



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

HERRAMIENTA PARA LA SIMULACIÓN DE TRAYECTORIAS ESTUDIANTILES:
APLICACIÓN AL CÁLCULO DE LOS INGRESOS POR GRATUIDAD DE LA
UNIVERSIDAