulares subvencionados. Sin embargo, la capacidad del sistema no es lo suficientemente elástica, por lo que en términos de asignación (SAE Admisión), la repartición entre dependencias administrativas de destino sigue siendo similares a las del escenario Pre SAE.

Al comparar variable a nivel de estudiante, se observa que en cambios hacia municipales (*clusters* A y c) hay mayor representación masculina, mientras que en cambios hacia

particulares subvencionados (*clusters* B y D) hay mayor representación femenina.

En general, la calidad de los *cluster* es buena excepto para *cluster* E, el cual integra flujos variados entre dependencias administrativas, lo que hace reducir su coeficiente de silueta.

Patrones de Cambio en Puerto Montt - Puerto Varas

Al comparar todos los escenarios y casos correspondiente a esta aglomeración urbana, se observa claramente una mayor estructura en el caso Forzado: tanto en el número de *clusters* como en su caracterización según transiciones de dependencia administrativa. En este mismo sentido, sólo en el caso No Forzado aparecen *clusters* de estudiantes de GSE alto.

Al contrastar los tipos de cambio Forzado y No Forzado, se destaca, que en el segundo la proporción de cambios desde la dependencia municipal es mucho menor, esto pues hay muchos establecimientos municipales que entregan educación sólo hasta los niveles estudiados, y por ende, caen en la categoría de cambios Forzados.

Para los casos SAE, en ambos tipos de cambio hay una demanda por resultados SIMCE y por menor vulnerabilidad (IVE) que no es completamente cubierta. Además, en ambos casos hay una demanda mayor por establecimientos particulares subvencionados que no logra ser cubierta por el sistema. Sin embargo, en ambos casos la proporción de estudiantes asignados (SAE Admisión) a establecimientos particulares subvencionados es mayor a la que alcanzaban estos en el escenario Pre SAE.

Capítulo 6

Conclusiones

El nuevo Sistema de Admisión Escolar (SAE) chileno ha cambiado el paradigma bajo el cual funcionaba el cambio de escuelas en el país. El que está bajo un periodo de transición, desde un mecanismo donde los establecimientos educacionales seleccionaban a los alumnos, hacia uno donde son las familias las que, mediante sus preferencias declaradas, escogen escuelas para sus hijos a través de un sistema único que se las asigna a través del algoritmo (sección 2.3.2.1).

En este contexto, el presente trabajo busca caracterizar la elección escolar en términos de atributos de los EE y estudiantes, y analizar cómo esta se ha modificado producto de la implementación del SAE.

Con este fin, se utilizaron modelos de aprendizaje no supervisado sobre los datos de cambio de escuelas durante el periodo 2013-2019 en cuatro aglomeraciones urbanas del país. Mediante una metodología rigurosa de exploración de datos, selección de atributos, selección y optimización de modelos, y visualizaciones, se ha indagado acerca de los patrones de movilidad escolar en los niveles de sexto y octavo básico, considerando situaciones en que el cambio es Forzado, porque la escuela no ofrece los grados siguientes, y no Forzado, en cuyo caso los estudiantes se cambian aun cuando la escuela sí ofrecía el siguiente nivel. Se consideraron tres escenarios: lo que ocurría antes de la implementación del SAE (Pre SAE), las preferencias declaradas por los padres en la postulación a los establecimientos educacionales (SAE Preferencias) y lo que ocurre en el proceso de admisión (SAE Admisión). Estas tendencias han sido caracterizadas bajo esta metodología, la que determina *clusters* de cambio de establecimiento en base a las características de los colegios de origen y destino de los estudiantes que cambian de escuela. Este trabajo también analiza las características de los estudiantes que cambian de escuela, según el *cluster* que los identifica.

Dados los escenarios considerados (ver Figura 5.1), se han podido establecer comparaciones de patrones de cambio de escuela en diferentes dimensiones para cada conurbación, entre las que se consideran las siguientes:

- Escenario Pre SAE vs. Escenario SAE Admisión: permite identificar, en términos de los cambios de escuelas efectivos, de qué manera se ha modificado la movilidad escolar

producto del nuevo sistema de admisión.

- Escenario SAE Preferencias vs. Escenario SAE Admisión: permite contrastar lo que las familias declaran que quieren para sus hijos (escenario SAE Preferencias) con aquello que finalmente ocurre en términos de asignación (escenario SAE Admisión).
- Cambios Forzados vs. Cambios No Forzados: muestra las diferencias estructurales entre los cambios de escuela forzados y los No Forzados.

En términos generales, los *clusters* quedan definidos en gran medida por la dependencia administrativa de los establecimientos de origen y destino, quedando el resto de las variables de agrupación supeditadas a estas. Es decir, en la mayoría de los escenarios son cuatro los *clusters*, definidos por las transiciones: municipal - municipal, municipal - particular subvencionado, particular subvencionado - municipal y particular subvencionado - particular subvencionado.

En todas las aglomeraciones urbanas excepto en Rancagua - Machalí, en los casos de cambio No Forzado, al comparar el escenario Pre SAE con los escenarios SAE, se observa que estos últimos tienen mayor flujo (relativo) de cambios hacia establecimientos particular subvencionados, tanto en SAE Preferencias como en SAE Admisión. En tanto que la distribución de cambios según dependencia administrativa de establecimiento de origen, permanece constante a través de todos los escenarios. Es decir, se observa una fuga de estudiantes desde el sector público hacia el particular subvencionado, mismo fenómeno identificado por Canals (2015) (en la Región Metropolitana).

Para el caso de cambios Forzados, este efecto sólo se observa en Puerto Montt - Puerto Varas, en donde además, la preferencia por establecimientos particulares subvencionados (SAE Preferencias) es considerablemente mayor a la admisión (SAE Admisión), lo que sugiere que hay una demanda por EE particulares subvencionados que no está siendo cubierta por la oferta del sistema educativo¹. Esto responde a la menor oferta pública de vacantes en enseñanza media, brecha que fue resuelta con el lucro en educación, que incentivó que se establecieran EE particulares (subvencionados o pagados). Por ende, este patrón detectado no significa sólo una marcada preferencia por establecimientos particulares subvencionados, sino que también es consecuencia de una cualidad estructural del sistema educativo.

Exceptuando los cambios no Forzados en Rancagua - Machalí, en todos los escenarios y en todas las conurbaciones, se observa una mayor demanda (SAE Preferencias) por establecimientos con mejores resultados SIMCE, consistente con lo que encuentran Gallego y Hernando (2009); Chumacero et al. (2011) (en el periodo Pre SAE). Dado que la demanda excede la oferta, el sistema no puede cubrirla totalmente (SAE Admisión). Esto no ocurre simplemente porque el SAE sea deficiente en el proceso de asignación, sino más bien porque la capacidad de los establecimientos más preferidos es limitada y rígida. Sin embargo, se observa que la asignación a establecimientos de mayores resultados SIMCE mejora levemente con respecto al escenario Pre SAE, lo que concuerda con los resultados de Carrasco y Honey (2019), aunque desafía lo encontrado por Urzua (2019), quien plantea que los resultados SIMCE no siempre son relevantes en la elección de escuela bajo el sistema actual.

¹"Sistema educativo" hace referencia a la estructura formada por los diversos componentes involucrados en educación. No debe confundirse con "sistema de admisión escolar", el cual se refiere específicamente al mecanismo de asignación de establecimientos a estudiantes que cambian de escuela.

Asimismo, en concordancia con Makovec et al. (2010), en los casos de cambios Forzados, en todas las aglomeraciones urbanas y escenarios, hay mayor movilidad desde municipales que en los casos No Forzados. Complementariamente, en los casos No Forzados hay más cambios desde establecimientos particulares subvencionados que en los casos de cambio Forzado.

Por otra parte, en todas las conurbaciones estudiadas, los casos No Forzados exhiben una menor estructura o están menos claramente definidos respecto al caso Forzado, lo que se evidencia a través de los menores coeficientes de silueta de los *clusters* en estos casos, y su mayor número de *clusters*, en general. Esta menor estructura responde a la naturaleza más heterogénea de los cambios No Forzados, donde la decisión de cambio de escuela que toman las familias está condicionada por razones no observables. Sin embargo, esto no impide que los casos No Forzados sean caracterizados por el modelo. Contrariamente, el caso Forzado es más estructurado en los patrones que exhibe.

Se observa en todos los casos y escenarios que hay una importante correlación entre variables socioeconómicas de los estudiantes y el tipo de colegios hacia el cual se cambian (o prefieren): estudiantes con mayor Ingreso del Hogar, mayor Educación de los Padres, no Beneficiarios SEP, con mejores puntajes SIMCE, son los que forman parte y prefieren establecimientos con mayor rendimiento SIMCE, menos vulnerables (IVE) y de dependencia particular subvencionada, acorde con lo planteado por Canales et al. (2016); Villalobos y Valenzuela (2012); Hsieh y Urquiola (2006), entre otros. Más aún, los resultados están acorde a Eyzaguirre et al. (2019a), quien estudia esta relación también el escenario SAE.

Por ende, la segregación socioeconómica en estos primeros años del SAE sigue siendo muy marcada. Sin embargo, este efecto se ve atenuado en los escenarios SAE en todas las aglomeraciones urbanas estudiadas, es decir, el escenario SAE Admisión deja levemente mejor a las familias que el escenario Pre SAE (lo mismo encuentra Carrasco y Honey (2019)). Esto sugiere que el sistema está en una etapa de transición y ajuste a las nuevas reglas impuestas por el SAE, tendencia también identificada por Eyzaguirre et al. (2019a).

Por consiguiente, dado que este proceso está en sus primeras etapas, este estudio solo entrega conclusiones preliminares sobre los efectos más profundos que esta política pública está generando. En este sentido, futuras investigaciones que consideren más periodos de aplicación del SAE y un mayor número de territorios serán necesarias para dilucidar los efectos del nuevo sistema a nivel nacional.

Otro aspecto de la segregación socioeconómica, es la presencia de *clusters* de cambio de estudiantes de GSE alto en todas las aglomeraciones urbanas y en todos los escenarios de los casos de cambio No Forzado. Por un lado, en las conurbaciones Iquique - Alto Hospicio y Coquimbo - La Serena, estas agrupaciones aparecen en los escenarios SAE, lo que se explica por los cambios de dependencia administrativa desde particular subvencionado hacia particular pagado de algunos establecimientos, los que abarcan el 5% y 8% de la matrícula total de dichas conurbaciones en el año 2017 (ver Tablas 3.9 y 3.10), respectivamente. Para Iquique - Alto Hospicio, este cambio conlleva que el *cluster* de GSE alto, presente en el escenario Pre SAE del caso de cambio Forzado, es eliminado en los escenarios SAE, lo que ocurre producto de la salida de dichos establecimientos particulares subvencionados del SAE, retirando de la oferta educativa a aquellos establecimientos de copago alto que eran escogidos por los estudiantes de dicho *cluster*.

Por otro lado, se observa que en Puerto Montt - Puerto Varas, este *cluster* de GSE alto está presente en todos los escenarios del caso de cambio No Forzado, lo que indica la persistencia de la estructura de la demanda y oferta educativa local, caracterizada en parte por la movilidad escolar entre este tipo de establecimientos.

Asimismo, se destaca que en los casos de cambios No Forzados, los estudiantes tienen notoriamente menor rendimiento académico (variable Percentil Promedio de Notas), misma relación encontrada por Canals (2015), lo que sugiere que esta es una razón que está detrás de la decisión de cambiar de establecimiento educativo.

Todo lo anterior lleva a concluir, además, la presencia de una diferenciación importante en los patrones de cambio de escuela según caso, escenario y territorio. Se observa que distintos territorios tienen diferentes necesidades en términos de demanda educativa (escenario SAE Preferencias) y de capacidad de cada sistema escolar. Estas mismas diferencias hacen que la comparación entre territorios pierda sentido si no se consideran estas dimensiones en el análisis. Así, contextualizar las acciones de mejora hacia el sistema optimizaría esfuerzos y traería mayores beneficios a los estudiantes de todo el país. Por ende, futuros esfuerzos por entender los efectos de programas y políticas educativas escolares deben tomar en cuenta los contextos de los distintos territorios, y no solo mirar la información de manera agregada.

En este sentido, el presente trabajo tiene dos grandes contribuciones: (i) permite comparar este nuevo escenario con presencia del SAE, con lo que ocurría antes de la implementación del sistema, y (ii) propone una metodología nueva para estudiar cambios de escuela, directamente agrupando cambios de establecimientos según las características de estos, caracterizando la movilidad escolar mediante flujos de estudiantes entre determinados tipos de establecimientos.

Una arista de análisis de la situación que queda inexplorada en este trabajo, pues escapa a los objetivos del mismo, corresponde al sesgo de género que se observa en varios *clusters*, es decir, en determinadas agrupaciones hay un desbalance significativo en la proporción mujer - hombre con respecto a la proporción propia de la aglomeración urbana. Esto queda propuesto para ser desarrollado en un trabajo futuro.

Si bien este estudio hace un aporte en términos de la identificación y cuantificación del efecto de la Ley de Inclusión Escolar sobre las preferencias de los padres en las 4 aglomeraciones involucradas, tiene ciertas limitaciones que son importantes de considerar. En efecto, la muestra utilizada para la elaboración de modelos y el análisis no es representativa de la realidad nacional, pues sólo se centra en algunas zonas urbanas, lo que responde a la implementación gradual que ha tenido el SAE desde el año 2015 (sección 2.3.2). Por otra parte, el estudio llevado a cabo no es de carácter causal, nos informa bien sobre lo que ocurre, pero no acerca de las influencias de las distintas variables en los resultados. Un análisis de causalidad debe ser llevado a cabo en una etapa posterior del estudio, utilizando metodologías apropiadas para tal efecto.

En términos generales, se concluye que los territorios estudiados necesitan fortalecer su educación pública y mejorar sus resultados educacionales (medidos por el SIMCE), de tal manera que la oferta educativa pública mejore en esta dimensión y se frene la fuga de estudiantes hacia establecimientos particulares subvencionados. Más aún, la situación de preferencias de cambios no es algo sostenible en el tiempo, pues la Ley de Inclusión Escolar, por un lado, ha

suprimido en gran medida los incentivos a crear establecimientos particulares subvencionados (fin al lucro), y por otro, quita a este tipo de escuelas las herramientas que tenían para generar mejores resultados SIMCE: la selección y el copago. Urge entonces que el sistema educativo público retome un rol protagónico en términos de cobertura y calidad (medida por SIMCE).

En la misma línea, se propone también el desarrollo de políticas públicas focalizadas en contrarrestar los comportamientos de autoexclusión y autosegregación encontrados por Elacqua (2005); Canales et al. (2016); Gallego y Hernando (2009), entre otros. Para así revalorizar socialmente la educación pública, promoviendo su elección por parte de las familias, complementando este tipo de medidas con mejoras en calidad (medida por SIMCE).

Al mirar el comportamiento de las familias como tomadores de decisiones, se concluye que están buscando (mediante sus preferencias declaradas) sobre todo establecimientos con mayores resultados SIMCE. Sin embargo, se constata que las desigualdades en preferencias según nivel socioeconómico persisten, lo que también concluyen Eyzaguirre et al. (2019b); Carrasco et al. (2019).

En síntesis, los resultados del análisis realizado evidencian que en los procesos de cambio de escuela inciden las características de los establecimientos, los rasgos individuales de los estudiantes, el tipo de cambio (Forzado o No Forzado), y la particularidad de sus territorios. Aunque el SAE provee un escenario más justo en lo que respecta a elección de escuela, los resultados muestran que hay una brecha insalvable entre oferta y demanda que el sistema no puede satisfacer por ahora, entiendo que tampoco lo hacía antes.

Si bien las conclusiones se centran en observaciones de carácter cualitativo de los patrones identificados, los resultados encontrados con esta nueva metodología son coherentes en gran medida con lo que plantea la literatura, sobre todo con trabajos recientes sobre el SAE. Por consiguiente, los resultados de este trabajo tienen sentido en función de lo que se conoce sobre *school choice*, particularmente en Chile. Más aún, el presente trabajo provee una amplia base para el estudio más en profundidad de cada uno de los casos y escenarios considerados, así como la expansión de los mismos hacia otros territorios y contextos no considerados, como lo son, por ejemplo, la movilidad escolar en sectores rurales y la consideración del resto de los niveles escolares.

Bibliografía

- Abdulkadiroğlu, A., y Sonmez, T. (2003). School Choice: A Mechanism Design Approach. *American Economic Review*, 93(3), 729–747.
- Abdulkadiroğlu, A., Pathak, P. A., y Roth, A. E. (2005). The New York City High School Match. *The American Economic Review*, 95(2).
- Abdulkadiroğlu, A., Pathak, P. A., Roth, A. E., y Sönmez, T. (2005). The Boston Public School Match. *The American Economic Review*, 95(2), 368–372.
- Aguirre, J. (2018). Can Progressive Vouchers Help the Poor Benefit from School Choice? Evidence from the Chilean Voucher System. *SSRN Electronic Journal*, 1–51. doi: 10.2139/ssrn.3123670
- Arce-trigatti, B. P., Lincove, J. A., y Harris, D. N. (2016). *Is there school choice in school choice? School Program Offerings in the New Orleans Public Schools Market* (Inf. Téc.). Education Research Alliance for New Orleans.
- Baker, R. S. J. D. (2010). *Data Mining for Education* (Third Edit ed.). Elsevier.
- Baker, R. S. J. D. (2013). Learning, Schooling, and Data Analytics. En *Handbook on innovations in learning* (p. 179–190). Philadelphia,: M. Murphy, S. Redding, & J. Twyman. Descargado de <http://www.centeril.org/>
- Batista, G. E., y Monard, M. C. (2004). An analysis of four missing data treatment methods for supervised learning. *Applied Artificial Intelligence*, 17(5-6), 519–533. doi: 10.1080/713827181
- Bellei, C. (2013). El estudio de la segregación socioeconómica y académica de la educación chilena. *Estudios Pedagógicos XXXIX*, N° 1(1), 325–345. Descargado de <https://scielo.conicyt.cl/pdf/estped/v39n1/art19.pdf>
- Bernhardt, B. (2013). *Data analysis for continuous school improvement* (Third ed.; New York: Routledge, Ed.).
- Berry, B. Y. S., Levinsohn, J., y Pakes, A. (1995). Automobile Prices in Market Equilibrium. *Econometrica*, 63(4), 841–890.
- Biblioteca del Congreso Nacional de Chile. (2014). *Ley de Subvención Escolar Preferencial*.

Descargado de <https://www.leychile.cl/N?i=269001&f=2018-08-25&p=>

Biblioteca del Congreso Nacional de Chile. (2015). *Ley de Inclusión Escolar*. Descargado de <https://www.leychile.cl/N?i=1078172&f=2019-04-25&p=>

Bovo, A., Sanchez, S., Heguy, O., y Duthen, Y. (2013). Clustering Moodle data as a tool for profiling students. En *2nd international conference on e-learning and e-technologies in education* (pp. 121–126). Descargado de <https://hal.archives-ouvertes.fr/hal-01166940> doi: 10.1109/ICeLeTE.2013.6644359

Calsamiglia, C., Fu, C., y Güell, M. (2017). Structural Estimation of a Model of School Choices: The Boston Mechanism vs. Its Alternatives. *Journal of Political Economy*. doi: 10.1086/704573

Canales, M., Bellei, C., y Orellana, V. (2016). ¿Por qué elegir una escuela privada subvencionada? Sectores medios emergentes y elección de escuela en un sistema de mercado. *Estudios Pedagógicos*, 42(3), 89–109. doi: 10.4067/s0718-07052016000400005

Canals, C. (2015). *Cambio de Escuelas en el Gran Santiago* (Inf. Téc.). Santiago: Universidad de Chile.

Cao, L. (2017). Data science: A comprehensive overview. *ACM Computing Surveys*, 50(3). doi: 10.1145/3076253

Carrasco, A., y Honey, N. (2019). Nuevo Sistema de Admisión Escolar y su capacidad de atenuar la desigualdad de acceso a colegios de calidad: al inicio de un camino largo. *Estudios en Justicia Educativa*(01), 1–35. Descargado de <http://centrojusticiaeducacional.cl/wp-content/uploads/2019/03/estudios-n1.pdf>

Carrasco, A., Mizala, A., Contreras, D., y Santos, H. (2014). *Hacia un sistema escolar más inclusivo: Cómo reducir la segregación escolar en Chile* (Inf. Téc. n.º 3). Espacio Público.

Carrasco, A., Oyarzún, J. d. D., Bonilla, A., Honey, N., y Díaz, B. (2019). La experiencia de las familias con el nuevo sistema de admisión escolar: un cambio cultural. *Estudios en Justicia Educativa*, Centro de Justicia Educativa, 02(August), 1–29.

CASEN. (2002). *Encuesta de Caracterización Socioeconómica Nacional* (Inf. Téc.). Ministerio de Planificación y Cooperación.

Chubb, J., y Moe, T. (1991). Politics, Markets and Schools Politics, Markets and America's Schools. *British Journal of Sociology of Education*, 28(2), 257–264. doi: 10.1111/j.1467-9752.1994.tb00332.x

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. doi: 10.1016/j.econedurev.2011.05.015

Contreras, D., Sepúlveda, P., y Bustos, S. (2010). When schools are the ones that choose: The effects of screening in Chile. *Social Science Quarterly*, 91(5), 1349–1368. doi: 10.1111/

- Córdoba, C. (2006). Elección de escuela en sectores pobres de Chile. En *Simposio: Políticas educacionales y gestión de sistemas escolares - pens/edu*.
- Correa, J., Epstein, R., Escobar, J., Rios, I., Bahamondes, B., Bonet, C., ... Epstein, B. (2019). School choice in Chile. En *Proceedings of the 2019 acm conference on economics and computation* (pp. 325–343). doi: 10.1145/3328526.3329580
- Cox, C. (1997). La Reforma de la educación chilena. *Colección de Estudios Cieplan*, 1(45), 5–32. Descargado de <https://dialnet.unirioja.es/servlet/articulo?codigo=5420503>
- Dalwani, K. (2015). *Using Silhouette analysis for selecting the number of cluster for K-means clustering*. Descargado de <https://kapilddatascience.wordpress.com/2015/11/10/using-silhouette-analysis-for-selecting-the-number-of-cluster-for-k-means-clustering/>
- de Haan, M., Gautier, P. A., Oosterbeek, H., y Klaauw, B. v. d. (2016). The performance of school assignment mechanisms in practice. *IZA Discussion Paper*(9118), 38.
- Elacqua, G. (2005). *School Choice in Chile: An Analysis of Parental Preferences and Search Behavior* (Inf. Téc. n.º May). New York, NY, USA: Teachers College, Columbia University.
- Elacqua, G. (2009). International Journal of Educational Development The impact of school choice and public policy on segregation: Evidence from Chile. *International Journal of Educational Development*, 32(3), 444–453. Descargado de <http://dx.doi.org/10.1016/j.ijedudev.2011.08.003> doi: 10.1016/j.ijedudev.2011.08.003
- Elacqua, G., y Martínez, M. (2015). *Documentos de Trabajo 16 Searching for schools in a low quality market : Evidence from Chile* (n.º February). Santiago.
- Eyzaguirre, S., Hernando, A., Razmilic, S., Blanco, N., Figueroa, N., Tagle, F., y Icaran, R. (2019a). ¿Existe sesgo socioeconómico en las preferencias de colegios? Descargado de www.cepchile.cl
- Eyzaguirre, S., Hernando, A., Razmilic, S., Blanco, N., Figueroa, N., Tagle, F., y Icaran, R. (2019b). ¿Qué explica las diferencias socioeconómicas en las preferencias escolares? *Punto de Referencia*(514), 1–23.
- Friedman, M. (1955). The Role of Government in a Free Society. En *Capitalism and freedom* (pp. 85–107). University of Chicago Press.
- Friedman, M. (1962). *The Role of Government in Education*. Descargado de <http://la.utexas.edu/users/hcleaver/330T/350kPEEFriedmanRoleOfGovttable.pdf>
- Gale, D., y Shapley, L. (1962). College Admissions and the Stability of Marriage Author. *Mathematical Association of America*, 69(1), 9–15. doi: 10.2307/2312726

- Gallego, F., y Hernando, A. (2009). School choice in Chile: Looking at the Demand Side. *Documento de trabajo Instituto Economía UC*, 356(July), 1 - 46.
- Gašević, D., Jovanović, J., Pardo, A., y Dawson, S. (2017). Detecting Learning Strategies with Analytics: Links with Self-Reported Measures and Academic Performance. *Journal of Learning Analytics*, 4(2), 113–128. doi: 10.18608/jla.2017.42.10
- Gómez, D., Chumacero, R. A., y Paredes, R. D. (2012). School choice and information. *Estudios de Economía*, 39(2), 143–157. Descargado de <http://www.redalyc.org/resumen.oa?id=22125058003>
- Gutiérrez, G., Jerrim, J., y Torres, R. (2017). School segregation across the world: has any progress been made in reducing the separation of the rich from the poor? , 1–23. Descargado de https://johnjerrim.files.wordpress.com/2017/12/working_paper_international_segregation1.pdf
- Harris, R., Johnston, R., y Burgess, S. (2007). Neighborhoods, ethnicity and school choice: Developing a statistical framework for geodemographic analysis. *Population Research and Policy Review*, 26(5-6), 553–579. doi: 10.1007/s11113-007-9042-9
- Hastie, T., Tibshirani, R., y Friedman, J. (2001). *Analysis on critical nodes in controlling complex networks using dominating sets*. New York, NY, USA: Springer New York Inc. doi: 10.1109/SITIS.2013.106
- Hernández, M., y Raczynski, D. (2015). Elección de escuela en Chile: De las dinámicas de distinción y exclusión a la segregación socioeconómica del sistema escolar. *Estudios Pedagógicos*, XLI(2), 127–141. Descargado de <http://www.scielo.cl/pdf/estped/v41n2/art08.pdf> doi: 10.4067/S0718-07052015000200008
- Hsieh, C. T., y Urquiola, M. (2006). The effects of generalized school choice on achievement and stratification: Evidence from Chile’s voucher program. *Journal of Public Economics*, 90(8-9), 1477–1503. doi: 10.1016/j.jpubeco.2005.11.002
- Hurwitz, J., y Kirsch, D. (2018). *Machine Learning for dummies*. John Wiley & Sons, Inc.
- Instituto Nacional de Estadísticas. (2017). *Síntesis de Resultados CENSO 2017* (Inf. Téc.). Instituto Nacional de Estadísticas. Descargado de <https://www.censo2017.cl/descargas/home/sintesis-de-resultados-censo2017.pdf>
- Inzunza, J., y Campos, J. (2016). El SIMCE en Chile: Historia, Problematicación y Resistencia. En U. P. Nacional (Ed.), *Xi seminario internacional de la red estrado*.
- J. MacQUEEN. (1967). Some Methods for Classification and Analysis of Multivariate Observations. En *Proceedings of the symposium on mathematics and probability* (Vol. 1, pp. 281–297). Berkely: University of California Press, Berkeley. doi: 10.1007/s11665-016-2173-6
- Jyothi, N., Bhan, K., Mothukuri, U., Jain, S., y Jain, D. (2012). A recommender system assisting instructor in building learning path for personalized learning system. *Proceedings - 2012 IEEE 4th International Conference on Technology for Education, T4E 2012*, 228–

230. doi: 10.1109/T4E.2012.51

Klašnja-Milićević, A., Ivanović, M., y Budimac, Z. (2017). Data science in education: Big data and learning analytics. *Computer Applications in Engineering Education*, 25(6), 1066–1078. doi: 10.1002/cae.21844

Kleinberg, J., Ludwig, J., Mullainathan, S., y Obermeyer, Z. (2015). Prediction policy problems. *American Economic Review*, 105(5), 491–495. doi: 10.1257/aer.p20151023

Kojima, F., y Ünver, M. U. (2011). *The 'Boston'school choice mechanism*. Stanford, CA. Descargado de <https://fmwww.bc.edu/ec-p/wp729.pdf>

Lara, B., Mizala, A., y Repetto, A. (2011). The Effectiveness of Private Voucher Education: Evidence From Structural School Switches. *Educational Evaluation and Policy Analysis*, 33(2), 119–137. doi: 10.3102/0162373711402990

Lee, J. H., Shin, J., y Realff, M. J. (2018). Machine learning: Overview of the recent progresses and implications for the process systems engineering field. *Computers and Chemical Engineering*. doi: 10.1016/j.compchemeng.2017.10.008

Magana, A. J., y Boutin, M. (2018). A Principled Approach to Using Machine Learning in Qualitative Education Research. En *Proceedings - frontiers in education conference, fie* (pp. 1–7). IEEE. doi: 10.1109/FIE.2018.8659073

Mahajan, M., Nimbhorkar, P., y Varadarajan, K. (2009). The Planar k-means Problem is NP-hard. En *Theoretical computer science* (Vol. 442, p. 274 - 285). Descargado de <http://www.springerlink.com/index/10.1007/978-3-642-00202-1%5Cnhttp://link.springer.com/10.1007/978-3-642-00202-1> doi: 10.1007/978-3-642-00202-1

Makovec, M., Mizala, A., y Barrera, A. (2010). *Parental decisions in a choice based school system: Analyzing the transition between primary and secondary school*. Santiago.

Marwala, T. (2018). *Gaussian Mixture Models*. Descargado de <https://scikit-learn.org/stable/modules/mixture.html#gmm> doi: 10.1142/9789813271234{_}0013

McFadden, D. (1974). The measurement of urban travel demand. *Journal of Public Economics*, 3(4), 303–328. doi: 10.1016/0047-2727(74)90003-6

MINEDUC. (2015). *La Reforma Educacional está en marcha* (Inf. Téc.). Ministerio de Educación de Chile. Descargado de <http://sitios.mineduc.cl/CuentaP/files/assets/common/downloads/CuentaP.pdf>

MINEDUC. (2016). *Revisión OCDE para mejorar la efectividad del uso de recursos en las escuelas* (Vol. 17) (n.º 188).

MINEDUC. (2017). *El primer gran debate de la reforma educacional: Ley de Inclusión Escolar* (Inf. Téc.). Ministerio de Educación de Chile.

- Mizala, A. (2014). La Política Chilena de Elección de Colegios al Pizarrón. *Revista Anales de la Universidad de Chile, Séptima Se(7)*, 1–34. Descargado de <https://anales.uchile.cl/index.php/ANUC/issue/view/3184>
- Mizala, A., y Torche, F. (2012). Bringing the schools back in: The stratification of educational achievement in the Chilean voucher system. *International Journal of Educational Development*, 32(1), 132–144. Descargado de <http://dx.doi.org/10.1016/j.ijedudev.2010.09.004> doi: 10.1016/j.ijedudev.2010.09.004
- Mizala, A., y Torche, F. (2013). ¿Logra la subvención escolar preferencial igualar los resultados educativos? *Documento de Referencia 09*. Descargado de www.espaciopublico.cl
- Moya, R. (2016). *Selección del número óptimo de Clusters*. Descargado de <https://jarroba.com/seleccion-del-numero-optimo-clusters/>
- OCDE. (2018). *Estudios Económicos de la OCDE* (Vol. 18; Inf. Téc. n.º S2). doi: 10.1017/S1431927612008112
- OPECH. (2015). Análisis crítico de la Ley que regula la admisión de los y las estudiantes, elimina el financiamiento compartido y prohíbe el lucro en establecimientos educacionales que reciben aportes del Estado. *Foro Chileno por el Derecho a la Educación, Abril, 2015*(2011).
- Paredes, R., Opazo, M., Volante, P., y Zubizarreta, J. R. (2013). Shared financing in the public education in Chile. *Policy proposals to Chile, Centro de Políticas Públicas, Universidad Católica, Chile*.
- Peckham, T., y McCalla, G. (2012). Mining Student Behavior Patterns in Reading Comprehension Tasks. En *Proceedings of the 5th international conference on educational data mining, edm 2012* (pp. 87–94).
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., . . . Duchesnay, E. (2011). Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research*, 12, 2825–2830.
- Psaromiligkos, Y., Orfanidou, M., Kytagiias, C., y Zafiri, E. (2011). Mining log data for the analysis of learners' behaviour in web-based learning management systems. *Operational Research*, 11(2), 187–200. doi: 10.1007/s12351-008-0032-4
- Rodríguez, P., Valenzuela, J. P., y Suchan, K. (2016). *Determinando el acceso real de los estudiantes a establecimientos educacionales efectivos para generar políticas públicas que mejoren la provisión de educación de calidad*.
- Roth, A. E. (1982). The Economics of Matching : Stability and Incentives. *Mathematics of Operations Research*, 7(4), 617–628.
- Roth, A. E. (1984). The Evolution of the Labor Market for Medical Interns and Residents: A Case Study in Game Theory. *Journal of Political Economy*, 92(6), 991–1016. doi: 10.1086/261272

- Roth, A. E. (2008). What have we learned from market design? *Economic Journal*, 118(527), 285–310. doi: 10.1111/j.1468-0297.2007.02121.x
- Sancllemente, M. (2008). *Nomadismo escolar en el sistema educacional chileno (2003-2007)* (Inf. Téc.). Georgetown University.
- Shavelson, R. J. (1979). Applications of Cluster Analysis in Educational Research: looking for a needle in a haystack. *British Educational Research Journal*, 5(1), 45–53. doi: 10.1080/0141192790050105
- Sistema de Admisión Escolar. (2018). *Resultados del Proceso de Admisión Escolar 2017* (Inf. Téc.). Santiago, Chile: Ministerio de Educación.
- Suartana, I. M., y Hidayat, A. I. (2017). Analysis of New Student Selection using Clustering Algorithms. En *Iop conference series: Materials science and engineering*. doi: 10.1088/1757-899X/288/1/012079
- Subramanian, N. (2019). *Optimal k in K-means*. Descargado de <https://aiaspirant.com/optimal-k-in-k-menas/>
- The Learning Machine. (2019). *Unsupervised Learning: Clustering* (Vol. 505) (n.º 146-148). Descargado de <https://www.thelearningmachine.ai/kmeans>
- Thorndike, R. L. (1953). Who belongs in the family. *Psychometrika*, 18, 267–276. doi: 10.1007/BF02289263
- Tie, Z. X., Jin, R., Zhuang, H., y Wang, Z. Q. (2010). The research on teaching method of basics course of computer based on cluster analysis. En *Proceedings - 10th IEEE international conference on computer and information technology, cit-2010* (pp. 2001–2004). doi: 10.1109/CIT.2010.338
- Urzua, S. (2019). *Analysis of Chile's New Centralized School Admission System (working paper)*.
- Valenzuela, J. P., Villarroel, G., y Villalobos, C. (2013). Preferential School Subsidy Law (SEP): Some Preliminary Results of its Implementation. *Pensamiento Educativo: Revista de Investigación Educativa Latinoamericana*, 50(2), 113–131. doi: 10.7764/pel.50.2.2013.7
- Vargas, C. (2018). *Perfilamiento de estudiantes y escuelas a través de modelos de aprendizaje automático para la elección de escuelas en Chile* (Tesis Doctoral no publicada). Universidad Técnica Federico Santa María.
- Vickers, M., Rees, P., y Birkin, M. (2005). *Creating the National Classification of Census Output Areas: Data, Methods and Results*. Leeds. Descargado de <http://eprints.whiterose.ac.uk/5003/>
- Villalobos, C., y Valenzuela, J. P. (2012). Polarization and social cohesion in the Chilean school system. *Revista de Analisis Economico*, 27(2), 145–172. doi: 10.4067/

S0718-88702012000200005

- Wook, M., Yahaya, Y. H., Wahab, N., Isa, M. R. M., Awang, N. F., y Seong, H. Y. (2009). Predicting NDUM Student's Academic Performance Using Data Mining Techniques. En *International conference on computer and electrical engineering* (pp. 357–361). IEEE. doi: 10.1109/ICCEE.2009.168
- Xu, D., y Tian, Y. (2015). A Comprehensive Survey of Clustering Algorithms. *Annals of Data Science*, 2(2), 165–193. doi: 10.1007/s40745-015-0040-1

Apéndices

A. Descripción de Variables

Variables de establecimientos			
Variable	Descripción	Valores	Fuente
rbd	Rol de base de datos del establecimiento.	Entero	MINEDUC
Comuna	Código de comuna a la cual pertenece el establecimiento.	Entero	MINEDUC
Dependencia administrativa	Dependencia administrativa del establecimiento.	1: Municipal 2: Particular subvencionado 3: Particular Pagado 4: Corporación de administración delegada 5: Servicio local de educación	MINEDUC
Copago	Cobro mensual realizado a las familias por el establecimiento.	Entero	MINEDUC
IVE	Índice de Vulnerabilidad Escolar del Establecimiento.	Decimal entre 0 y 1	MINEDUC
SIMCE	Puntaje promedio SIMCE de lenguaje y matemáticas de los estudiantes del establecimiento	Entero	MINEDUC
NSE	Nivel socioeconómico del establecimiento	1: bajo 2: medio bajo 3: medio 4: medio alto 5: alto	MINEDUC
lat-lon rbd	Coordenadas geográficas del establecimiento	grados, minutos, segundos	MINEDUC
Es rural	Indica si el establecimiento está ubicado en una zona rural.	0: no es rural 1: es rural	MINEDUC
Nivel máximo	Nivel máximo que imparte el establecimiento	1: prekinder 0: kinder 1: primero básico 2: segundo básico 3: tercero básico 4: cuarto básico 5: quinto básico 6: sexto básico 7: séptimo básico 8: octavo básico 9: primero medio 10: segundo medio 11: tercero medio 12: cuarto medio	MINEDUC

Tabla 6.1: Variables de establecimientos.

Variables de estudiantes			
Variable	Descripción	Valores	Fuente
mrun	Identificador del estudiante.	Entero	MINEDUC
Promedio general	Promedio de notas.	Decimal entre 1,0 y 7,0	MINEDUC
Género Alumno	Género del estudiante.	1: Masculino 2: Femenino	MINEDUC
Curso	Nivel que cursa el estudiante.	-1: prekinder 0: kinder 1: primero básico 2: segundo básico 3: tercero básico 4: cuarto básico 5: quinto básico 6: sexto básico 7: séptimo básico 8: octavo básico 9: primero medio 10: segundo medio 11: tercero medio 12: cuarto medio	MINEDUC
Beneficiario SEP	Indica si el alumno es beneficiario de la ley de Subvención Escolar Preferencial.	0: no es beneficiario 1: si es beneficiario	MINEDUC
latitud-longitud estudiante	Coordenadas geográficas con error del hogar del estudiante.	grados, minutos, segundos	MINEDUC
Puntaje SIMCE	Puntaje promedio en pruebas de lenguaje y matemáticas de estudiante.	Entero	MINEDUC
Distancia hogar-colegio	Distancia desde el hogar del estudiante a su establecimiento.	Entero, medida en metros	MINEDUC/ CIAE
Estudio de los Padres	Nivel de estudio promedio de los padres del estudiante.	1: No estudió 2: 1er año de educación básica 3: 2o año de educación básica 4: 3o año de educación básica 5: 4o año de educación básica 6: 5o año de educación básica 7: 6o año de educación básica 8: 7o año de educación básica 9: 8o año de educación básica 10: I año de educación media 11: II año de educación media 12: III año de educación media 13: IV año de educación media científico-humanista 14: IV o V año de educación media técnico-profesional o vocacional 15: Educación incompleta en un centro de FT o IP 16: Titulado de un centro de FT o IP 17: Educación incompleta de una universidad 18: Educación universitaria completa 19: Grado de magister universitario 20: Grado de doctor universitario	MINEDUC
Ingreso del Hogar	Ingreso total del hogar del estudiante.	1: Menos de \$100.000 2: Entre \$100.001 y \$200.000 3: Entre \$200.001 y \$300.000 4: Entre \$300.001 y \$400.000 5: Entre \$400.001 y \$500.000 6: Entre \$500.001 y \$600.000 7: Entre \$600.001 y \$800.000 8: Entre \$800.001 y \$1.000.000 9: Entre \$1.000.001 y \$1.200.000 10: Entre \$1.200.001 y \$1.400.000 11: Entre \$1.400.001 y \$1.600.000 12: Entre \$1.600.001 y \$1.800.000 13: Entre \$1.800.001 y \$2.000.000 14: Entre \$2.000.001 y \$2.200.000 15: Más de \$2.200.000	MINEDUC

Tabla 6.2: Variables de estudiantes.

Variables de cambio de establecimientos			
Variable	Descripción	Valores	Fuente
Cambio	Indica si el estudiante cambia de establecimiento.	0: No se cambia 1: Se cambia	MINEDUC
Cambio Forzado	Indica si el estudiante está forzado a cambiarse de establecimiento.	0: Cambio No Forzado 1: Cambio Forzado	MINEDUC

Tabla 6.3: Variables de cambio de escuela.

B. Representatividad de Variables

Conurbación	Caso	Escenario	Beneficiario SEP	Estudio de los Padres	Ingreso del Hogar	Puntaje SIMCE alumno	Promedio de Notas	Género	Distancia Hogar-Colegio
Iquique - Alto Hospicio	Forzado	Pre SAE	6155	4855	4788	4754	6155	6155	1297
		SAE Preferencias	2575	1788	1763	1770	2575	2575	2399
	No Forzado	SAE Admisión	2580	1791	1766	1773	2580	2580	2403
		Pre SAE	6465	4828	4760	4820	6465	6465	1029
Coquimbo - La Serena	Forzado	SAE Preferencias	2782	1486	1482	1500	2590	2782	2347
		SAE Admisión	2783	1492	1488	1506	2589	2783	2347
	No Forzado	Pre SAE	9685	8312	8199	8102	9685	9685	1627
		SAE Preferencias	3720	3177	3122	3109	3720	3720	3332
Rancagua - Maachali	Forzado	SAE Admisión	3730	3187	3132	3120	3730	3730	3344
		Pre SAE	8841	6880	6783	6870	8841	8841	1477
	No Forzado	SAE Preferencias	3875	1765	1738	1759	3679	3875	3011
		SAE Admisión	3871	1763	1737	1757	3677	3871	3016
Puerto Montt- Puerto Varas	Forzado	Pre SAE	11807	10524	10403	10341	11807	11807	879
		SAE Preferencias	5289	4749	4679	4674	5111	5289	5084
	No Forzado	SAE Admisión	5265	4725	4656	4650	5087	5265	5061
		Pre SAE	6509	5192	5113	5195	6509	6509	474
Puerto Montt- Puerto Varas	Forzado	SAE Preferencias	3143	1239	1231	1233	2899	3143	2648
		SAE Admisión	3137	1235	1227	1229	2900	3137	2648
	No Forzado	Pre SAE	8802	7307	7244	7054	8802	8802	1284
		SAE Preferencias	4778	4239	4203	4080	4777	4778	4399
Puerto Montt- Puerto Varas	Forzado	SAE Admisión	4783	4247	4211	4086	4782	4783	4402
		Pre SAE	3938	3022	2976	2985	3938	3938	429
	No Forzado	SAE Preferencias	1786	892	873	876	1678	1786	1319
		SAE Admisión	1791	895	875	879	1683	1791	1321

Conurbación	Caso	Escenario	Variables de Establecimientos de origen y destino con las que se entrenan los modelos de clustering
Iquique - Alto Hospicio	Forzado	Pre SAE	6155
		SAE Preferencias	2575
		SAE Admisión	2580
	No Forzado	Pre SAE	6465
		SAE Preferencias	2782
		SAE Admisión	2783
Coquimbo - La Serena	Forzado	Pre SAE	9685
		SAE Preferencias	3720
		SAE Admisión	3730
	No Forzado	Pre SAE	8841
		SAE Preferencias	3875
		SAE Admisión	3871
Rancagua - Maachalí	Forzado	Pre SAE	11807
		SAE Preferencias	5289
		SAE Admisión	5265
	No Forzado	Pre SAE	6509
		SAE Preferencias	3143
		SAE Admisión	3137
Puerto Montt- Puerto Varas	Forzado	Pre SAE	8802
		SAE Preferencias	4778
		SAE Admisión	4783
	No Forzado	Pre SAE	3938
		SAE Preferencias	1786
		SAE Admisión	1791

C. Tratamiento de Datos Faltantes

Variable	Método
IVE	- Utilizar IVE del mismo establecimiento del año anterior. - Tomar valor de base de datos del centro de estudios MINEDUC que entrega copago por tramos.
Copago	-Utilizar valor del mismo establecimiento del año anterior. - Establecimientos particulares pagados con valor %100.000
SIMCE Establecimiento	- En años no disponibles se utiliza valor del mismo establecimiento del año anterior, para mismo nivel.
Beneficiario SEP	- Si estudiante no aparece en base de datos de alumnos beneficiarios SEP, se considera como no beneficiario.

D. Optimización del Número de *Clusters*

D.1. Iquique - Alto Hospicio

Cambios Forzados

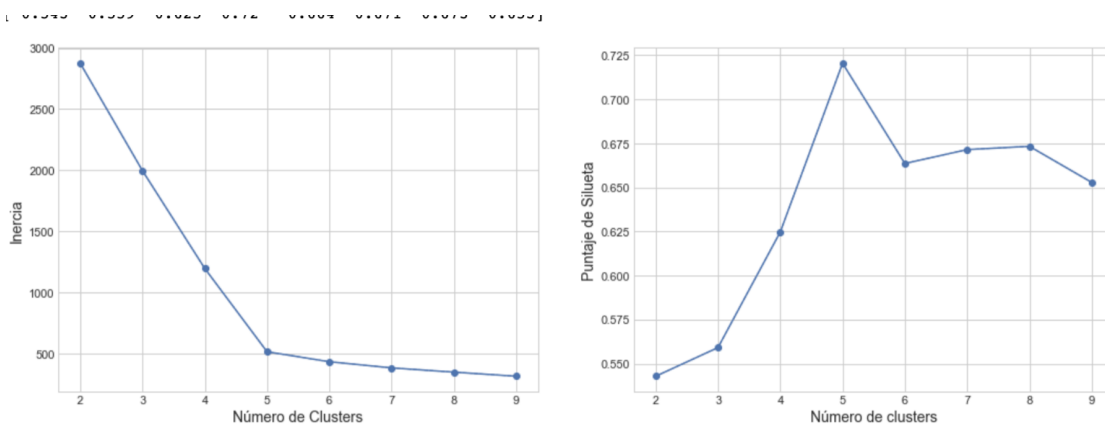


Figura 6.1: Inercia y coeficiente de silueta. Caso Iquique - Alto Hospicio, cambios Forzados, Pre SAE.

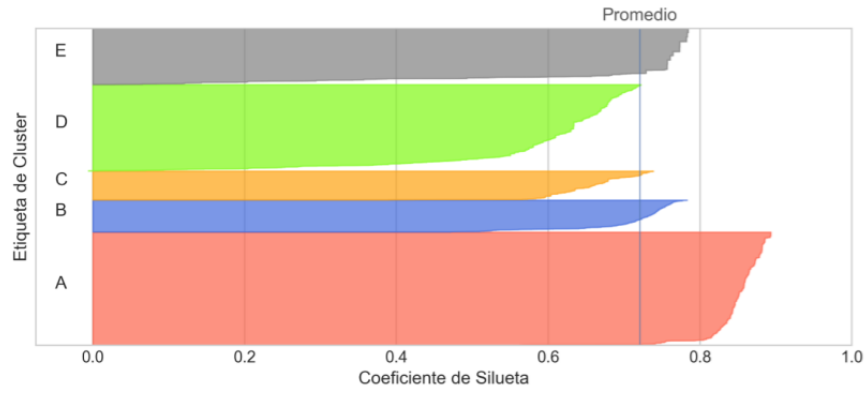


Figura 6.2: Coeficiente de silueta por *cluster*. Caso óptimo. Iquique - Alto Hospicio, cambios Forzados, Pre SAE.

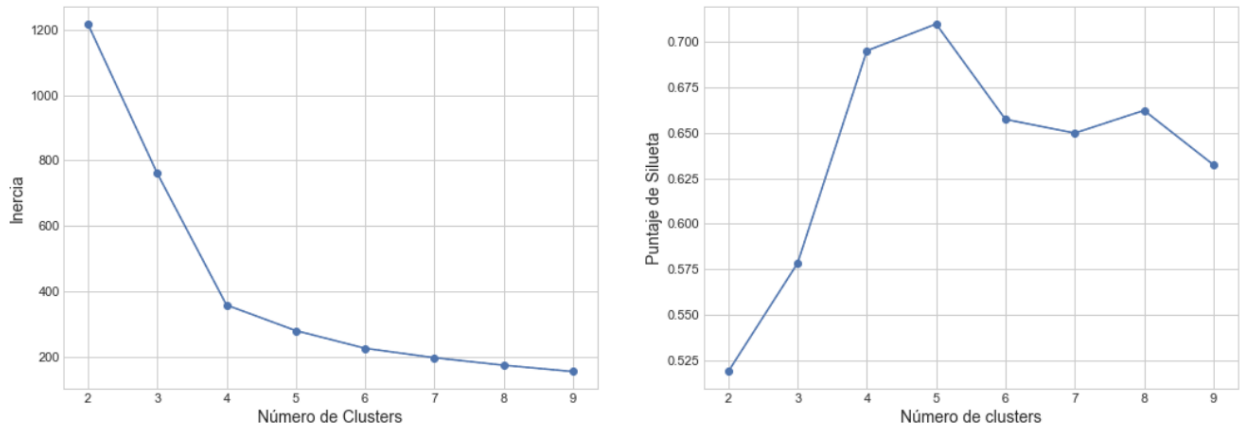


Figura 6.3: Inercia y coeficiente de silueta. Caso Iquique - Alto Hospicio, cambios Forzados, SAE Preferencias.

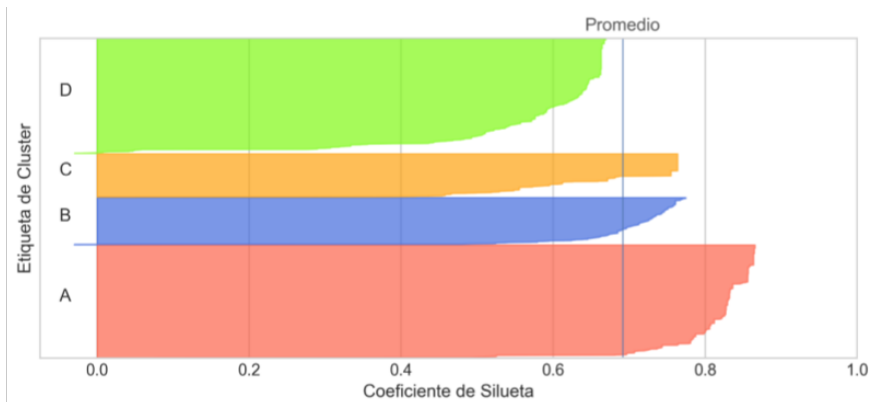


Figura 6.4: Coeficiente de silueta por *cluster*. Caso óptimo. Iquique - Alto Hospicio, cambios Forzados, SAE Preferencias.

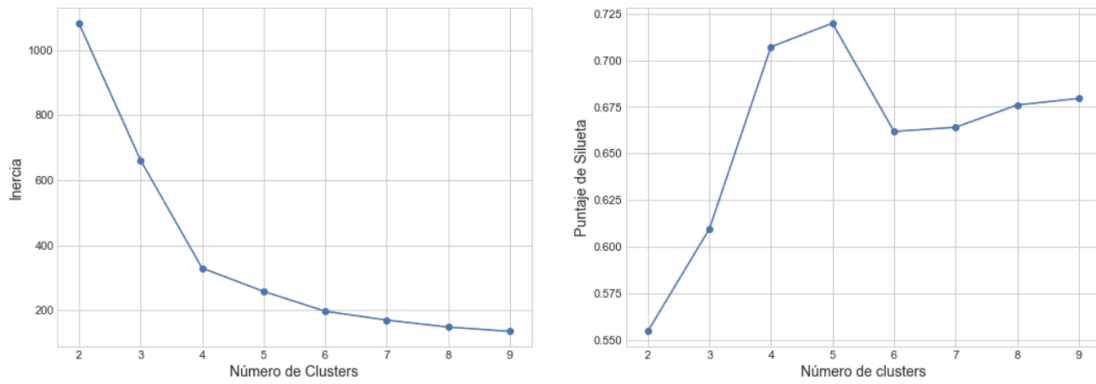


Figura 6.5: Inercia y coeficiente de silueta. Caso Iquique - Alto Hospicio, cambios Forzados, SAE Admisión.

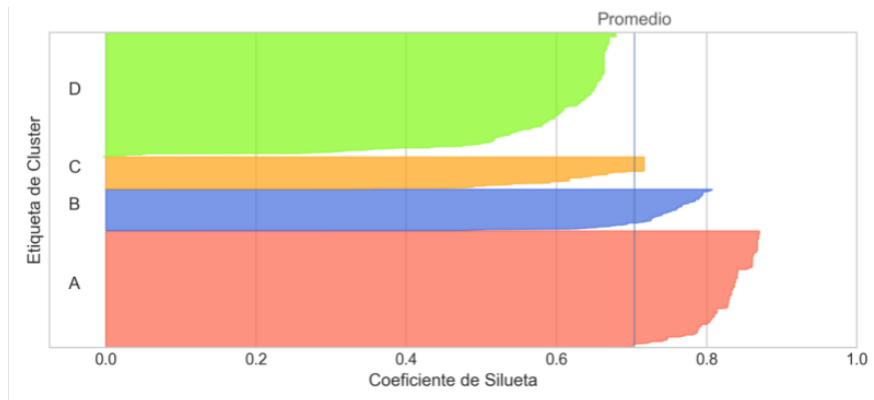


Figura 6.6: Coeficiente de silueta por *cluster*. Caso óptimo. Iquique - Alto Hospicio, cambios Forzados, SAE Admisión.

Cambios No Forzados

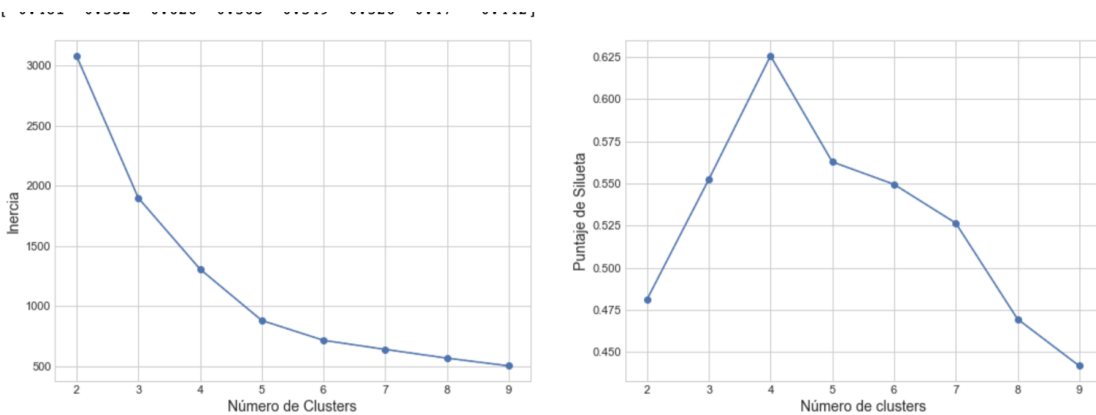


Figura 6.7: Inercia y coeficiente de silueta. Caso Iquique - Alto Hospicio, cambios No Forzados, Pre SAE.

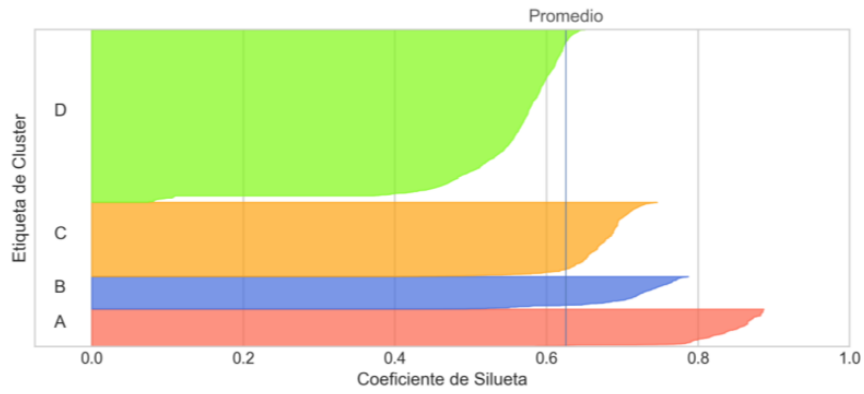


Figura 6.8: Coeficiente de silueta por *cluster*. Caso óptimo. Iquique - Alto Hospicio, cambios No Forzados, Pre SAE.

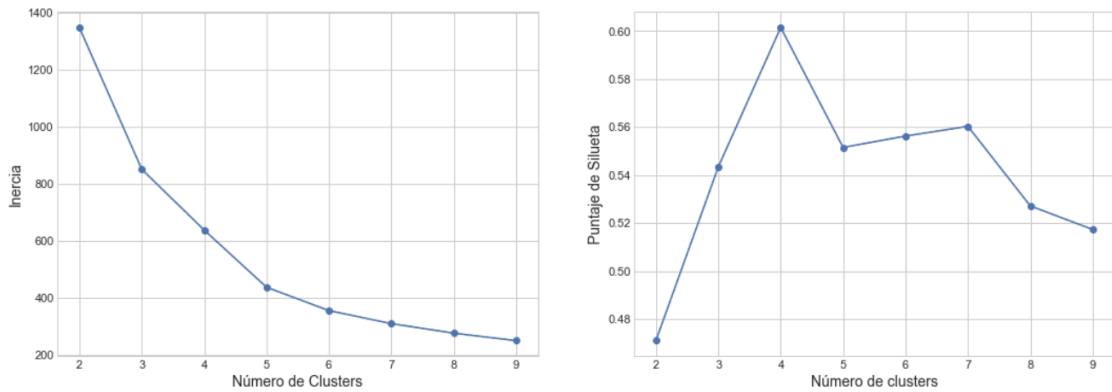


Figura 6.9: Inercia y coeficiente de silueta. Caso Iquique - Alto Hospicio, cambios No Forzados, SAE Preferencias.

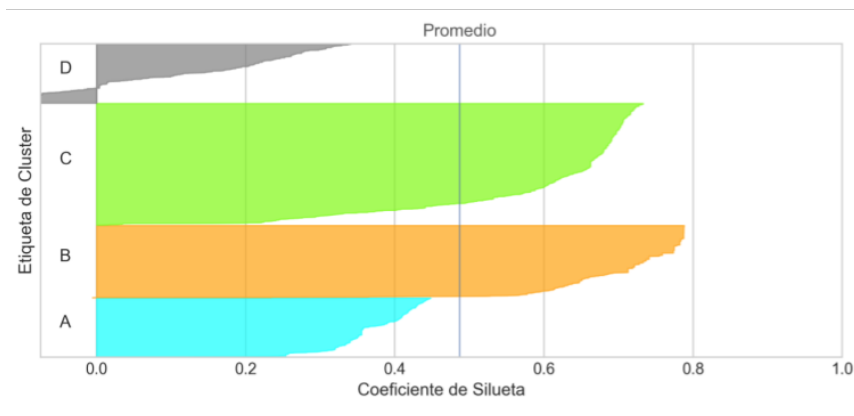


Figura 6.10: Coeficiente de silueta por *cluster*. Caso óptimo. Iquique - Alto Hospicio, cambios No Forzados, SAE Preferencias.

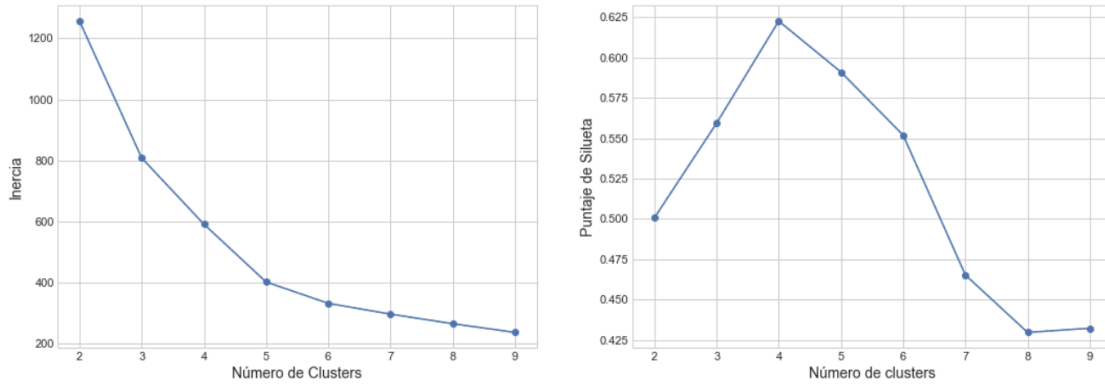


Figura 6.11: Inercia y coeficiente de silueta. Caso Iquique - Alto Hospicio, cambios No Forzados, SAE Admisión.

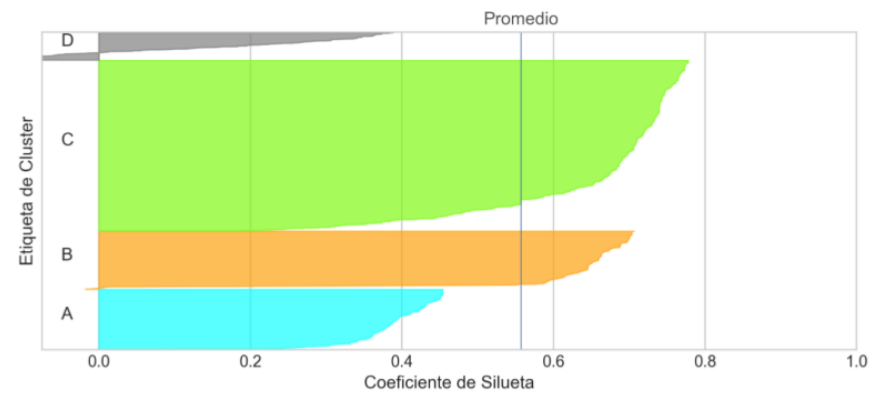


Figura 6.12: Coeficiente de silueta por *cluster*. Caso óptimo. Iquique - Alto Hospicio, cambios No Forzados, SAE Admisión.

D.2. Coquimbo - La Serena

Cambios Forzados

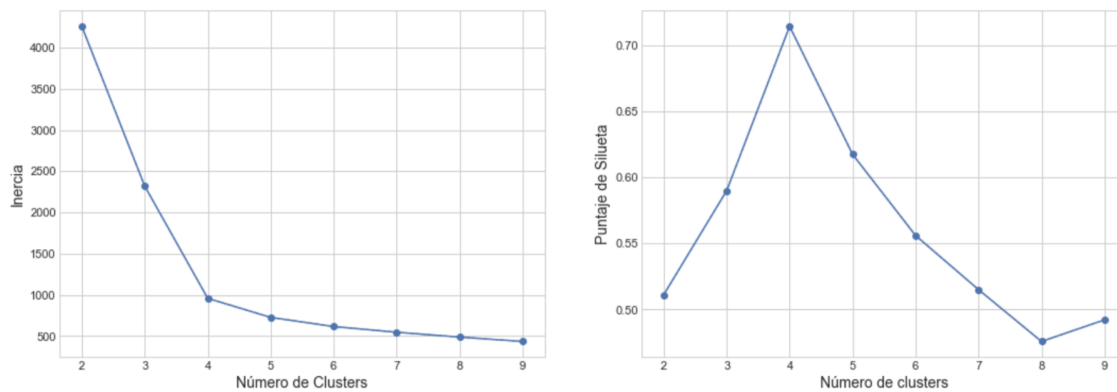


Figura 6.13: Inercia y coeficiente de silueta. Caso Coquimbo - La Serena, cambios Forzados, Pre SAE.

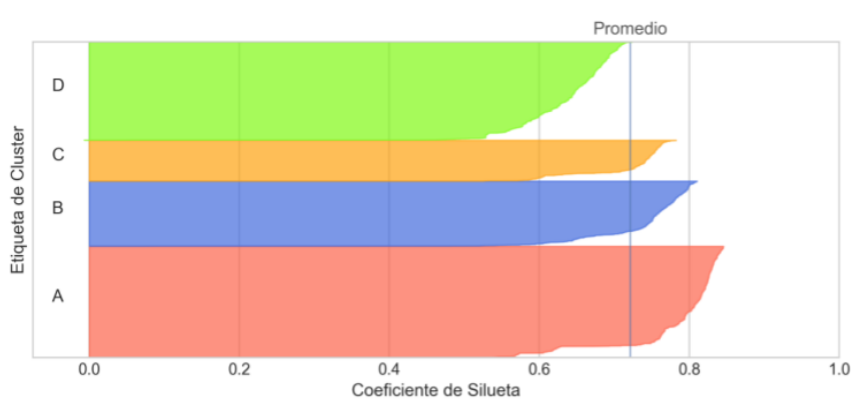


Figura 6.14: Coeficiente de silueta por *cluster*. Caso óptimo. Coquimbo - La Serena, cambios Forzados, Pre SAE.

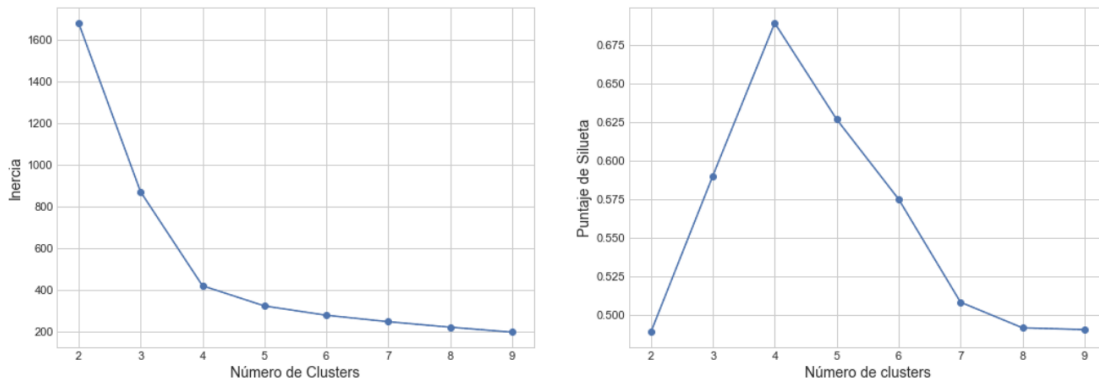


Figura 6.15: Inercia y coeficiente de silueta. Caso Coquimbo - La Serena, cambios Forzados, SAE Preferencias.

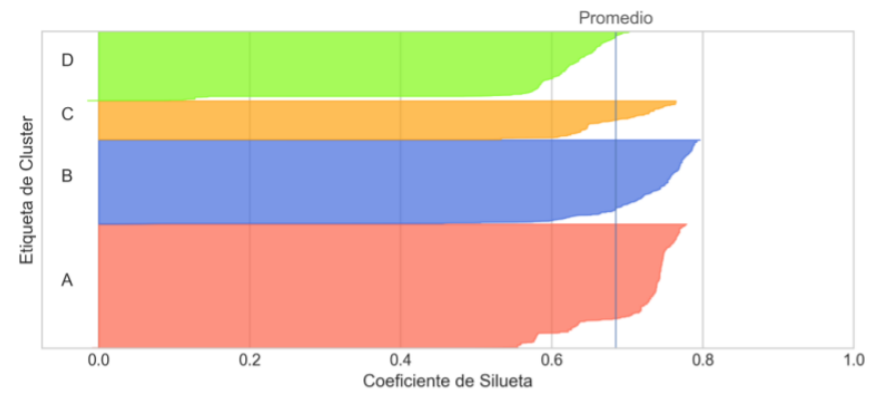


Figura 6.16: Coeficiente de silueta por *cluster*. Caso óptimo. Coquimbo - La Serena, cambios Forzados, SAE Preferencias.

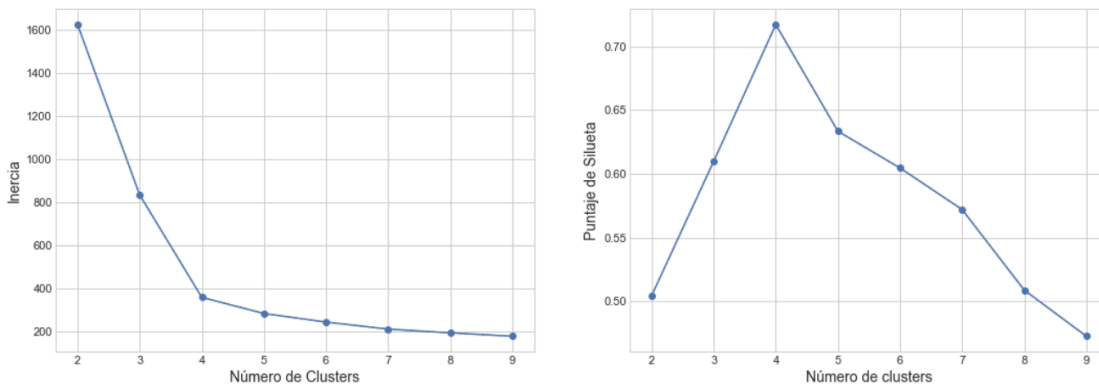


Figura 6.17: Inercia y coeficiente de silueta. Caso Coquimbo - La Serena, cambios Forzados, SAE Admisión.

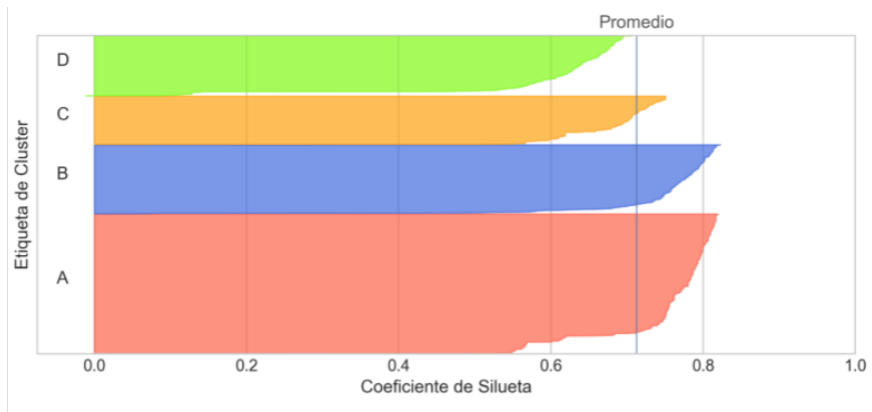


Figura 6.18: Coeficiente de silueta por *cluster*. Caso óptimo. Coquimbo - La Serena, cambios Forzados, SAE Admisión.

Cambios No Forzados

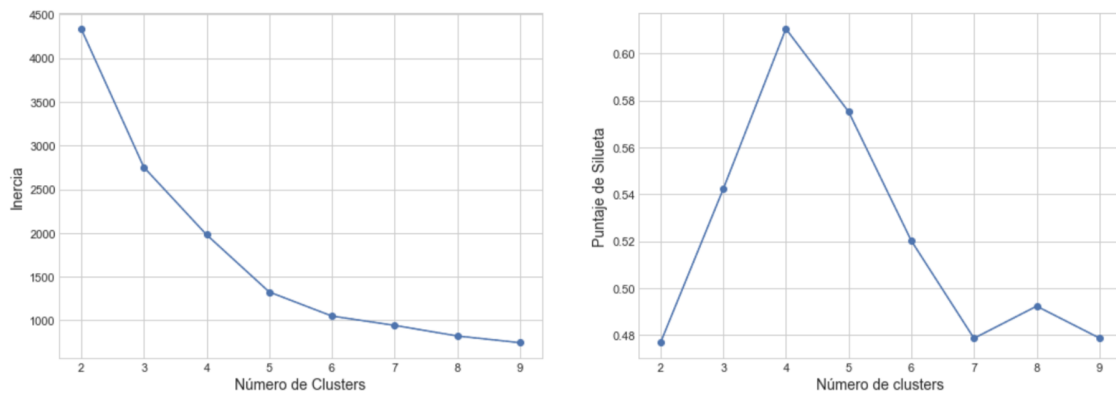


Figura 6.19: Inercia y coeficiente de silueta. Caso Coquimbo - La Serena, cambios No Forzados, Pre SAE.

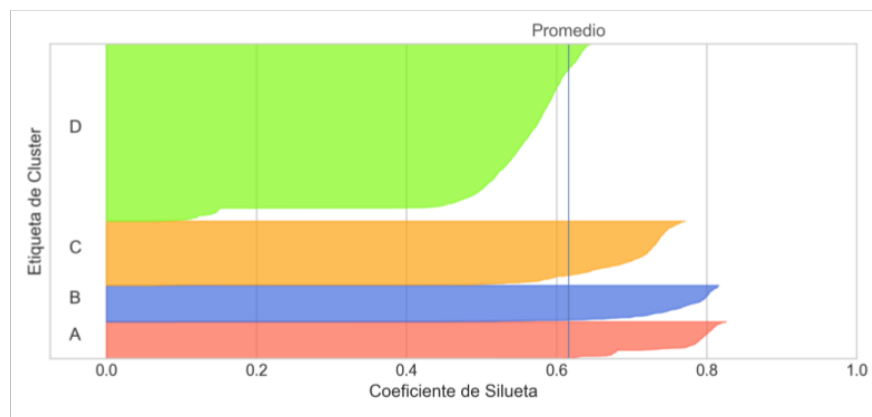


Figura 6.20: Coeficiente de silueta por *cluster*. Caso óptimo. Coquimbo - La Serena, cambios No Forzados, Pre SAE.

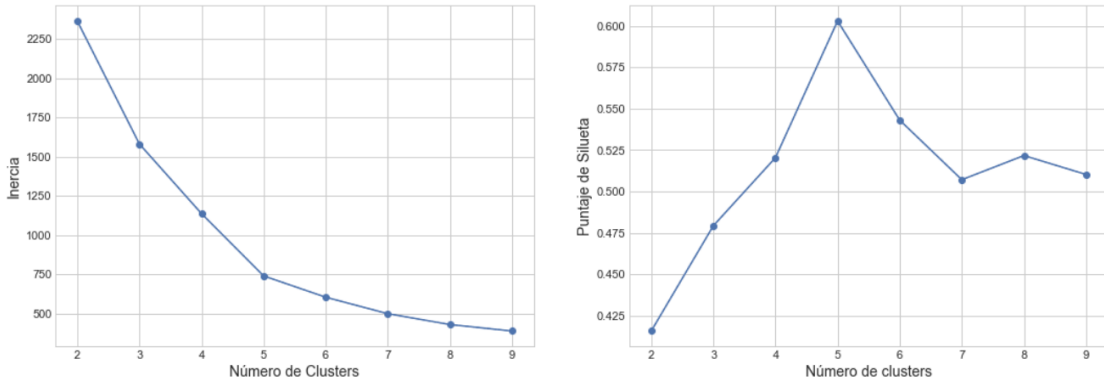


Figura 6.21: Inercia y coeficiente de silueta. Caso Coquimbo - La Serena, cambios No Forzados, SAE Preferencias.

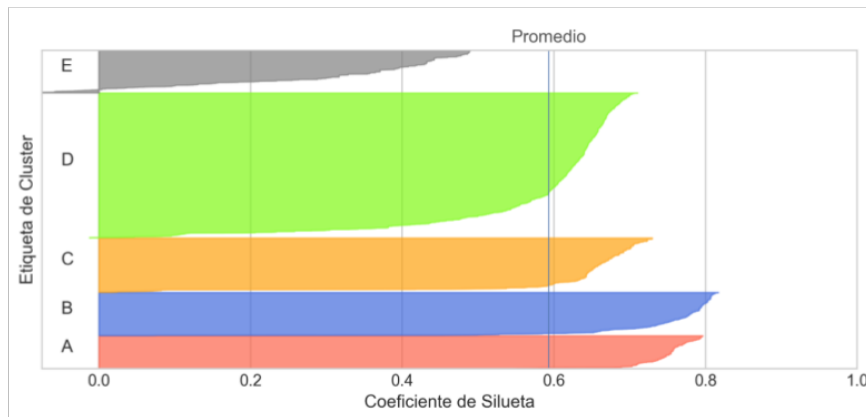


Figura 6.22: Coeficiente de silueta por *cluster*. Caso óptimo. Coquimbo - La Serena, cambios No Forzados, SAE Preferencias.

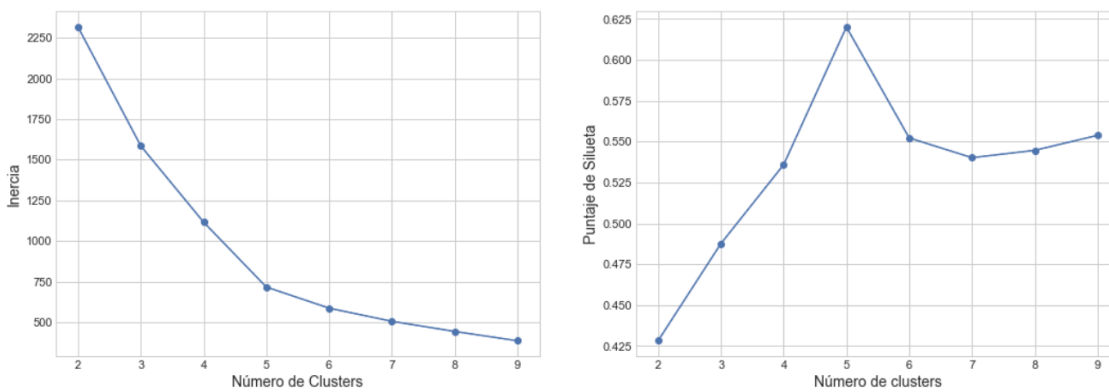


Figura 6.23: Inercia y coeficiente de silueta. Caso Coquimbo - La Serena, cambios No Forzados, SAE Admisión.

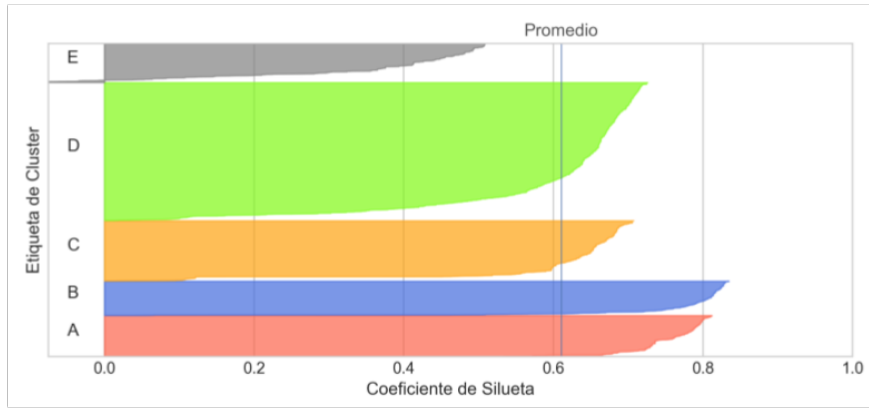


Figura 6.24: Coeficiente de silueta por *cluster*. Caso óptimo. Coquimbo - La Serena, cambios No Forzados, SAE Admisión.

D.3. Rancagua - Machalí

Cambios Forzados

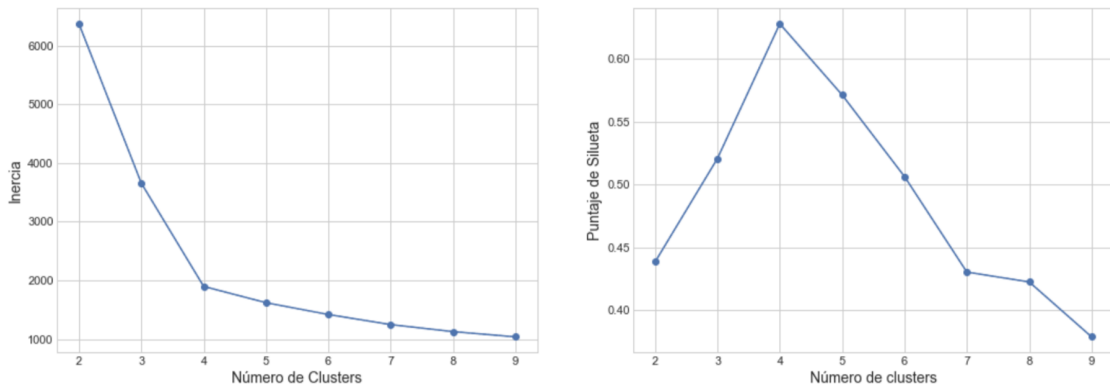


Figura 6.25: Inercia y coeficiente de silueta. Caso Rancagua - Machalí, cambios Forzados, Pre SAE.

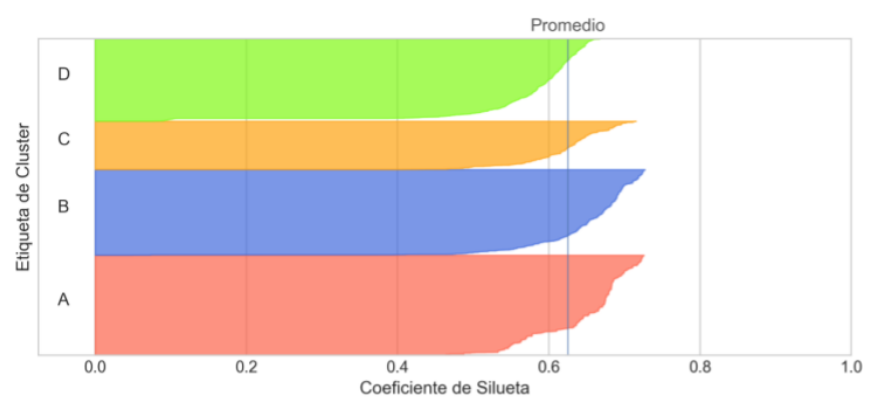


Figura 6.26: Coeficiente de silueta por *cluster*. Caso óptimo. Rancagua - Machalí, cambios Forzados, Pre SAE.

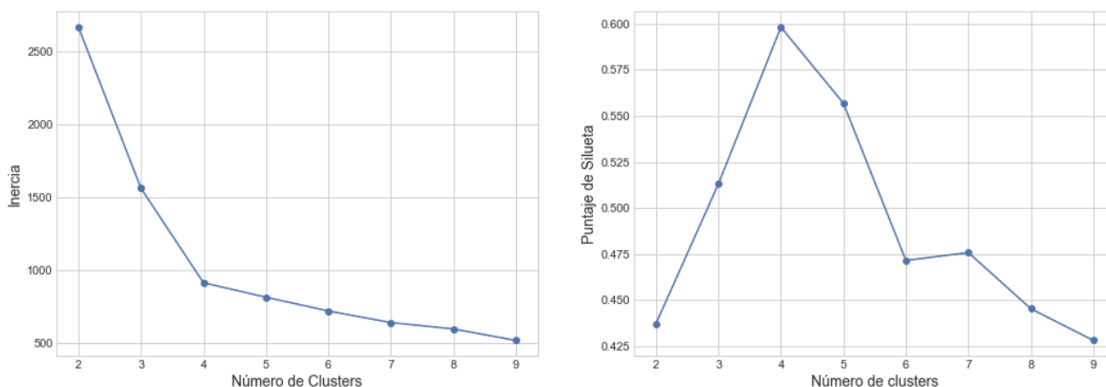


Figura 6.27: Inercia y coeficiente de silueta. Caso Rancagua - Machalí, cambios Forzados, SAE Preferencias.

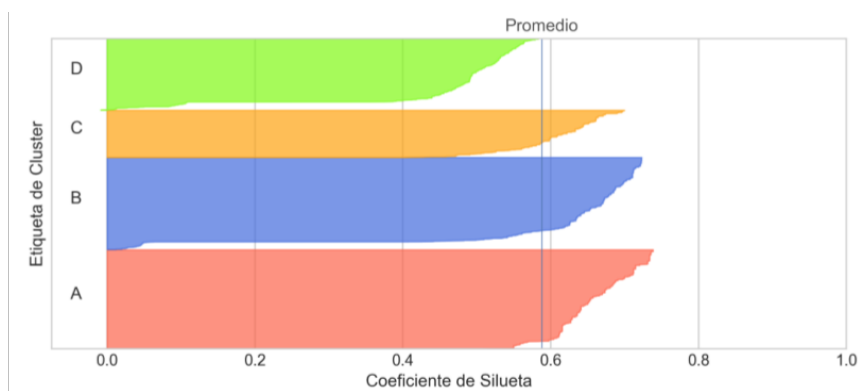


Figura 6.28: Coeficiente de silueta por *cluster*. Caso óptimo. Rancagua - Machalí, cambios Forzados, SAE Preferencias.

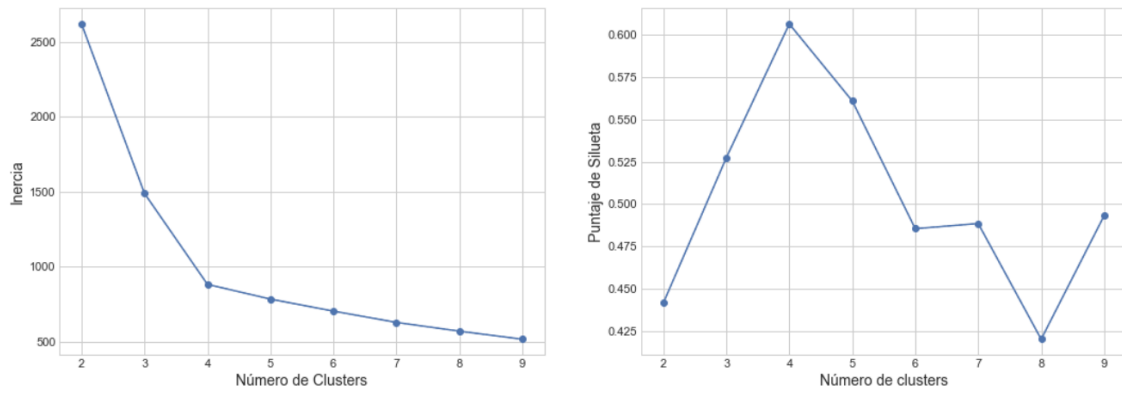


Figura 6.29: Inercia y coeficiente de silueta. Caso Rancagua - Machalí, cambios Forzados, SAE Admisión.

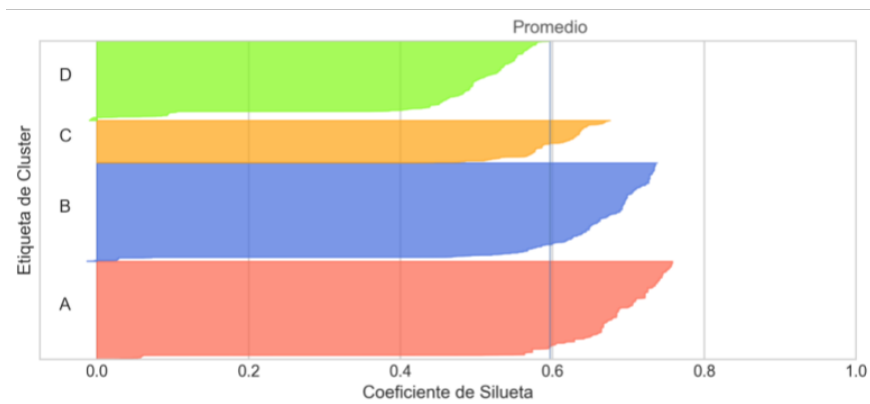


Figura 6.30: Coeficiente de silueta por *cluster*. Caso óptimo. Rancagua - Machalí, cambios Forzados, SAE Admisión.

Cambios No Forzados

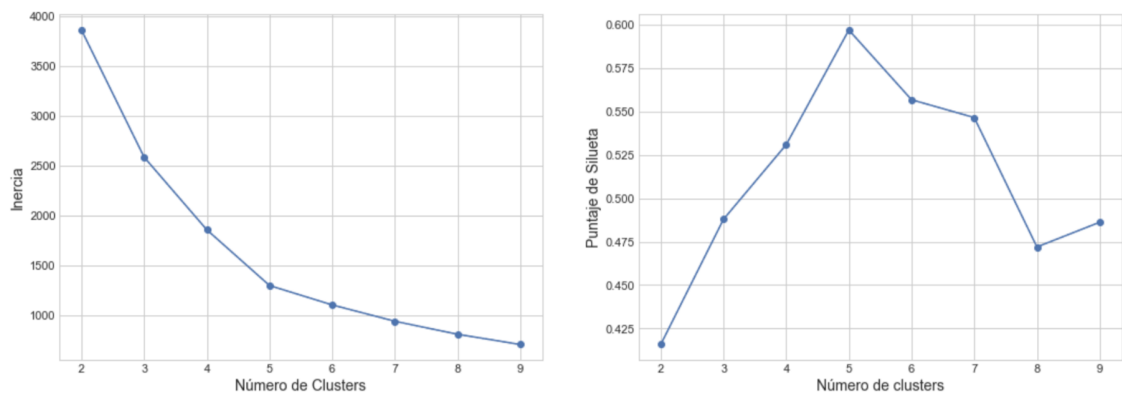


Figura 6.31: Inercia y coeficiente de silueta. Caso Rancagua - Machalí, cambios No Forzados, Pre SAE.

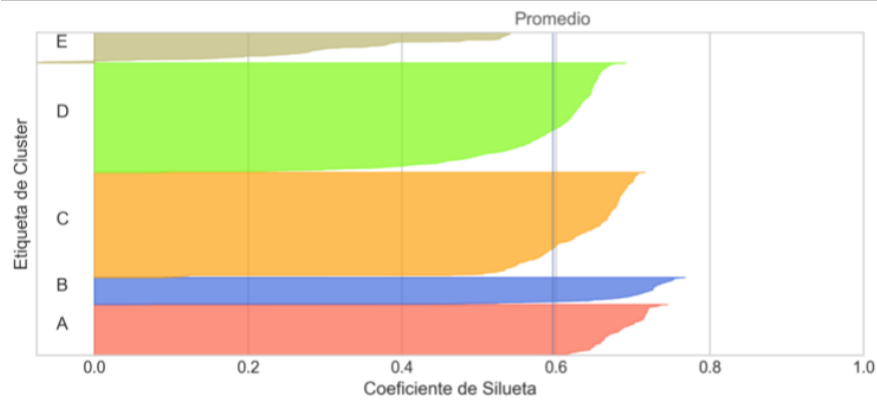


Figura 6.32: Coeficiente de silueta por *cluster*. Caso óptimo. Rancagua - Machalí, cambios No Forzados, Pre SAE.

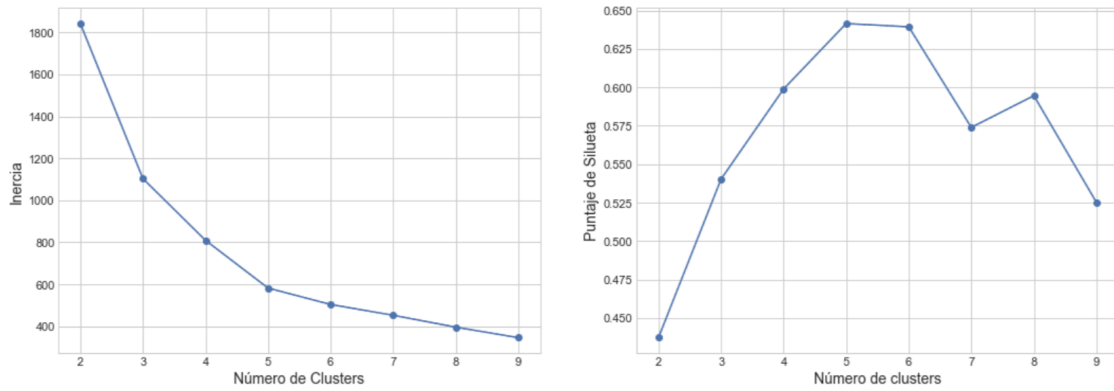


Figura 6.33: Inercia y coeficiente de silueta. Caso Rancagua - Machalí, cambios No Forzados, SAE Preferencias.

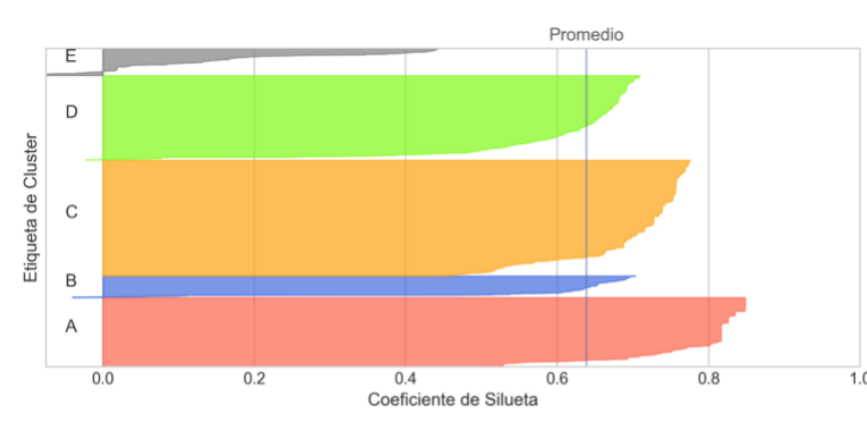


Figura 6.34: Coeficiente de silueta por *cluster*. Caso óptimo. Rancagua - Machalí, cambios No Forzados, SAE Preferencias.

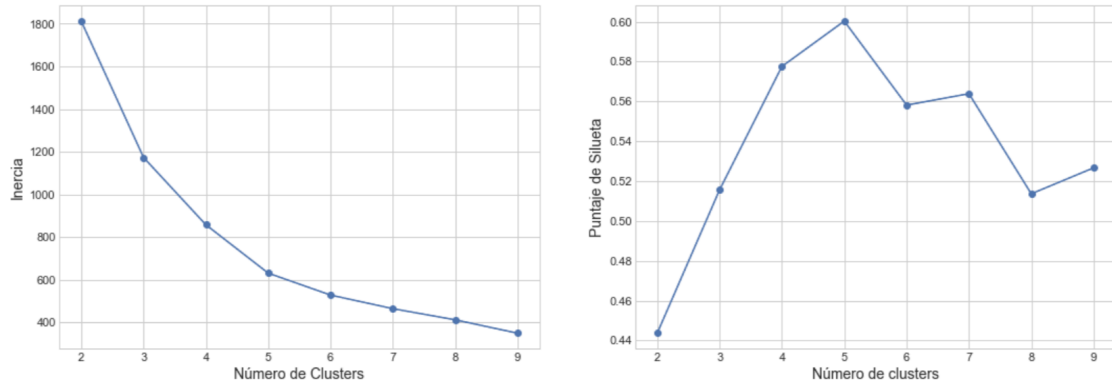


Figura 6.35: Inercia y coeficiente de silueta. Caso Rancagua - Machalí, cambios No Forzados, SAE Admisión.

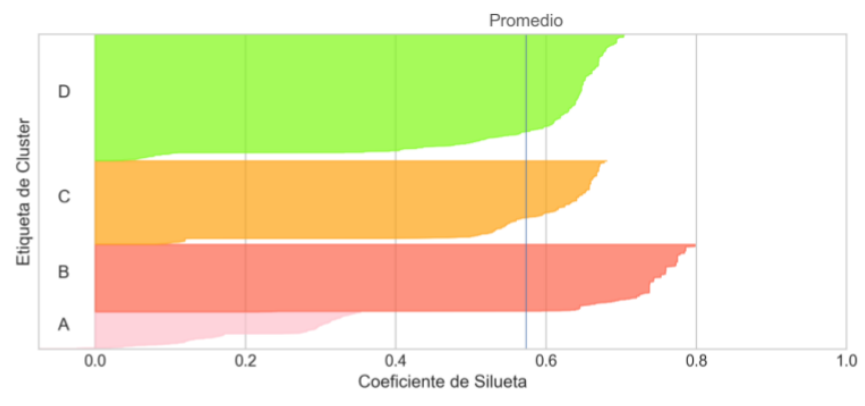


Figura 6.36: Coeficiente de silueta por *cluster*. Caso óptimo. Rancagua - Machalí, cambios No Forzados, SAE Admisión.

D.4. Puerto Montt - Puerto Varas

Cambios Forzados

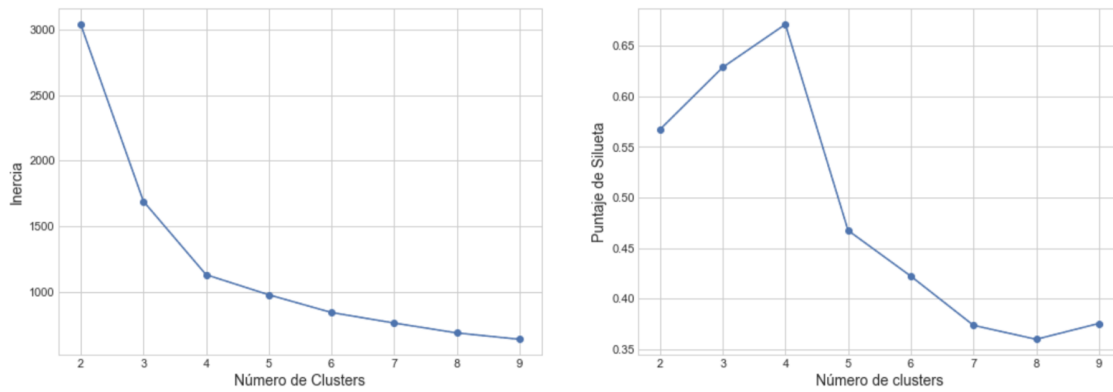


Figura 6.37: Inercia y coeficiente de silueta. Caso Puerto Montt - Puerto Varas, cambios Forzados, Pre SAE.

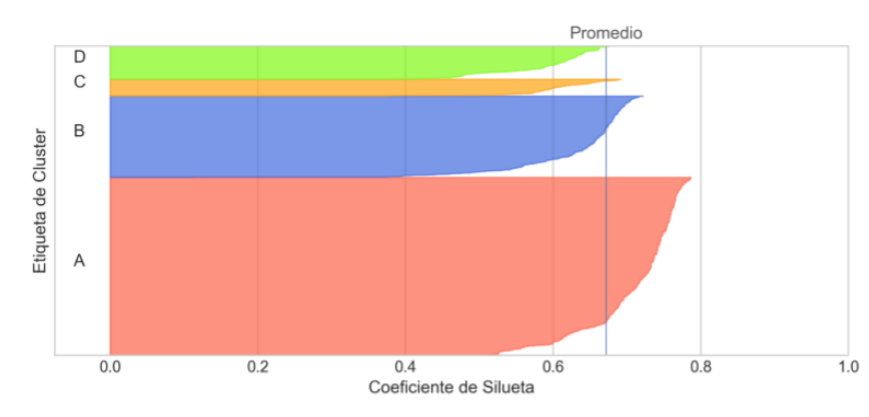


Figura 6.38: Coeficiente de silueta por *cluster*. Caso óptimo. Puerto Montt - Puerto Varas, cambios Forzados, Pre SAE.

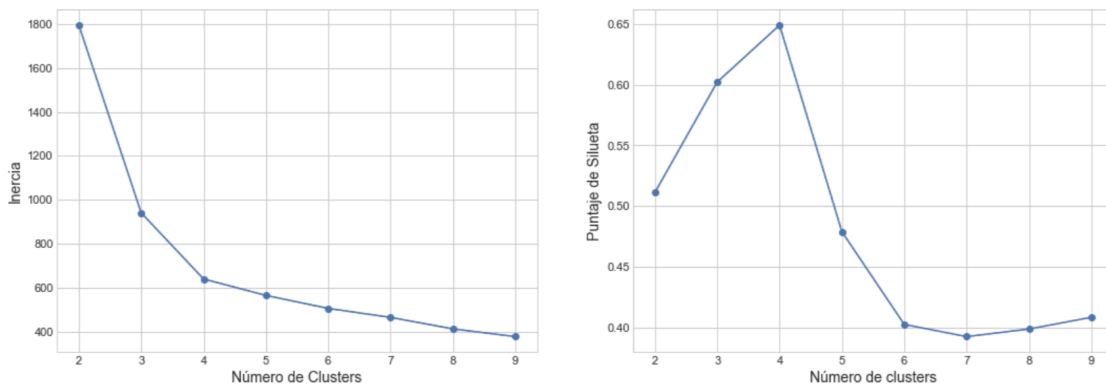


Figura 6.39: Inercia y coeficiente de silueta. Caso Puerto Montt - Puerto Varas, cambios Forzados, SAE Preferencias.

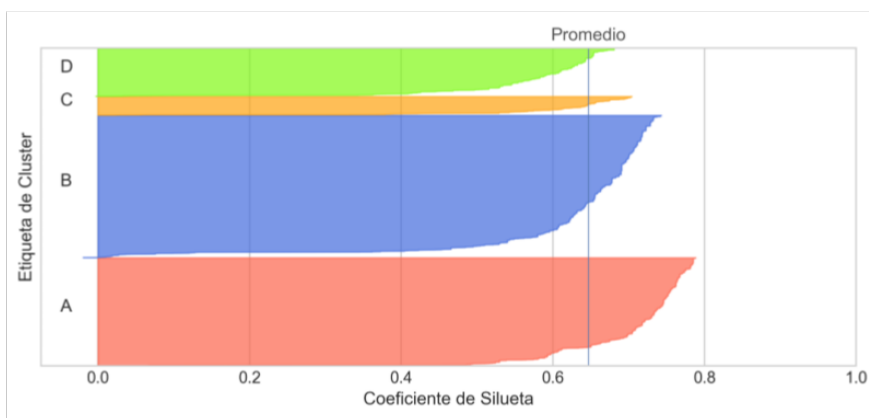


Figura 6.40: Coeficiente de silueta por *cluster*. Caso óptimo. Puerto Montt - Puerto Varas, cambios Forzados, SAE Preferencias.

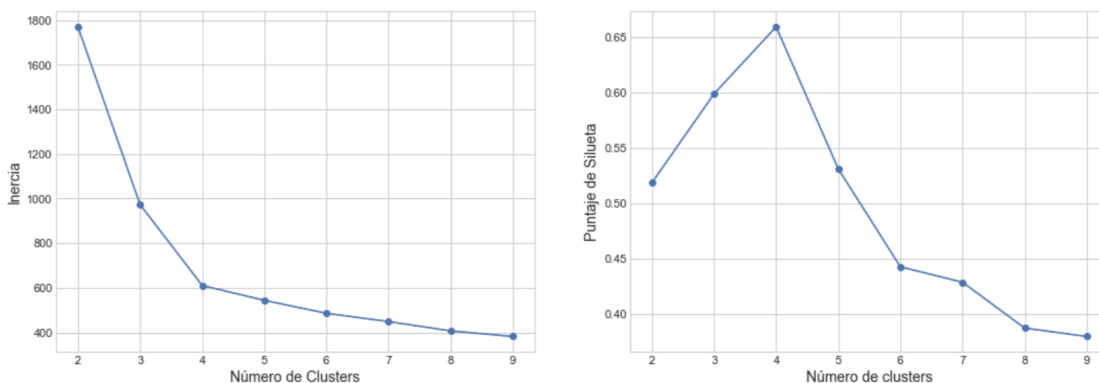


Figura 6.41: Inercia y coeficiente de silueta. Caso Puerto Montt - Puerto Varas, cambios Forzados, SAE Admisión.

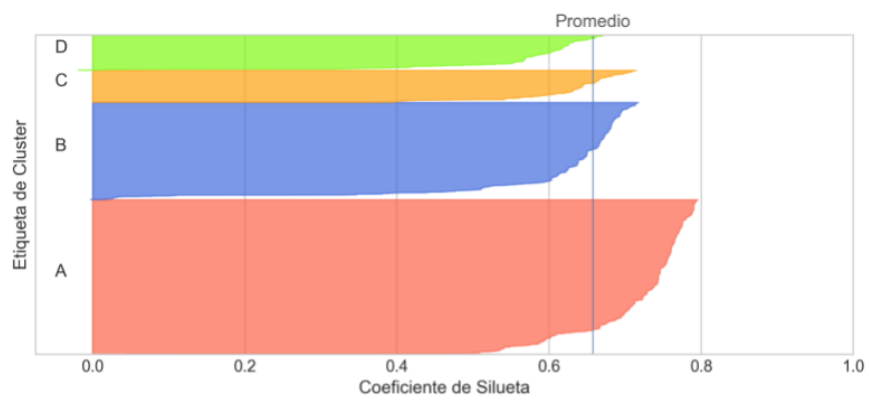


Figura 6.42: Coeficiente de silueta por *cluster*. Caso óptimo. Puerto Montt - Puerto Varas, cambios Forzados, SAE Admisión.

Cambios No Forzados

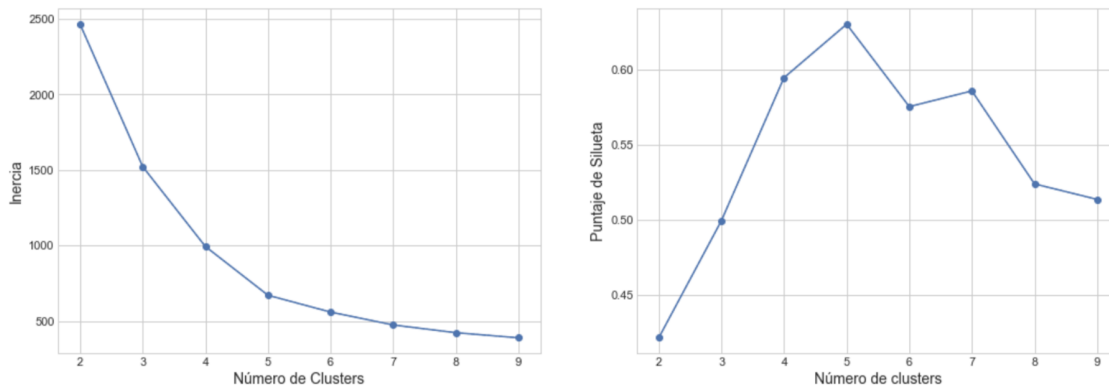


Figura 6.43: Inercia y coeficiente de silueta. Caso Puerto Montt - Puerto Varas, cambios No Forzados, Pre SAE.

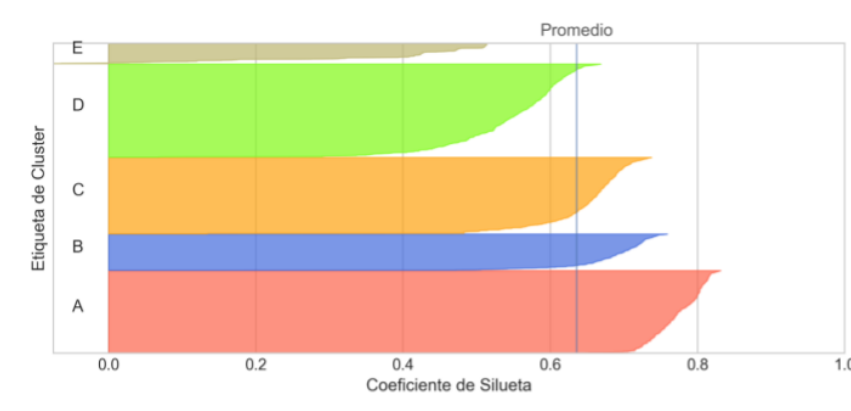


Figura 6.44: Coeficiente de silueta por *cluster*. Caso óptimo. Puerto Montt - Puerto Varas, cambios No Forzados, Pre SAE.

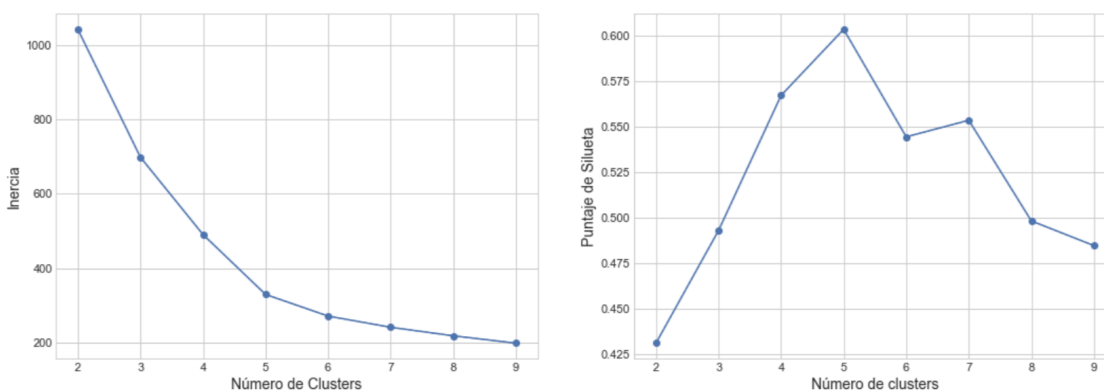


Figura 6.45: Inercia y coeficiente de silueta. Caso Puerto Montt - Puerto Varas, cambios No Forzados, SAE Preferencias.

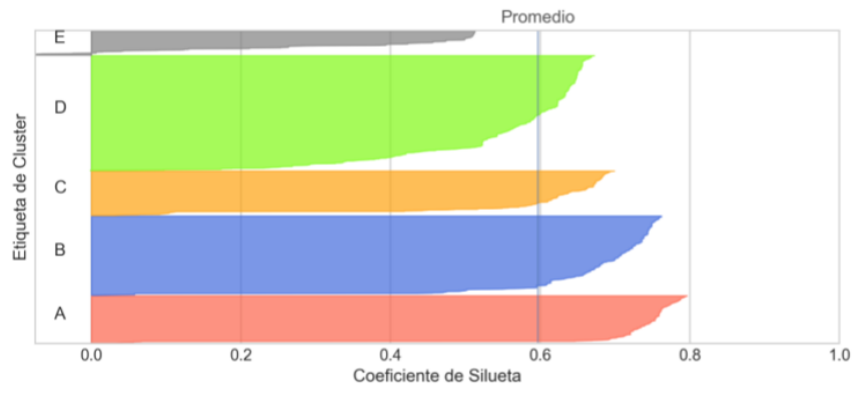


Figura 6.46: Coeficiente de silueta por *cluster*. Caso óptimo. Puerto Montt - Puerto Varas, cambios No Forzados, SAE Preferencias.

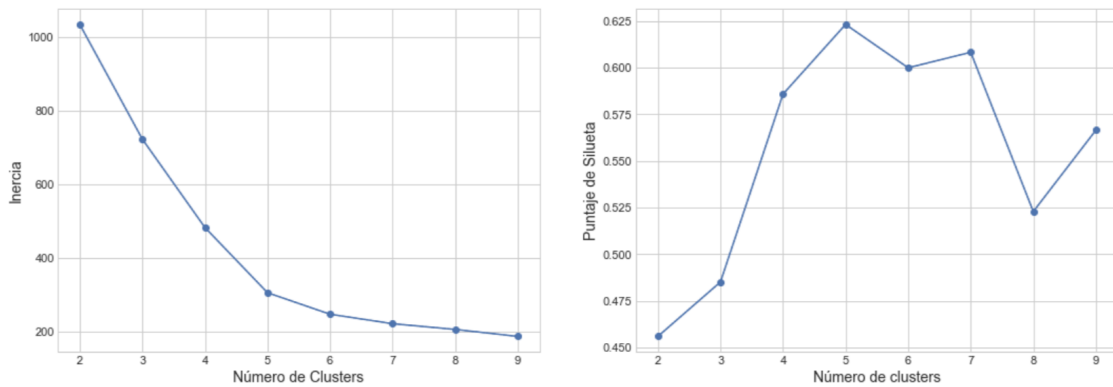


Figura 6.47: Inercia y coeficiente de silueta. Caso Puerto Montt - Puerto Varas, cambios No Forzados, SAE Admisión.

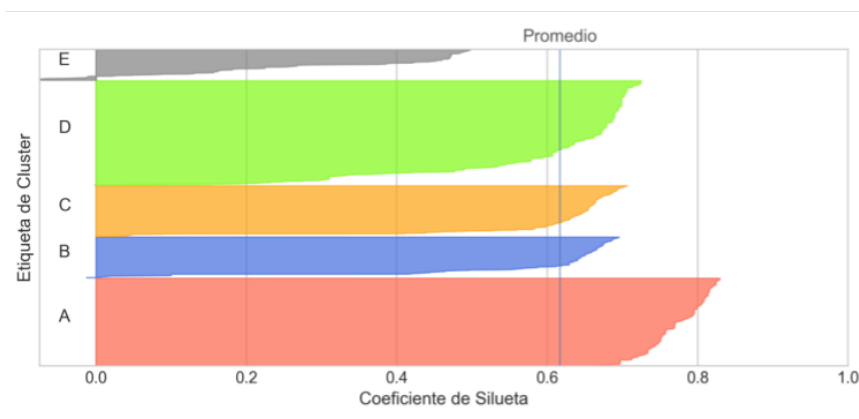


Figura 6.48: Coeficiente de silueta por *cluster*. Caso óptimo. Puerto Montt - Puerto Varas, cambios No Forzados, SAE Admisión.

D.5. Punta Arenas

Cambios Forzados

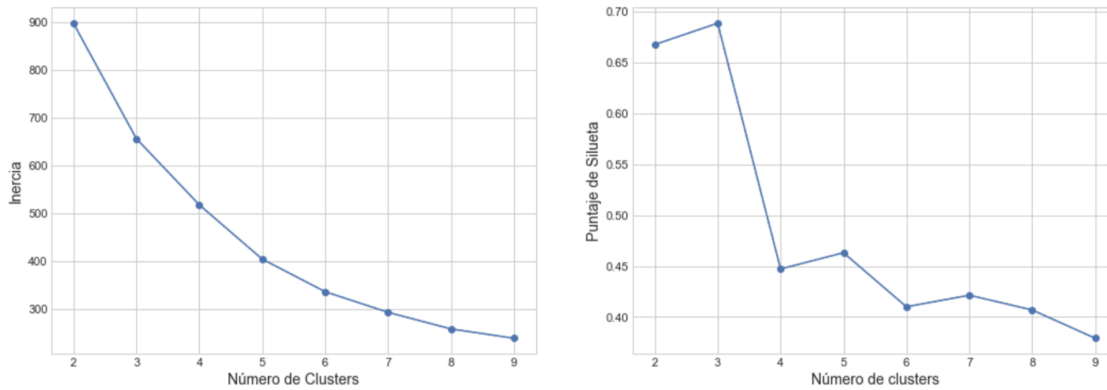


Figura 6.49: Inercia y coeficiente de silueta. Caso Punta Arenas, cambios Forzados, Pre SAE.

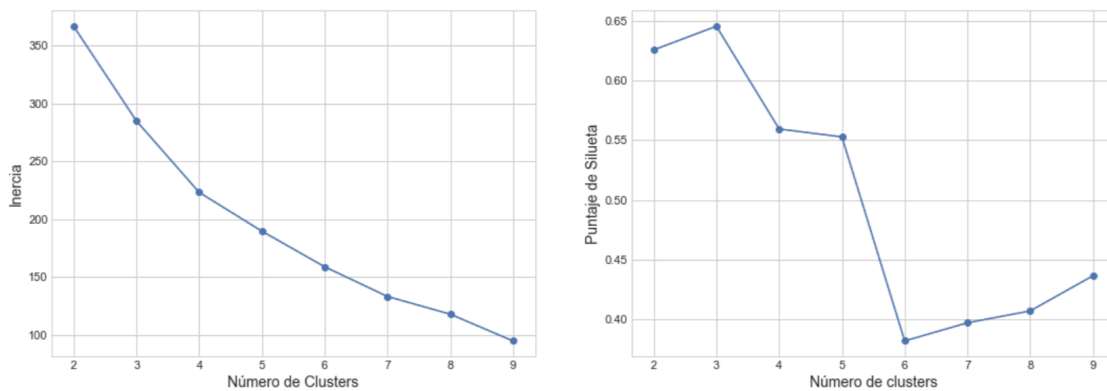


Figura 6.50: Inercia y coeficiente de silueta. Caso Punta Arenas, cambios Forzados, SAE Preferencias.

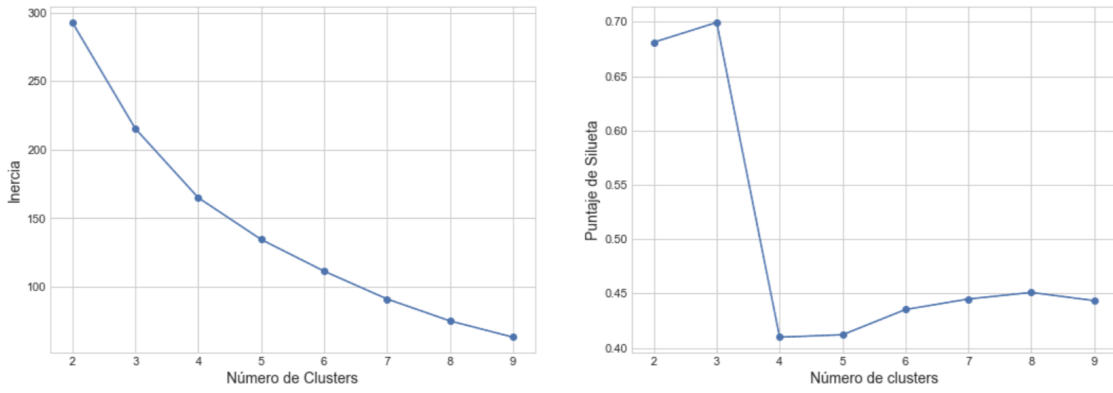


Figura 6.51: Inercia y coeficiente de silueta. Caso Punta Arenas, cambios Forzados, SAE Admisión.

Cambios No Forzados

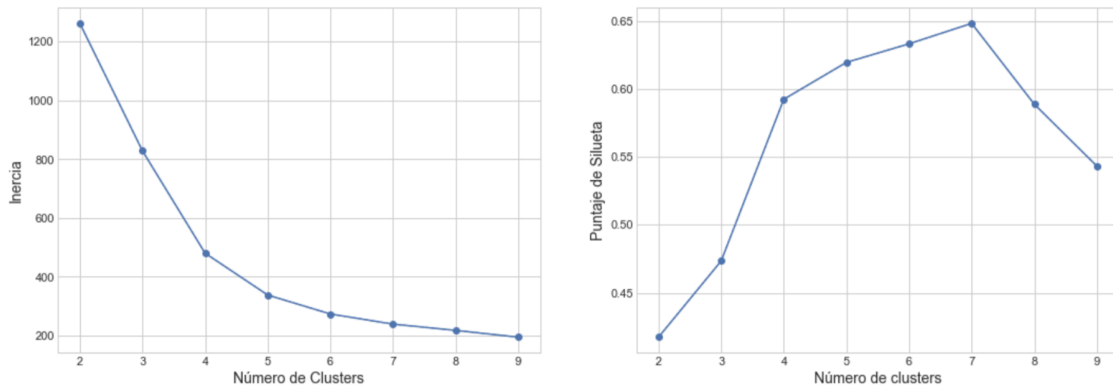


Figura 6.52: Inercia y coeficiente de silueta. Caso Punta Arenas, cambios No Forzados, Pre SAE.

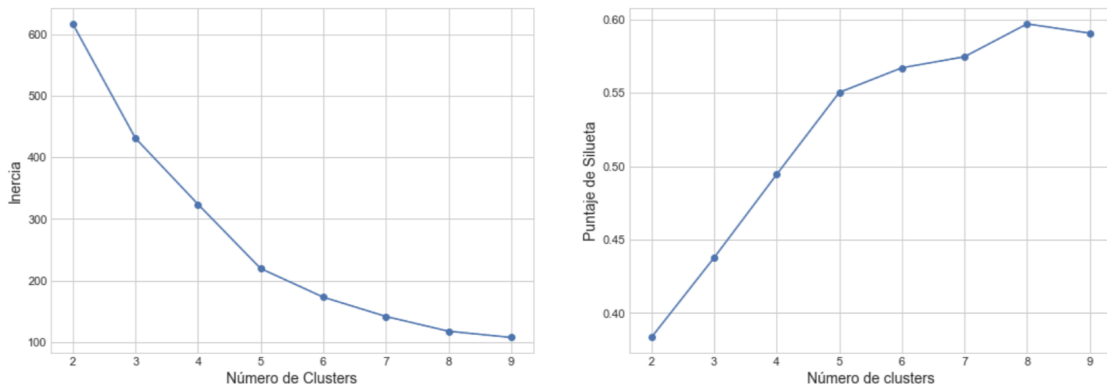


Figura 6.53: Inercia y coeficiente de silueta. Caso Punta Arenas, cambios No Forzados, SAE Preferencias.

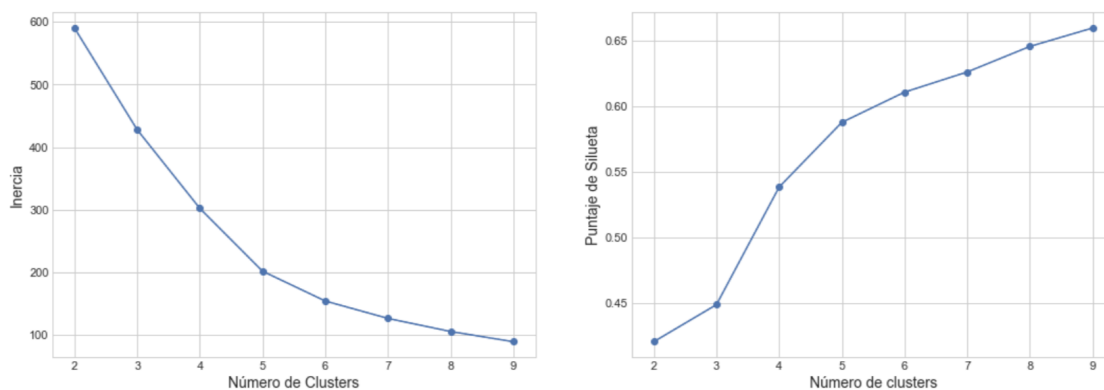


Figura 6.54: Inercia y coeficiente de silueta. Caso Punta Arenas, cambios No Forzados, SAE Admisión.

Resultados de Optimización

Conurbación	Caso	Escenario		
		Pre SAE	SAE Preferencias	SAE Admisión
Iquique - Alto Hospicio	Forzado	5	4	4
	No Forzado	4	4	4
Coquimbo - La Serena	Forzado	4	4	4
	No Forzado	4	5	5
Rancagua- Machalí	Forzado	4	4	4
	No Forzado	5	5	4
Pto. Montt - Pto. Varas	Forzado	4	4	4
	No Forzado	5	5	5

Tabla 6.4: Número óptimo de *clusters* para cada caso estudiado

Conurbación	Caso	Escenario		
		Pre SAE	SAE Preferencias	SAE Admisión
Iquique - Alto Hospicio	Forzado	Green	Green	Green
	No Forzado	Yellow	Yellow	Yellow
Coquimbo - La Serena	Forzado	Green	Green	Green
	No Forzado	Yellow	Yellow	Yellow
Rancagua- Machalí	Forzado	Green	Green	Green
	No Forzado	Yellow	Green	Green
Pto. Montt - Pto. Varas	Forzado	Green	Green	Green
	No Forzado	Green	Yellow	Yellow
Punta Arenas	Forzado	Red	Red	Red
	No Forzado	Red	Red	Red

Tabla 6.5: Calidad de resultados de *clustering* considerando Inercia y Coeficiente de Silueta.

E. Resultados de Escenarios SAE por Año

Aglomeración	Criterio	Tipo de cambio	
		Forzado	No Forzado
Iquique – Alto Hospicio	Preferencias	En 2018 se observa mayor preferencia por cambios desde EE municipal hacia EE particular subvencionado, simultaneo a una disminución de la preferencia por cambios entre particulares subvencionado.	En 2018 se observa mayor preferencia por cambios desde EE municipal hacia EE particular subvencionado, simultaneo a una disminución de la preferencia por cambios entre particulares subvencionado.
	Admisión	Mismo comportamiento ambos años.	Mismo comportamiento ambos años.
Coquimbo – La Serena	Preferencias	En 2018 aparece agrupación de cambio desde particular EE subvencionado hacia EE municipales de alto rendimiento SIMCE.	Aparece en 2018 nueva segmentación de estudiantes que se cambian entre de colegios caros, con muy poca vulnerabilidad y alto SIMCE.
	Admisión	En 2018 hay mayor admisión en cambios desde particular subvencionado a municipal. Consistente con la aparición de demanda por EE municipales de alto desempeño SIMCE.	Comportamiento similar en ambos años.
Rancagua - Machalí	Preferencias	Mismo comportamiento ambos años.	Mismo comportamiento ambos años.
	Admisión	Mismo comportamiento ambos años.	Mismo comportamiento ambos años.
Puerto Montt – Puerto Varas	Preferencias	Mismo comportamiento ambos años.	En 2018 crece agrupación de cambio desde EE subvencionados hacia EE municipales de alto rendimiento SIMCE.
	Admisión	Mismo comportamiento ambos años.	En 2018 hay mayor cambio de estudiantes desde la dependencia municipal.

Figura 6.55: Comparación de comportamiento de la movilidad escolar, entre los procesos SAE de admisión 2018 y 2019, según aglomeración, criterio y tipos de cambio estudiado.

Principales observaciones:

- En Iquique – Alto Hospicio aumenta demanda por EE subvencionados.
- En Coquimbo – La Serena aparece demanda por EE municipales con buenos puntajes SIMCE. En el caso No Forzado surge aglomeración de preferencia de cambios formada por estudiantes de GSE alto.
- En Rancagua – Machalí se observan comportamientos muy similares en ambos procesos.
- En Puerto Montt – Puerto Varas el escenario de cambio No Forzado aumenta en desorden: crece demanda por EE municipales con buenos puntajes SIMCE y aumenta el cambio desde la dependencia municipal.

IQUIQUE - ALTO HOSPICIO

2017

FORZADO

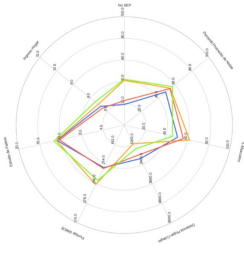
2018

PREFERENCIA

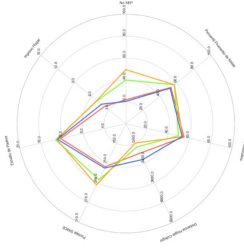
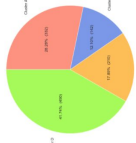
ADMISION

PREFERENCIA

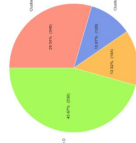
ADMISION



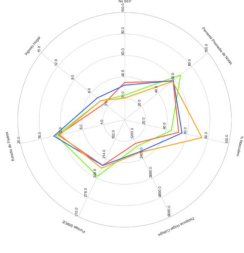
OPCION	ADMISION	PREFERENCIA	ADMISION	PREFERENCIA
A	100	100	100	100
B	100	100	100	100
C	100	100	100	100
D	100	100	100	100



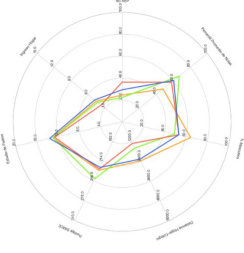
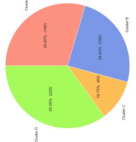
OPCION	ADMISION	PREFERENCIA	ADMISION	PREFERENCIA
A	100	100	100	100
B	100	100	100	100
C	100	100	100	100
D	100	100	100	100



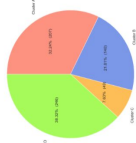
- Municipal -> Municipal
- Municipal -> Part. Subv.
- Part. Subv. -> Municipal
- Part. Subv. -> Part. Subv. 1
- Part. Subv. -> Part. Subv. 2
- Part. Subv. -> Subv./Priv.
- Subv./Priv. -> Privado
- Munic./Priv. -> Part. Subv.
- Municipal -> Muni./Subv.
- Muni./Subv. -> Municipal



OPCION	ADMISION	PREFERENCIA	ADMISION	PREFERENCIA
A	100	100	100	100
B	100	100	100	100
C	100	100	100	100
D	100	100	100	100



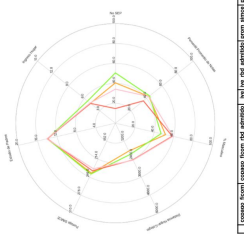
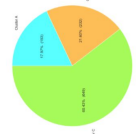
OPCION	ADMISION	PREFERENCIA	ADMISION	PREFERENCIA
A	100	100	100	100
B	100	100	100	100
C	100	100	100	100
D	100	100	100	100



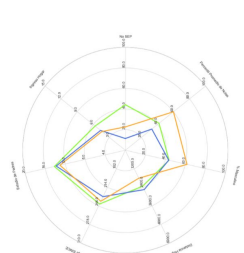
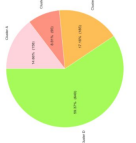
NO FORZADO



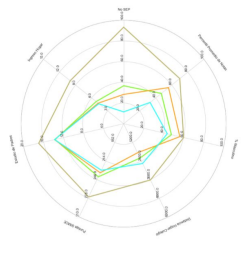
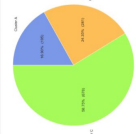
OPCION	ADMISION	PREFERENCIA	ADMISION	PREFERENCIA
A	100	100	100	100
B	100	100	100	100
C	100	100	100	100
D	100	100	100	100



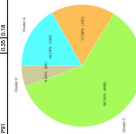
OPCION	ADMISION	PREFERENCIA	ADMISION	PREFERENCIA
A	100	100	100	100
B	100	100	100	100
C	100	100	100	100
D	100	100	100	100



OPCION	ADMISION	PREFERENCIA	ADMISION	PREFERENCIA
A	100	100	100	100
B	100	100	100	100
C	100	100	100	100
D	100	100	100	100



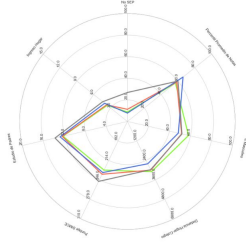
OPCION	ADMISION	PREFERENCIA	ADMISION	PREFERENCIA
A	100	100	100	100
B	100	100	100	100
C	100	100	100	100
D	100	100	100	100



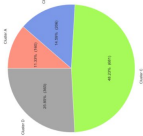
COQUIMBO - LA SERENA

2017

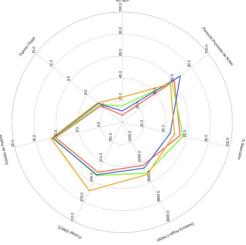
PREFERENCIA



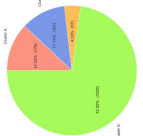
Indicador	Objetivo	Realizado	Porcentaje
A) Cobertura	100%	95%	95%
B) Satisfacción	90%	85%	94%
C) Retención	90%	85%	94%
D) Eficiencia	90%	85%	94%



ADMISIÓN



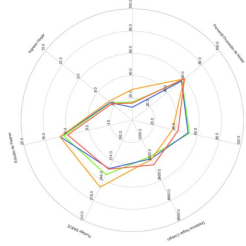
Indicador	Objetivo	Realizado	Porcentaje
A) Cobertura	100%	95%	95%
B) Satisfacción	90%	85%	94%
C) Retención	90%	85%	94%
D) Eficiencia	90%	85%	94%



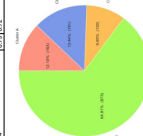
- Municipal -> Municipal
- Municipal -> Part. Subv.
- Part. Subv. -> Municipal
- Part. Subv. -> Part. Subv. 1
- Part. Subv. -> Part. Subv. 2
- Subv./Priv. -> Subv./Priv.
- Subv./Priv. -> Privado
- Munic./Priv. -> Part. Subv.
- Municipal-> Muni./Subv.
- Muni./Subv.-> Municipal

2018

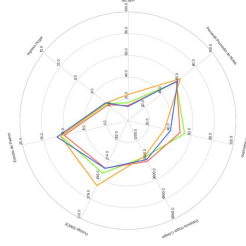
PREFERENCIA



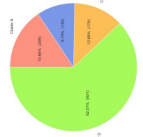
Indicador	Objetivo	Realizado	Porcentaje
A) Cobertura	100%	95%	95%
B) Satisfacción	90%	85%	94%
C) Retención	90%	85%	94%
D) Eficiencia	90%	85%	94%



ADMISIÓN



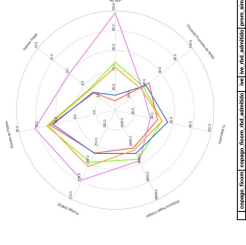
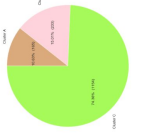
Indicador	Objetivo	Realizado	Porcentaje
A) Cobertura	100%	95%	95%
B) Satisfacción	90%	85%	94%
C) Retención	90%	85%	94%
D) Eficiencia	90%	85%	94%



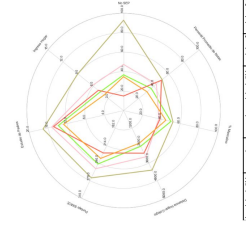
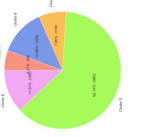
NO FORZADO



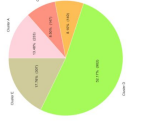
Indicador	Objetivo	Realizado	Porcentaje
A) Cobertura	100%	95%	95%
B) Satisfacción	90%	85%	94%
C) Retención	90%	85%	94%
D) Eficiencia	90%	85%	94%



Indicador	Objetivo	Realizado	Porcentaje
A) Cobertura	100%	95%	95%
B) Satisfacción	90%	85%	94%
C) Retención	90%	85%	94%
D) Eficiencia	90%	85%	94%



Indicador	Objetivo	Realizado	Porcentaje
A) Cobertura	100%	95%	95%
B) Satisfacción	90%	85%	94%
C) Retención	90%	85%	94%
D) Eficiencia	90%	85%	94%



RANCAGUA - MACHALI

2017

FORZADO

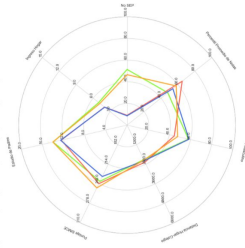
2018

PREFERENCIA

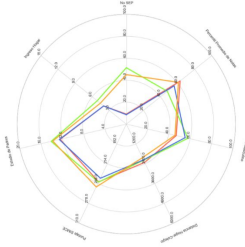
ADMISION

PREFERENCIA

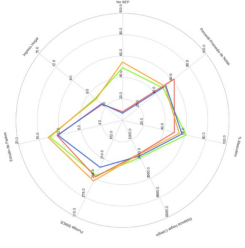
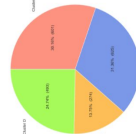
ADMISION



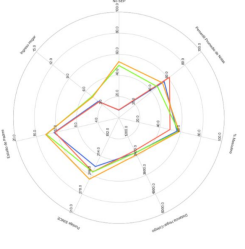
OPCION	OPCION	OPCION	OPCION	OPCION	OPCION	OPCION	OPCION	OPCION	OPCION
A	100	100	100	100	100	100	100	100	100
B	100	100	100	100	100	100	100	100	100
C	100	100	100	100	100	100	100	100	100
D	100	100	100	100	100	100	100	100	100



OPCION	OPCION	OPCION	OPCION	OPCION	OPCION	OPCION	OPCION	OPCION	OPCION
A	100	100	100	100	100	100	100	100	100
B	100	100	100	100	100	100	100	100	100
C	100	100	100	100	100	100	100	100	100
D	100	100	100	100	100	100	100	100	100



OPCION	OPCION	OPCION	OPCION	OPCION	OPCION	OPCION	OPCION	OPCION	OPCION
A	100	100	100	100	100	100	100	100	100
B	100	100	100	100	100	100	100	100	100
C	100	100	100	100	100	100	100	100	100
D	100	100	100	100	100	100	100	100	100

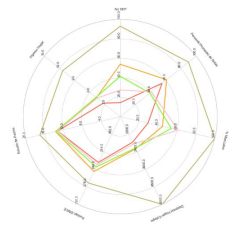


OPCION	OPCION	OPCION	OPCION	OPCION	OPCION	OPCION	OPCION	OPCION	OPCION
A	100	100	100	100	100	100	100	100	100
B	100	100	100	100	100	100	100	100	100
C	100	100	100	100	100	100	100	100	100
D	100	100	100	100	100	100	100	100	100

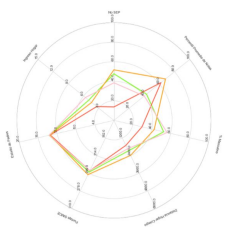
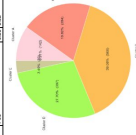


- Municipal -> Municipal
- Municipal -> Part. Subv.
- Part. Subv. -> Municipal
- Part. Subv. -> Part. Subv. 1
- Part. Subv. -> Part. Subv. 2
- Part. Subv. -> Subv./Ptv.
- Subv./Ptv. -> Privado
- Munic./Ptv. -> Part. Subv.
- Municipal -> Muni./Subv.
- Muni./Subv. -> Municipal
- Privado -> Subv./Ptv.

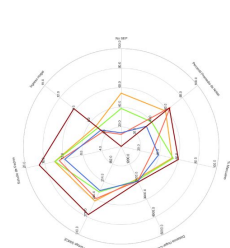
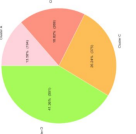
NO FORZADO



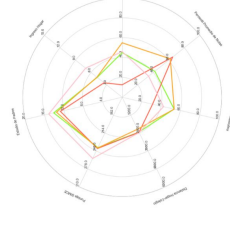
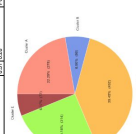
OPCION	OPCION	OPCION	OPCION	OPCION	OPCION	OPCION	OPCION	OPCION	OPCION
A	100	100	100	100	100	100	100	100	100
B	100	100	100	100	100	100	100	100	100
C	100	100	100	100	100	100	100	100	100
D	100	100	100	100	100	100	100	100	100



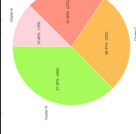
OPCION	OPCION	OPCION	OPCION	OPCION	OPCION	OPCION	OPCION	OPCION	OPCION
A	100	100	100	100	100	100	100	100	100
B	100	100	100	100	100	100	100	100	100
C	100	100	100	100	100	100	100	100	100
D	100	100	100	100	100	100	100	100	100



OPCION	OPCION	OPCION	OPCION	OPCION	OPCION	OPCION	OPCION	OPCION	OPCION
A	100	100	100	100	100	100	100	100	100
B	100	100	100	100	100	100	100	100	100
C	100	100	100	100	100	100	100	100	100
D	100	100	100	100	100	100	100	100	100



OPCION	OPCION	OPCION	OPCION	OPCION	OPCION	OPCION	OPCION	OPCION	OPCION
A	100	100	100	100	100	100	100	100	100
B	100	100	100	100	100	100	100	100	100
C	100	100	100	100	100	100	100	100	100
D	100	100	100	100	100	100	100	100	100



PUERTO MONTT - PUERTO VARAS

2017

2018

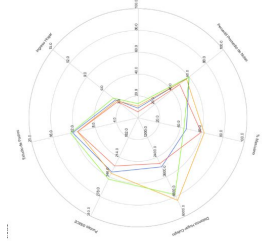
PREFERENCIA

FORZADO

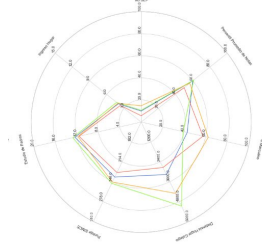
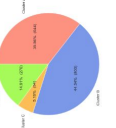
ADMISIÓN

PREFERENCIA

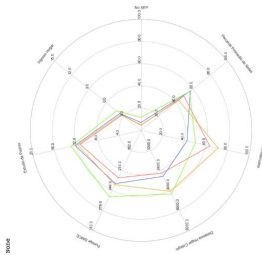
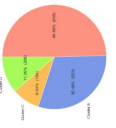
ADMISIÓN



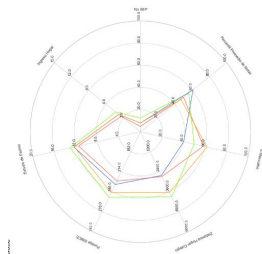
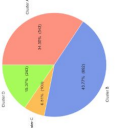
CATEGORÍA	ADMISIÓN	NO ADMISIÓN	TOTAL
A	100	0	100
B	100	0	100
C	100	0	100
D	100	0	100



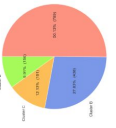
CATEGORÍA	ADMISIÓN	NO ADMISIÓN	TOTAL
A	100	0	100
B	100	0	100
C	100	0	100
D	100	0	100



CATEGORÍA	ADMISIÓN	NO ADMISIÓN	TOTAL
A	100	0	100
B	100	0	100
C	100	0	100
D	100	0	100

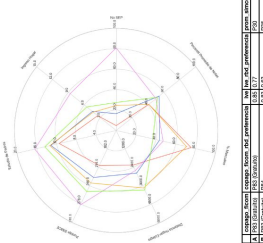


CATEGORÍA	ADMISIÓN	NO ADMISIÓN	TOTAL
A	100	0	100
B	100	0	100
C	100	0	100
D	100	0	100

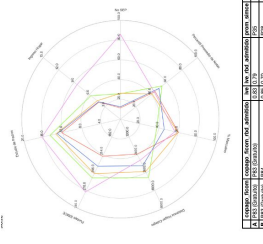


- Municipal -> Municipal
- Municipal -> Part. Subv.
- Part. Subv. -> Municipal
- Part. Subv. -> Part. Subv. 1
- Part. Subv. -> Part. Subv. 2
- Part. Subv. -> Subv./Priv.
- Subv./Priv. -> Privado
- Munic./Priv. -> Part. Subv.
- Municipal -> Muni./Subv.
- Muni./Subv. -> Municipal

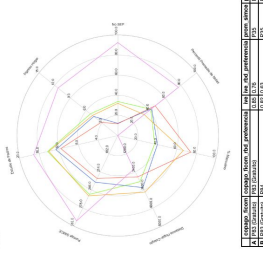
NO FORZADO



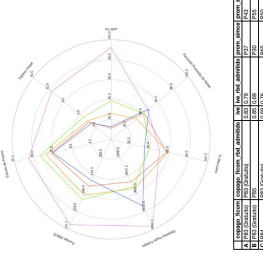
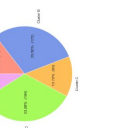
CATEGORÍA	ADMISIÓN	NO ADMISIÓN	TOTAL
A	100	0	100
B	100	0	100
C	100	0	100
D	100	0	100



CATEGORÍA	ADMISIÓN	NO ADMISIÓN	TOTAL
A	100	0	100
B	100	0	100
C	100	0	100
D	100	0	100



CATEGORÍA	ADMISIÓN	NO ADMISIÓN	TOTAL
A	100	0	100
B	100	0	100
C	100	0	100
D	100	0	100



CATEGORÍA	ADMISIÓN	NO ADMISIÓN	TOTAL
A	100	0	100
B	100	0	100
C	100	0	100
D	100	0	100



F. Resultados por Aglomeración Urbana y Caso

IQUIQUE - ALTO HOSPICIO

FORZADO

CONURBACIÓN

CAMBIO

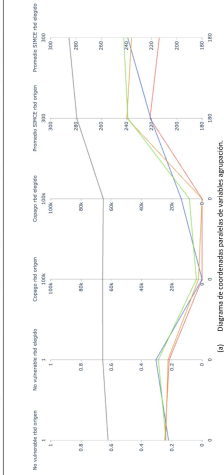
ESCENARIO

PRE-SAE

SAE-PREFERENCIAS

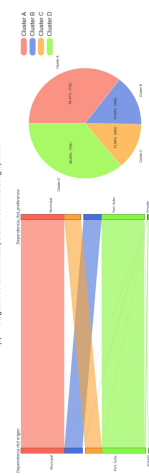
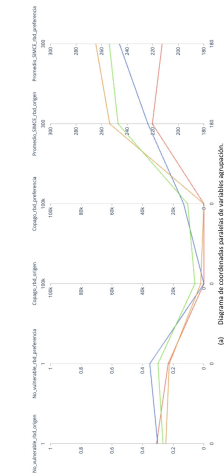
SAE-ADMISIÓN

Cluster	Compro el origen	Compro el destino	NS	Procedo del origen	Procedo del destino	Procedo del origen	Procedo del destino
A	NS (Gremio)	NS (Gremio)	0.21	NS (Gremio)	NS (Gremio)	NS (Gremio)	NS (Gremio)
B	NS (Gremio)	NS (Gremio)	0.25	NS (Gremio)	NS (Gremio)	NS (Gremio)	NS (Gremio)
C	NS (Gremio)	NS (Gremio)	0.27	NS (Gremio)	NS (Gremio)	NS (Gremio)	NS (Gremio)
D	NS (Gremio)	NS (Gremio)	0.28	NS (Gremio)	NS (Gremio)	NS (Gremio)	NS (Gremio)
E	NS (Gremio)	NS (Gremio)	0.31	NS (Gremio)	NS (Gremio)	NS (Gremio)	NS (Gremio)



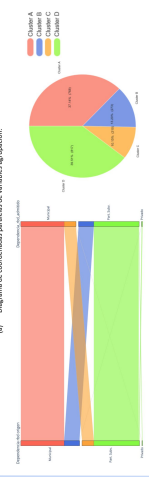
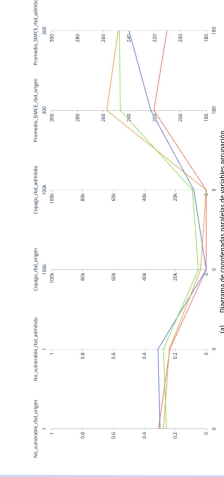
(b) Flujo de cambios entre dependencias administrativas. (c) Gráfico circular del tamaño de cada cluster.

Cluster	Compro el origen	Compro el destino	NS	Procedo del origen	Procedo del destino	Procedo del origen	Procedo del destino
A	NS (Gremio)	NS (Gremio)	0.03	NS (Gremio)	NS (Gremio)	NS (Gremio)	NS (Gremio)
B	NS (Gremio)	NS (Gremio)	0.07	NS (Gremio)	NS (Gremio)	NS (Gremio)	NS (Gremio)
C	NS (Gremio)	NS (Gremio)	0.15	NS (Gremio)	NS (Gremio)	NS (Gremio)	NS (Gremio)
D	NS (Gremio)	NS (Gremio)	0.18	NS (Gremio)	NS (Gremio)	NS (Gremio)	NS (Gremio)



(b) Flujo de cambios entre dependencias administrativas. (c) Gráfico circular del tamaño de cada cluster.

Cluster	Compro el origen	Compro el destino	NS	Procedo del origen	Procedo del destino	Procedo del origen	Procedo del destino
A	NS (Gremio)	NS (Gremio)	0.03	NS (Gremio)	NS (Gremio)	NS (Gremio)	NS (Gremio)
B	NS (Gremio)	NS (Gremio)	0.10	NS (Gremio)	NS (Gremio)	NS (Gremio)	NS (Gremio)
C	NS (Gremio)	NS (Gremio)	0.10	NS (Gremio)	NS (Gremio)	NS (Gremio)	NS (Gremio)
D	NS (Gremio)	NS (Gremio)	0.12	NS (Gremio)	NS (Gremio)	NS (Gremio)	NS (Gremio)
E	NS (Gremio)	NS (Gremio)	0.14	NS (Gremio)	NS (Gremio)	NS (Gremio)	NS (Gremio)



(b) Flujo de cambios entre dependencias administrativas. (c) Gráfico circular del tamaño de cada cluster.

VARIABLES COLEGIO

TAMAÑO CLUSTERS

VARIABLES ESTUDIANTES

CALIDAD DE CLUSTERS

IQUIQUE - ALTO HOSPICIO

NO FORZADO

CONURBACIÓN

CAMBIO

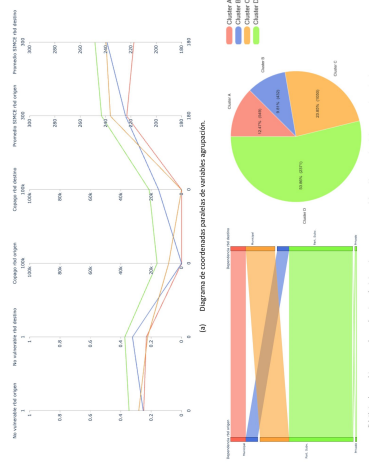
ESCENARIO

PRE-SAE

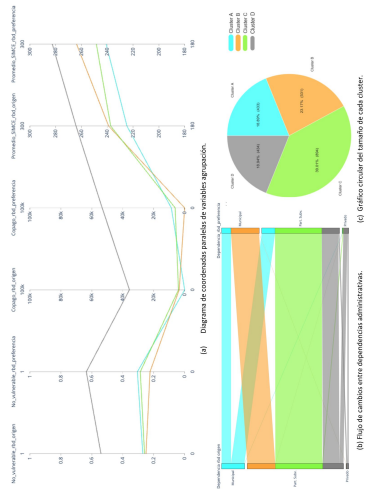
SAE-PREFERENCIAS

SAE-ADMISIÓN

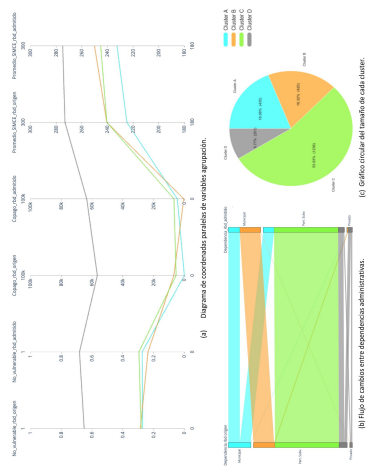
Cluster	Comp. del origen	Comp. del destino	Cl. del origen	Cl. del destino	Prevalencia SAE	Prevalencia SAE	Prevalencia SAE	Prevalencia SAE
	del origen	del destino	del origen	del destino	del origen	del destino	del origen	del destino
A	PS (Comun.)	PS (Comun.)	PS	PS	0.15	0.17	PS	PS
B	PS (Comun.)	PS (Comun.)	PS	PS	0.22	0.17	PS	PS
C	PS (Comun.)	PS (Comun.)	PS	PS	0.22	0.17	PS	PS
D	PS (Comun.)	PS (Comun.)	PS	PS	0.22	0.17	PS	PS



Cluster	Comp. del origen	Comp. del destino	Cl. del origen	Cl. del destino	Prevalencia SAE	Prevalencia SAE	Prevalencia SAE	Prevalencia SAE
	del origen	del destino	del origen	del destino	del origen	del destino	del origen	del destino
A	PS (Comun.)	PS (Comun.)	PS	PS	0.10	0.10	PS	PS
B	PS (Comun.)	PS (Comun.)	PS	PS	0.10	0.10	PS	PS
C	PS (Comun.)	PS (Comun.)	PS	PS	0.10	0.10	PS	PS
D	PS (Comun.)	PS (Comun.)	PS	PS	0.10	0.10	PS	PS



Cluster	Comp. del origen	Comp. del destino	Cl. del origen	Cl. del destino	Prevalencia SAE	Prevalencia SAE	Prevalencia SAE	Prevalencia SAE
	del origen	del destino	del origen	del destino	del origen	del destino	del origen	del destino
A	PS (Comun.)	PS (Comun.)	PS	PS	0.13	0.13	PS	PS
B	PS (Comun.)	PS (Comun.)	PS	PS	0.13	0.13	PS	PS
C	PS (Comun.)	PS (Comun.)	PS	PS	0.13	0.13	PS	PS
D	PS (Comun.)	PS (Comun.)	PS	PS	0.13	0.13	PS	PS



VARIABLES COLEGIO

TAMAÑO CLUSTERS

VARIABLES ESTUDIANTES

CALIDAD DE CLUSTERS

COQUIMBO - LA SERENA

FORZADO

CONURBACIÓN

CAMBIO

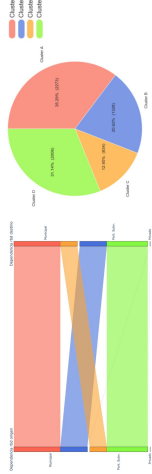
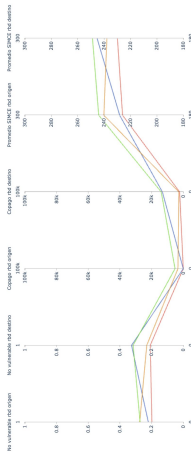
ESCENARIO

PRE-SAE

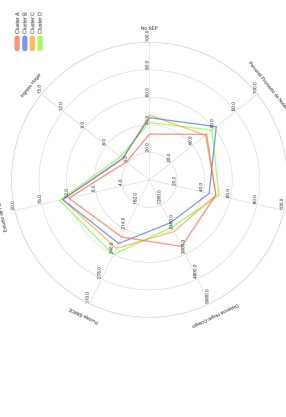
SAE-PREFERENCIAS

SAE-ADMISIÓN

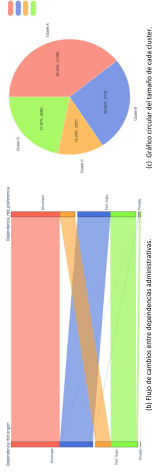
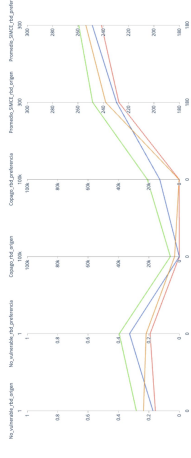
Cluster	Compro. Adm. (Adm.)	Compro. Econ. (Econ.)	Compro. Edu. (Edu.)	Compro. Soc. (Soc.)	Compro. Amb. (Amb.)	Compro. Pol. (Pol.)	Compro. Total (Total)	Dependencia (Dep.)	Dependencia (Dep.)	Dependencia (Dep.)	Dependencia (Dep.)	Dependencia (Dep.)
	Adm.	Econ.	Edu.	Soc.	Amb.	Pol.	Total	Adm.	Econ.	Edu.	Soc.	Amb.
A	0,22	0,73	0,06	0,17	0,17	0,27	0,52	P21	P22	P23	P24	P25
B	0,27	0,73	0,06	0,17	0,17	0,27	0,52	P26	P27	P28	P29	P30
C	0,27	0,73	0,06	0,17	0,17	0,27	0,52	P31	P32	P33	P34	P35
D	0,27	0,73	0,06	0,17	0,17	0,27	0,52	P36	P37	P38	P39	P40



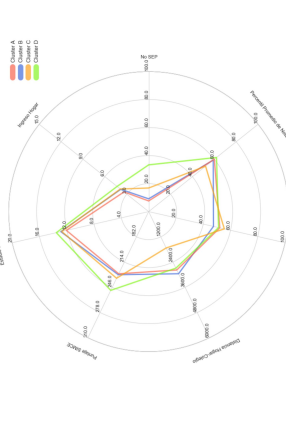
(c) Gráfico circular del tamaño de cada cluster.



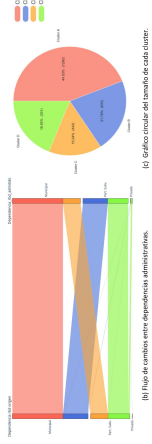
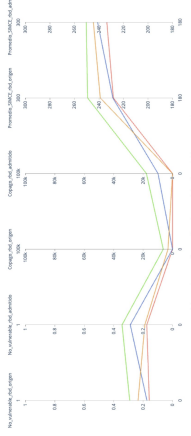
Cluster	Compro. Adm. (Adm.)	Compro. Econ. (Econ.)	Compro. Edu. (Edu.)	Compro. Soc. (Soc.)	Compro. Amb. (Amb.)	Compro. Pol. (Pol.)	Compro. Total (Total)	Dependencia (Dep.)	Dependencia (Dep.)	Dependencia (Dep.)	Dependencia (Dep.)	Dependencia (Dep.)
	Adm.	Econ.	Edu.	Soc.	Amb.	Pol.	Total	Adm.	Econ.	Edu.	Soc.	Amb.
A	0,27	0,73	0,06	0,17	0,17	0,27	0,52	P41	P42	P43	P44	P45
B	0,27	0,73	0,06	0,17	0,17	0,27	0,52	P46	P47	P48	P49	P50
C	0,27	0,73	0,06	0,17	0,17	0,27	0,52	P51	P52	P53	P54	P55
D	0,27	0,73	0,06	0,17	0,17	0,27	0,52	P56	P57	P58	P59	P60



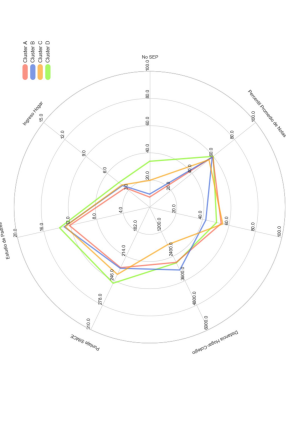
(c) Gráfico circular del tamaño de cada cluster.



Cluster	Compro. Adm. (Adm.)	Compro. Econ. (Econ.)	Compro. Edu. (Edu.)	Compro. Soc. (Soc.)	Compro. Amb. (Amb.)	Compro. Pol. (Pol.)	Compro. Total (Total)	Dependencia (Dep.)	Dependencia (Dep.)	Dependencia (Dep.)	Dependencia (Dep.)	Dependencia (Dep.)
	Adm.	Econ.	Edu.	Soc.	Amb.	Pol.	Total	Adm.	Econ.	Edu.	Soc.	Amb.
A	0,27	0,73	0,06	0,17	0,17	0,27	0,52	P61	P62	P63	P64	P65
B	0,27	0,73	0,06	0,17	0,17	0,27	0,52	P66	P67	P68	P69	P70
C	0,27	0,73	0,06	0,17	0,17	0,27	0,52	P71	P72	P73	P74	P75
D	0,27	0,73	0,06	0,17	0,17	0,27	0,52	P76	P77	P78	P79	P80



(c) Gráfico circular del tamaño de cada cluster.

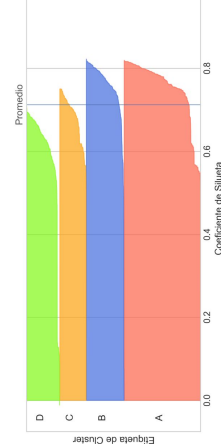
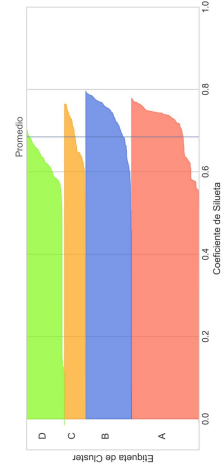
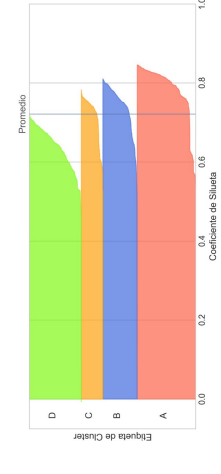


VARIABLES COLEGIO

TAMAÑO CLUSTERS

VARIABLES ESTUDIANTES

CALIDAD DE CLUSTERS



COQUIMBO - LA SERENA

NO FORZADO

CONURBACIÓN

CAMBIO

ESENARIO

CENTROS
CLUSTERS

VARIABLES
COLEGIO

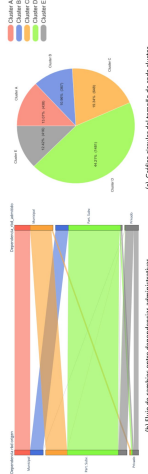
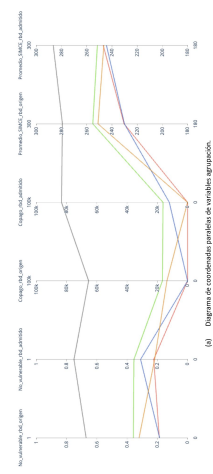
TAMAÑO
CLUSTERS

VARIABLES
ESTUDIANTES

CALIDAD
DE
CLUSTERS

PRE-SAE

Cluster	Campaño		INE	IVE		PROMEDIO SAE		PROMEDIO SAE		DEPENDENCIA	
	nd. origen	nd. destino		nd. origen	nd. destino	nd. origen	nd. destino	nd. origen	nd. destino	nd. origen	nd. destino
A	P21	P22	0.17	0.07	P27	P28	P23	P24	Municipal	Part. Subr.	Part. Subr.
B	P31	P32	0.10	0.31	P37	P38	P33	P34	Municipal	Part. Subr.	Part. Subr.
C	P41	P42	0.31	0.19	P47	P48	P43	P44	Municipal	Part. Subr.	Part. Subr.
D	P51	P52	0.10	0.10	P57	P58	P53	P54	Municipal	Part. Subr.	Part. Subr.
E	P61	P62	0.31	0.31	P67	P68	P63	P64	Municipal	Part. Subr.	Part. Subr.



(a) Diagrama de coordenadas paralelas de variables agregadas.

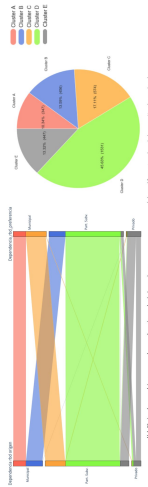
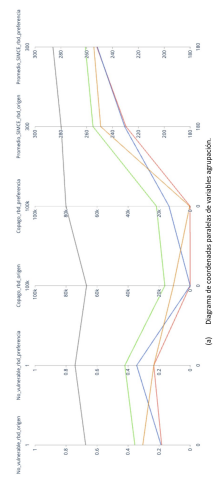
(b) Flujo de cambios entre dependencias administrativas.



(c) Grafico circular del tamaño de cada cluster.

SAE-PREFERENCIAS

Cluster	Campaño		INE	IVE		PROMEDIO SAE		PROMEDIO SAE		DEPENDENCIA	
	nd. origen	nd. destino		nd. origen	nd. destino	nd. origen	nd. destino	nd. origen	nd. destino	nd. origen	nd. destino
A	P21	P22	0.17	0.07	P27	P28	P23	P24	Municipal	Part. Subr.	Part. Subr.
B	P31	P32	0.10	0.31	P37	P38	P33	P34	Municipal	Part. Subr.	Part. Subr.
C	P41	P42	0.31	0.19	P47	P48	P43	P44	Municipal	Part. Subr.	Part. Subr.
D	P51	P52	0.10	0.10	P57	P58	P53	P54	Municipal	Part. Subr.	Part. Subr.
E	P61	P62	0.31	0.31	P67	P68	P63	P64	Municipal	Part. Subr.	Part. Subr.



(a) Diagrama de coordenadas paralelas de variables agregadas.

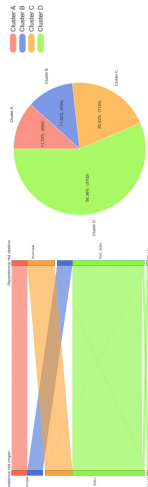
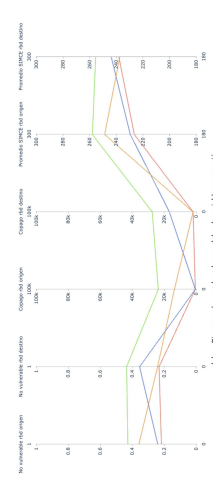
(b) Flujo de cambios entre dependencias administrativas.



(c) Grafico circular del tamaño de cada cluster.

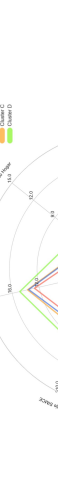
SAE-ADMISIÓN

Cluster	Campaño		INE	IVE		PROMEDIO SAE		PROMEDIO SAE		DEPENDENCIA	
	nd. origen	nd. destino		nd. origen	nd. destino	nd. origen	nd. destino	nd. origen	nd. destino	nd. origen	nd. destino
A	P21	P22	0.17	0.07	P27	P28	P23	P24	Municipal	Part. Subr.	Part. Subr.
B	P31	P32	0.10	0.31	P37	P38	P33	P34	Municipal	Part. Subr.	Part. Subr.
C	P41	P42	0.31	0.19	P47	P48	P43	P44	Municipal	Part. Subr.	Part. Subr.
D	P51	P52	0.10	0.10	P57	P58	P53	P54	Municipal	Part. Subr.	Part. Subr.
E	P61	P62	0.31	0.31	P67	P68	P63	P64	Municipal	Part. Subr.	Part. Subr.

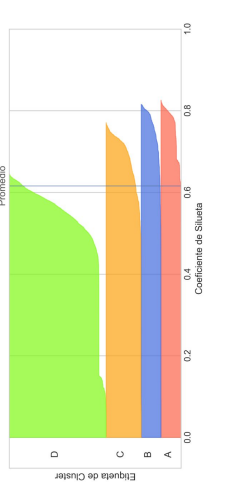
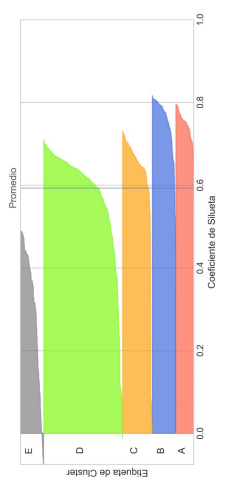
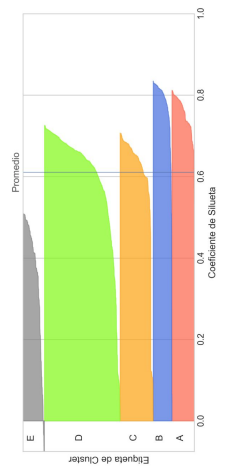


(a) Diagrama de coordenadas paralelas de variables agregadas.

(b) Flujo de cambios entre dependencias administrativas.



(c) Grafico circular del tamaño de cada cluster.



RANCAGUA - MACHALI

FORZADO

CONURBACIÓN

CAMBIO

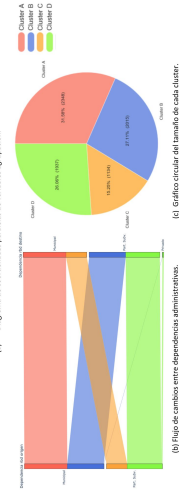
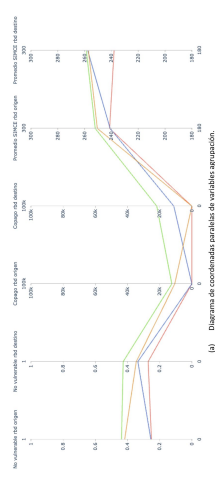
ESCENARIO

PRE-SAE

SAE-PREFERENCIAS

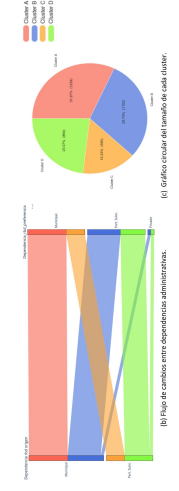
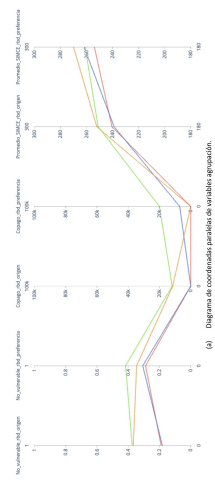
SAE-ADMISIÓN

Cluster	Compro	IVE	Promedio SAE	Dependencia	Dependencia
	rd. origen	rd. destino	rd. origen	rd. origen	rd. destino
A	P01 (Gomara)	0,25	0,23	P01	P01
B	P02 (Gomara)	0,25	0,23	P02	P02
C	P03 (Gomara)	0,13	0,06	P03	P03
D	P04 (Gomara)	0,14	0,14	P04	P04



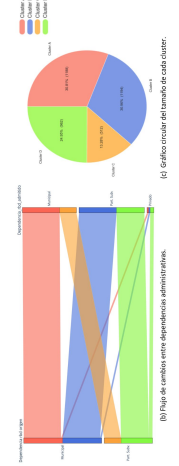
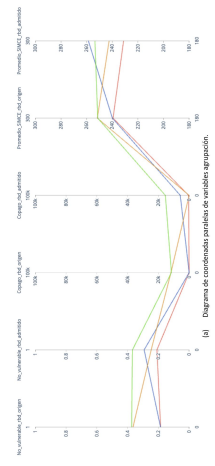
(a) Diagrama de coordenadas paralelas de variables agrupadas. (b) Flujo de cambios entre dependencias administrativas. (c) Gráfico circular del tamaño de cada cluster.

Cluster	Compro	IVE	Promedio SAE	Dependencia	Dependencia
	rd. origen	rd. preferencia	rd. origen	rd. origen	rd. preferencia
A	P01 (Gomara)	0,21	0,21	P01	P01
B	P02 (Gomara)	0,13	0,13	P02	P02
C	P03 (Gomara)	0,05	0,05	P03	P03
D	P04 (Gomara)	0,11	0,11	P04	P04



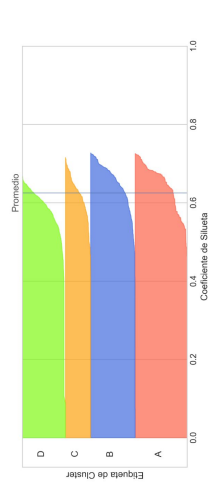
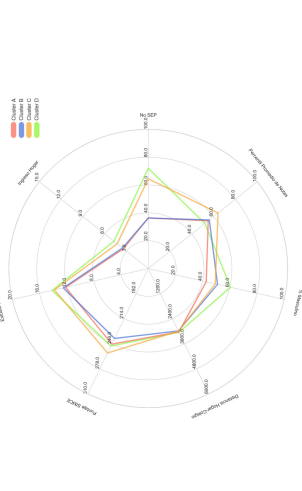
(a) Diagrama de coordenadas paralelas de variables agrupadas. (b) Flujo de cambios entre dependencias administrativas. (c) Gráfico circular del tamaño de cada cluster.

Cluster	Compro	IVE	Promedio SAE	Dependencia	Dependencia
	rd. origen	rd. admisión	rd. admisión	rd. origen	rd. admisión
A	P01 (Gomara)	0,14	0,14	P01	P01
B	P02 (Gomara)	0,13	0,13	P02	P02
C	P03 (Gomara)	0,03	0,03	P03	P03
D	P04 (Gomara)	0,04	0,04	P04	P04

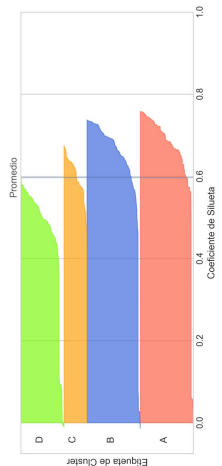
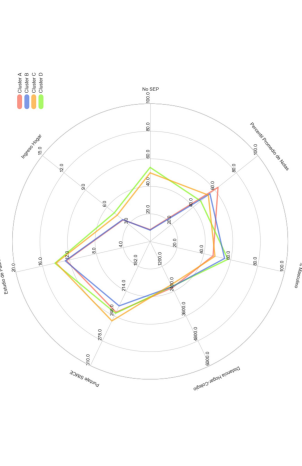
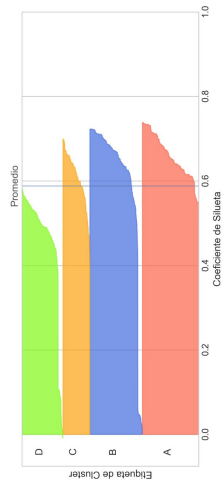
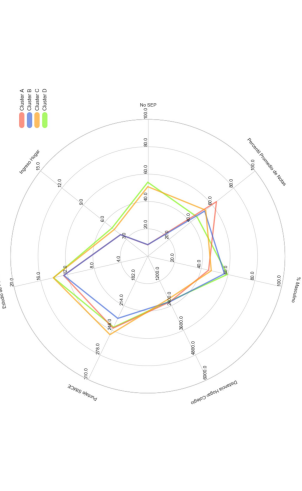


(a) Diagrama de coordenadas paralelas de variables agrupadas. (b) Flujo de cambios entre dependencias administrativas. (c) Gráfico circular del tamaño de cada cluster.

VARIABLES ESTUDIANTES



CALIDAD DE CLUSTERS



RANCAGUA - MACHALI

NO FORZADO

CONURBACIÓN

CAMBIO

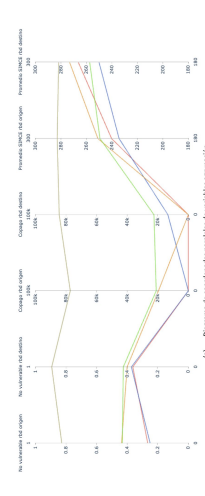
ESCENARIO

PRE-SAE

SAE-PREFERENCIAS

SAE-ADMISIÓN

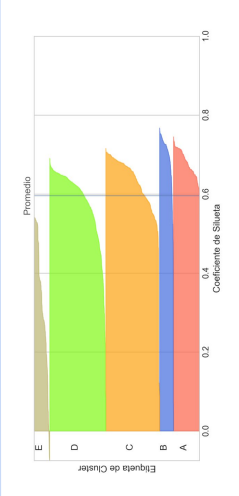
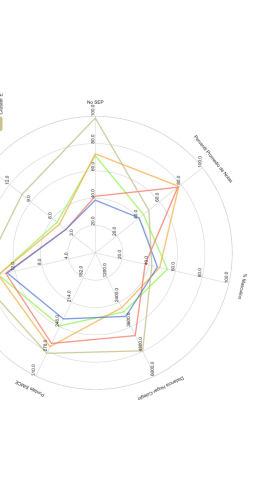
Clúster	Copango	IVE	Promedio SAE	Promedio SAE	Promedio SAE	Dependencia	Dependencia
	rd. origen	rd. destino	rd. origen	rd. destino	rd. origen	rd. destino	rd. destino
A	P90 (Gravado)	0,06	P90	P90	P90	Municipal	Part. Sube.
B	P90 (Gravado)	0,05	P90	P90	P90	Municipal	Part. Sube.
C	P90 (Gravado)	0,03	P90	P90	P90	Municipal	Part. Sube.
D	P90 (Gravado)	0,12	P90	P90	P90	Municipal	Part. Sube.
E	P90	0,17	P90	P90	P90	Municipal	Part. Sube.



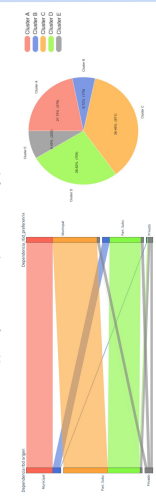
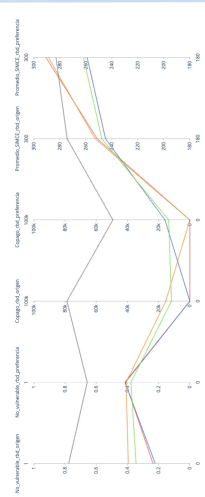
(a) Diagrama de coordenadas parciales de variables agrupadas.

(b) Flujo de cambios entre dependencias administrativas.

(c) Gráfico circular del tamaño de cada clúster.



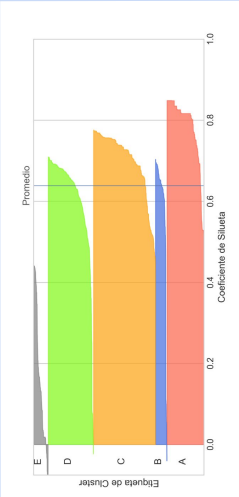
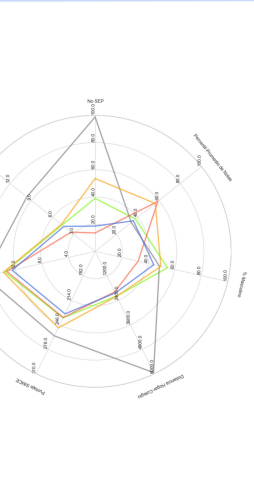
Clúster	Copango	IVE	Promedio SAE	Promedio SAE	Promedio SAE	Dependencia	Dependencia
	rd. origen	rd. preferencias	rd. origen	rd. preferencias	rd. origen	rd. destino	rd. destino
A	P90 (Gravado)	0,05	P90	P90	P90	Municipal	Part. Sube.
B	P90 (Gravado)	0,05	P90	P90	P90	Municipal	Part. Sube.
C	P90 (Gravado)	0,02	P90	P90	P90	Municipal	Part. Sube.
D	P90	0,02	P90	P90	P90	Municipal	Part. Sube.
E	P90	0,01	P90	P90	P90	Municipal	Part. Sube.



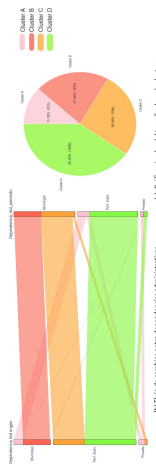
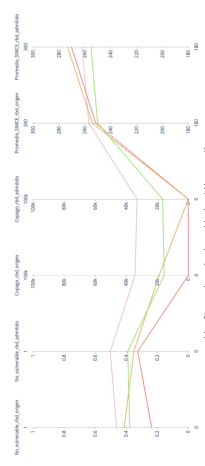
(a) Diagrama de coordenadas parciales de variables agrupadas.

(b) Flujo de cambios entre dependencias administrativas.

(c) Gráfico circular del tamaño de cada clúster.



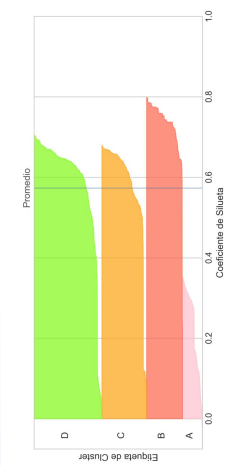
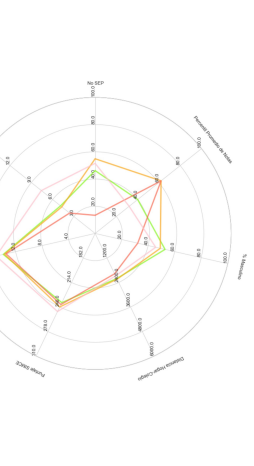
Clúster	Copango	IVE	Promedio SAE	Promedio SAE	Promedio SAE	Dependencia	Dependencia
	rd. origen	rd. admisión	rd. origen	rd. admisión	rd. origen	rd. destino	rd. destino
A	P90 (Gravado)	0,02	P90	P90	P90	Municipal	Part. Sube.
B	P90 (Gravado)	0,02	P90	P90	P90	Municipal	Part. Sube.
C	P90 (Gravado)	0,02	P90	P90	P90	Municipal	Part. Sube.
D	P90	0,02	P90	P90	P90	Municipal	Part. Sube.
E	P90	0,02	P90	P90	P90	Municipal	Part. Sube.



(a) Diagrama de coordenadas parciales de variables agrupadas.

(b) Flujo de cambios entre dependencias administrativas.

(c) Gráfico circular del tamaño de cada clúster.



PUERTO MONTT - PUERTO VARAS

FORZADO

CONURBACIÓN

CAMBIO

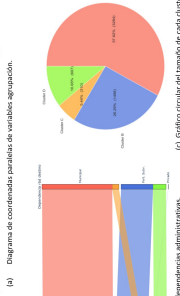
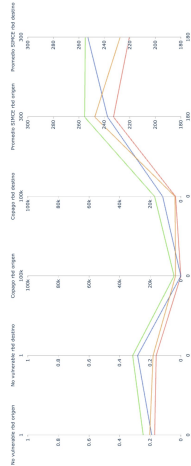
ESCENARIO

PRE-SAE

SAE-PREFERENCIAS

SAE-ADMISIÓN

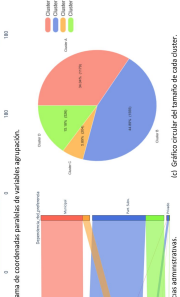
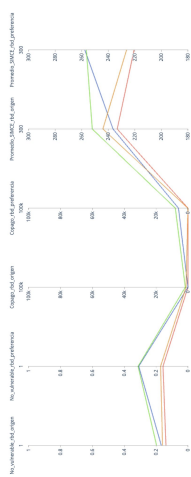
Cluster	Claves	INE	INVE	Procedimiento	SAE	Procedimiento	SAE	Procedimiento	SAE	Procedimiento	SAE	Procedimiento	SAE	Procedimiento	SAE	Procedimiento
	rdl destino	rdl origen	rdl destino	rdl origen	rdl destino	rdl origen	rdl destino	rdl origen	rdl destino	rdl origen	rdl destino	rdl origen	rdl destino	rdl origen	rdl destino	rdl origen
A	P01 (Gobierno)	0.83	P04	0.84	P01 (Gobierno)	0.86	P05	0.84	P01 (Gobierno)	0.85	P15	0.83	P01 (Gobierno)	0.85	P15	0.83
B	P02 (Gobierno)	0.75	P05	0.69	P02 (Gobierno)	0.82	P08	0.81	P02 (Gobierno)	0.82	P16	0.84	P02 (Gobierno)	0.82	P16	0.84
C	P03 (Gobierno)	0.75	P06	0.69	P03 (Gobierno)	0.82	P09	0.81	P03 (Gobierno)	0.82	P17	0.84	P03 (Gobierno)	0.82	P17	0.84
D	P04 (Gobierno)	0.75	P07	0.69	P04 (Gobierno)	0.82	P10	0.81	P04 (Gobierno)	0.82	P18	0.84	P04 (Gobierno)	0.82	P18	0.84



(a) Diagrama de coordenadas parciales de variables agrupadas.

(b) Flujo de cambios entre dependencias administrativas.

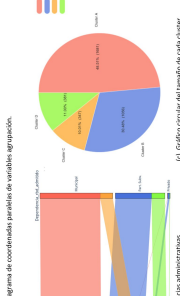
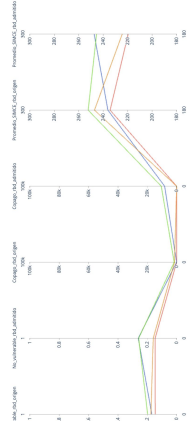
(c) Gráfico circular del tamaño de cada cluster.



(a) Diagrama de coordenadas parciales de variables agrupadas.

(b) Flujo de cambios entre dependencias administrativas.

(c) Gráfico circular del tamaño de cada cluster.

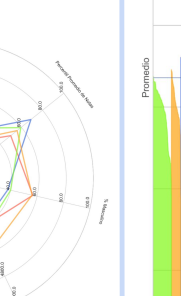


(a) Diagrama de coordenadas parciales de variables agrupadas.

(b) Flujo de cambios entre dependencias administrativas.

(c) Gráfico circular del tamaño de cada cluster.

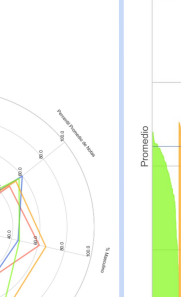
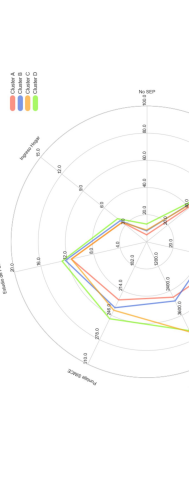
VARIABLES ESTUDIANTES



(a) Diagrama de coordenadas parciales de variables agrupadas.

(b) Flujo de cambios entre dependencias administrativas.

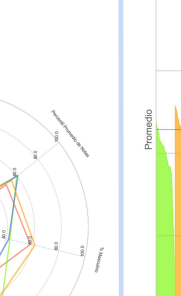
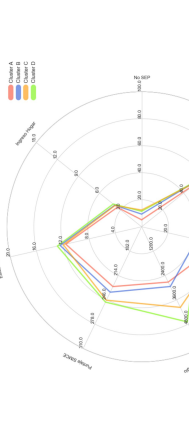
(c) Gráfico circular del tamaño de cada cluster.



(a) Diagrama de coordenadas parciales de variables agrupadas.

(b) Flujo de cambios entre dependencias administrativas.

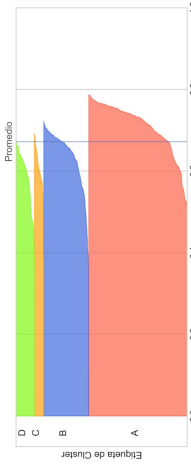
(c) Gráfico circular del tamaño de cada cluster.



(a) Diagrama de coordenadas parciales de variables agrupadas.

(b) Flujo de cambios entre dependencias administrativas.

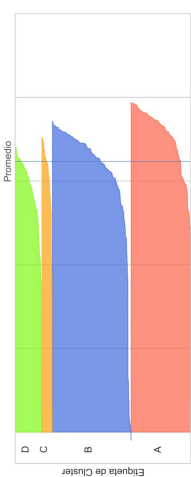
(c) Gráfico circular del tamaño de cada cluster.



(a) Diagrama de coordenadas parciales de variables agrupadas.

(b) Flujo de cambios entre dependencias administrativas.

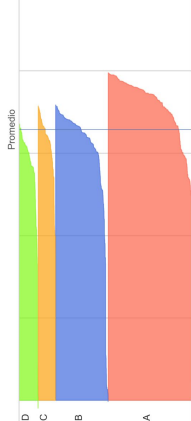
(c) Gráfico circular del tamaño de cada cluster.



(a) Diagrama de coordenadas parciales de variables agrupadas.

(b) Flujo de cambios entre dependencias administrativas.

(c) Gráfico circular del tamaño de cada cluster.



(a) Diagrama de coordenadas parciales de variables agrupadas.

(b) Flujo de cambios entre dependencias administrativas.

(c) Gráfico circular del tamaño de cada cluster.

CALIDAD DE CLUSTERS

PUERTO MONTT - PUERTO VARAS

CONURBACIÓN

CAMBIO

ESCENARIO

CENTROS CLUSTERS

VARIABLES COLEGIO

TAMAÑO CLUSTERS

VARIABLES ESTUDIANTES

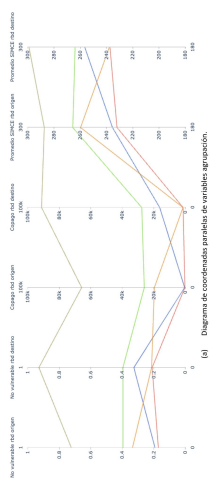
CALIDAD DE CLUSTERS

NO FORZADO

SAE-PREFERENCIAS

SAE-ADMISIÓN

Clase	Grupo	IVE	Promedio SAE	Promedio SAECE	Dependencia	Dependencia
		rel. origen	rel. destino	rel. origen	rel. destino	rel. destino
A	PS1 (Gratuito)	0.53	0.29	PS2	Municipal	PS1
B	PS2 (Gratuito)	0.53	0.29	PS3	Municipal	PS2
C	PS3 (Gratuito)	0.53	0.29	PS4	Municipal	PS3
D	PS4 (Gratuito)	0.53	0.29	PS5	Municipal	PS4
E	PS5 (Gratuito)	0.53	0.29	PS6	Municipal	PS5

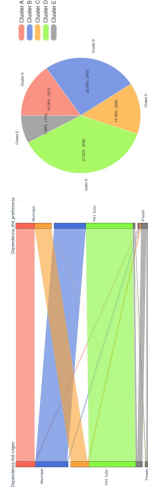
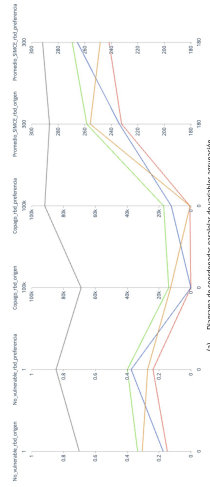


(a) Diagrama de coordenadas parciales de variables agregadas.

(b) Flujo de cambios entre dependencias administrativas.

(c) Gráfico circular del tamaño de cada cluster.

Clase	Grupo	IVE	Promedio SAE	Promedio SAECE	Dependencia	Dependencia
		rel. origen	rel. destino	rel. origen	rel. destino	rel. destino
A	PS1 (Gratuito)	0.51	0.27	PS2	Municipal	PS1
B	PS2 (Gratuito)	0.51	0.27	PS3	Municipal	PS2
C	PS3 (Gratuito)	0.51	0.27	PS4	Municipal	PS3
D	PS4 (Gratuito)	0.51	0.27	PS5	Municipal	PS4
E	PS5 (Gratuito)	0.51	0.27	PS6	Municipal	PS5

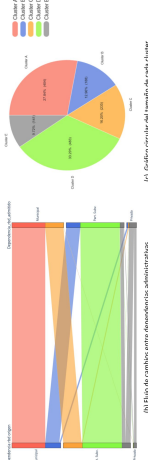
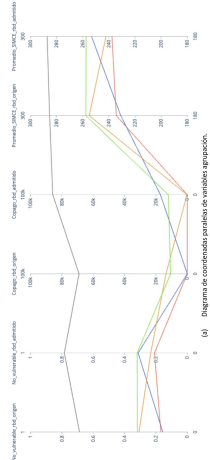


(a) Diagrama de coordenadas parciales de variables agregadas.

(b) Flujo de cambios entre dependencias administrativas.

(c) Gráfico circular del tamaño de cada cluster.

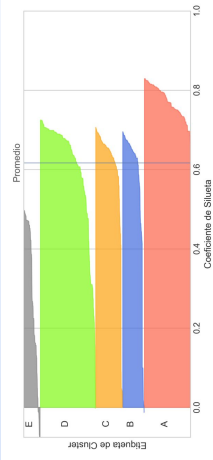
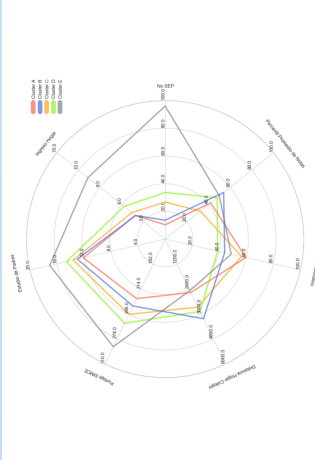
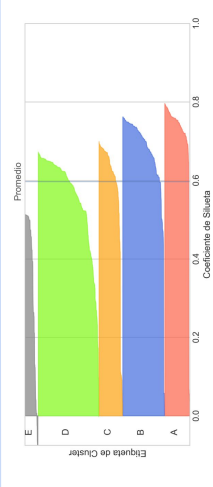
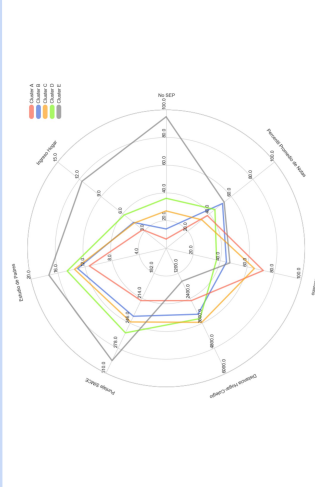
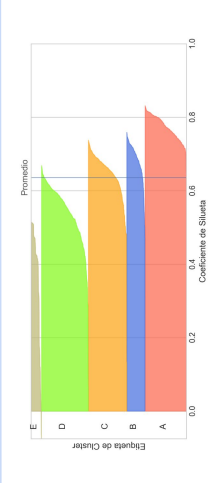
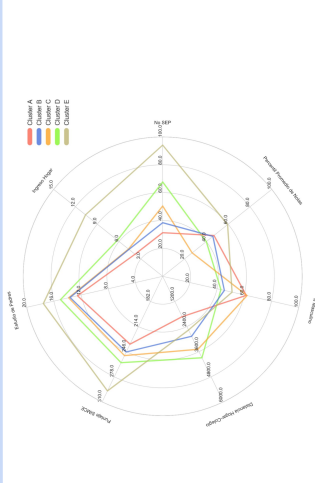
Clase	Grupo	IVE	Promedio SAE	Promedio SAECE	Dependencia	Dependencia
		rel. origen	rel. destino	rel. origen	rel. destino	rel. destino
A	PS1 (Gratuito)	0.58	0.28	PS2	Municipal	PS1
B	PS2 (Gratuito)	0.58	0.28	PS3	Municipal	PS2
C	PS3 (Gratuito)	0.58	0.28	PS4	Municipal	PS3
D	PS4 (Gratuito)	0.58	0.28	PS5	Municipal	PS4
E	PS5 (Gratuito)	0.58	0.28	PS6	Municipal	PS5



(a) Diagrama de coordenadas parciales de variables agregadas.

(b) Flujo de cambios entre dependencias administrativas.

(c) Gráfico circular del tamaño de cada cluster.



G. Códigos

```
def get_changed_students(dataframe):
    dataframe = dataframe[['agno', 'mrun', 'rbd']]
    # unique students
    unique_students = dataframe.mrun.unique()
    print('total =', len(unique_students))
    n = 0
    with open('/Users/Joaquin/Desktop/TESIS_CIAE/code/database/
    cambiados.csv', 'w') as file:
        writer = csv.writer(file)
        writer.writerow(['agno', 'mrun', 'rbd', 'rbd_elegido'])
        for student in unique_students:
            n = n + 1
            if n%1000==0:
                print(n)
            student_data = dataframe.loc[dataframe['mrun']==
            student].reset_index()
            student_data = student_data[['agno', 'mrun', 'rbd']]
            dataframe[dataframe.mrun != student]
            aparece = len(student_data)
            if aparece>1:
                for j in range(aparece-1):
                    rbd_elegido = student_data.get_value(
                        student_data.index[j+1], 'rbd')
                    rbd_actual = student_data.get_value(
                        student_data.index[j], 'rbd')
                    agno = student_data.get_value(student_data.
                    index[j], 'agno')
                    #print(agno)
                    data = student_data[j:j+1]
                    new_row = [agno, student, rbd_actual, rbd_elegido
                    ]
                    writer.writerow(new_row)

    file.close()
    cambiados.reset_index(drop=True, inplace=True)
    camb = pd.read_csv(dir + 'cambiados.csv')

    return camb

def plot_spider_chart(df_clust_dict):
    import plotly.graph_objects as go
    fig = go.Figure()
    for i in df_clust_dict.keys():
        categories =list(df_clust_dict[i].theta.values)
```

```

fig.add_trace(go.Scatterpolar(
    r=df_clust_dict[i].r,
    theta=categories,
    name='Cluster_' + i))

fig.update_layout(
    polar=dict(
        radialaxis=dict(
            visible=True,
            range=[0, 1]
        )
    ),
    showlegend=True
)

fig.show()

```

```

def plot_center_cluster_parcoords(centros_df, dict_cambio_index,
    my_colors):

    centros_df_2 = change_col_names_2(centros_df)

    centros_df_2['no_vulnerable_rbd'] = 1 - centros_df_2.ive
    centros_df_2['no_vulnerable_rbd_elegido'] = 1 - centros_df_2.
        ive_rbd_elegido

    centros_df_2.sort_values(by=['cod_depe2', '
        cod_depe2_rbd_elegido', 'copago_ficom'], inplace=True)
    centros_df_2.reset_index(inplace=True)
    centros_df_2['label'] = np.arange(len(centros_df))

    dict_by_items = {k: v for k, v in sorted(dict_cambio_index.
        items(), key=lambda item: item[1])}
    colors_ordered = []
    for col in dict_by_items.keys():
        colors_ordered = colors_ordered + [my_colors[col]]

    fig = go.Figure(data=
        go.Parcoords(
            line = dict(color = centros_df_2['label'], colorscale =
                colors_ordered), #,
            dimensions = list([
                dict(range = [0, 1],
                    label = 'No_vulnerable_rbd_origen', values =
                        centros_df_2['no_vulnerable_rbd']),
                dict(range = [0, 1],

```

```

        label = 'No_vulnerable_rbd_destino', values =
            centros_df_2['no_vulnerable_rbd_elegido']),
    dict(range = [0,100000],
        label = 'Copago_rbd_origen', values =
            centros_df_2['copago_ficom']),
    dict(range = [0,100000],
        label = 'Copago_rbd_destino', values =
            centros_df_2['copago_ficom_rbd_elegido']),
    dict(range = [180,300],
        label = 'Promedio_SIMCE_rbd_origen', values =
            centros_df_2['prom_simce']),
    dict(range = [180,300],
        label = 'Promedio_SIMCE_rbd_destino', values =
            centros_df_2['prom_simce_rbd_elegido']),
    ])
    )
)
fig.update_layout(
    plot_bgcolor = 'white',
    paper_bgcolor = 'white'
)
fig.update_layout(showlegend=True)

return fig

```

```

def plot_cod_depe_change_parcoords(data_2,y,dict_cambio_index,
my_colors):
    paralel_data = data_2.copy()
    paralel_data['label'] = y
    data_3 = change_col_names_2(paralel_data)

    dict_back = {'A':0, 'B':1, 'C':2, 'D':3, 'E':4, 'F':5, 'G':6, 'H':7}
    dict_cambio_numeros = {}
    for orig in dict_cambio_index.keys():
        new = dict_back[dict_cambio_index[orig]]
        dict_cambio_numeros[orig] = new

    data_3['label'] = data_3.label.map(dict_cambio_numeros)

    dict_by_items = {k: v for k, v in sorted(dict_cambio_index.
        items(), key=lambda item: item[1])}
    colors_ordered = []
    for col in dict_by_items.keys():
        colors_ordered = colors_ordered + [my_colors[col]]

    data_3.rename(columns={'cod_depe2': 'Dependencia_rbd_origen',

```

```

cod_depe2_rbd_elegido': 'Dependencia_rbd_destino'}, inplace=
True)

fig = px.parallel_categories(data_3.sort_values(by=['
Dependencia_rbd_origen', 'Dependencia_rbd_destino']),
dimensions= ['Dependencia_rbd_origen', 'Dependencia_rbd_
destino'],
color_continuous_scale=
colors_ordered, color='label',
)

fig.show()

def plot_spider_not_scaled(dict_spider, my_colors, name):
path = '/Volumes/Transcend/Tesis_jfarina/Base_de_datos/
repport_images/'
fig1 = plt.figure(figsize=(10, 10))
k = 0
for key in dict_spider.keys():

variables = ['No_SEP', 'Ingreso_Hogar', 'Estudio_de_Padres
', 'Puntaje_SIMCE', 'Distancia_Hogar-Colegio',
'%Masculino', 'Percentil_Promedio_de_Notas']

data = list(dict_spider[key].r[:7])
ranges = [(0,100), (0,15), (0,20), (150,310), (0,6000), (0,100)
,(0,100), (0,100), (0,100), (0,100), (0,100), (0,100)]

# plotting
radar = ComplexRadar(fig1, variables, ranges)
radar.plot(data, my_colors[k])
k+=1

plt.savefig(path + 'spider_' + name + ".svg", bbox_inches='tight'
)
plt.show()

```

```

## N_CLUSTERS OPTIMOS PARA CADA COMUNA Y CASO F Y NF - PRE SAE
n_optimal_pre_SAE= {}
# para cada comuna
for conurb in conurbaciones:
    con_name = conurb_names[str(conurb)]
    print(con_name)
    data_2 = data_all_pre_sae[data_all_pre_sae.cod_com_rbd.isin(
        conurb)]
    # casos f - nf
    for cambio in ['forzado', 'no_forzado']:
        print(cambio)
        if cambio == 'forzado':
            data = data_2[data_2.nivel_max_rbd == data_2.curso]
        else:
            data = data_2[data_2.nivel_max_rbd != data_2.curso]

    n_optimal = plot_silhouette_score(data, cont_feat_for_clust
        , schools_data_pre_sae_2[schools_data_pre_sae_2.
            cod_com_rbd.isin(conurb)])
    n_optimal_pre_SAE[con_name+'_'+cambio] = n_optimal

```

```
# ENTRENAR MODELO PRESAE PARA N_OPTIMO – LUEGO DEL SILHOUETTE PARA  
CADA CONURBACION
```

```
N_dataframe_pre_sae = pd.DataFrame(columns=features_radar)
my_colors_original = ['tomato', 'royalblue', 'orange', 'lawngreen',
                      'gray', 'violet', 'darkkhaki', 'pink', 'aqua', 'peru', 'darkred']
optimal_values_pre_SAE_dict = {}
dict_spiders_pre_sae = {}
dict_data_labels = {}
for conurb_cod in [conurbaciones[4]]:
    con_name = conurb_names[str(conurb_cod)]
    print(con_name)
    for case in ['_forzado', '_no_forzado']:
        print(case[1:])
        dict_data_labels[con_name + case] = {}
        key = con_name+case
        n_optimal = n_optimal_pre_SAE[key]
        data_2 = data_all_pre_sae[data_all_pre_sae.cod_com_rbd.
                                  isin(conurb_cod)]
        data_total_conurb = train_test_data_3[train_test_data_3.
                                                cod_com_rbd.isin(conurb_cod)]
        if '_no' in case:
            data = data_2[data_2.nivel_max_rbd != data_2.curso]
        else:
            data = data_2[data_2.nivel_max_rbd == data_2.curso]

        centros_df, centros_perc, labeled_data, model =
            get_centers_and_labels(n_optimal, data,
                                   cont_feat_for_clust, schools_data_pre_sae_2, clust_names)
        optimal_values_pre_SAE_dict[key] = [n_optimal, len(data),
                                             np.round(centros_df, 2), centros_perc, model]
        # get labeled data and venters
        centros_perc.sort_values(by=['cod_depe2', '
                                   cod_depe2_rbd_elegido', 'copago_ficom'], inplace=True)
        colors_for_case = get_colors_for_case(centros_perc,
                                               my_colors_original_dict)

        data_5 = merge_with_simce_and_distance(labeled_data,
                                                simce_tot, dist_16)

        # plot spider and percentile centers
        data_6 = data_5[features_radar+['label_all_data']]
        N_dataframe_pre_sae.loc[con_name+case] = data_6.count()
            [: -1]
        dict_data_labels[con_name + case]['pre_SAE'] = data_5
```

```

dict_cambio_index = get_dict_cambio_index(centros_perc ,
    clust_names)
print( 'data_info_:' ,data_6.info() )

spider_conurb_pre_SAE = make_spider_dict(n_optimal ,
    features_radar ,data_6 ,data_total_conurb , 'pre_SAE' ,
    center_name=dict_cambio_index)

dict_back = { 'A':0 , 'B':1 , 'C':2 , 'D':3 , 'E':4 , 'F':5 , 'G':6 , 'H':7 }
dict_cambio_numeros = {}
for orig in dict_cambio_index.keys() :
    new = dict_back[dict_cambio_index[orig]]
    dict_cambio_numeros[orig] = new
my_colors = [ colors_for_case[i] for i in [*
    dict_cambio_numeros.values() ] ]

name = 'PRESAE_' + con_name + case
plot_spider_not_scaled(spider_conurb_pre_SAE , my_colors ,
    name)

plot_pie_chart_single_cluster_size(centros_perc [ 'N' ] .
    values , dict_cambio_index , my_colors)
dict_spiders_pre_sae[key] = spider_conurb_pre_SAE

centros_perc.index = clust_names[:len(centros_perc)]

centros_perc.drop(columns=[ 'N' , 'N_%' ] , inplace=True)
centros_perc.columns = new_cols_centros_perc
display(centros_perc)

centros_perc.to_csv(path_2 + name + "_centros_perc.csv" ,
    index_label='Cluster')

# plot clustere uno a uno
scaler = MinMaxScaler()
data_2 = data[cont_feat_for_clust].dropna()
X = scaler.fit_transform(data_2)
y = model.predict(X)
fig , ax1= plt.subplots( figsize=(10,4) , dpi=300)
ax1 = silhouette_plot(X , y , dict_cambio_index , my_colors ,
    n_clusters=n_optimal , ax=ax1)
plt.show()

# plot cod_dpe2
plot_cod_depe_change(data_2 , y , dict_cambio_index , my_colors)

```

```
# plot centro cluster
centros_df_plot_fig = plot_center_cluster(centros_df,
    dict_cambio_index, my_colors)
centros_df_plot_fig.show()
plot_legend(my_colors, my_colors_original, legend_elements)
```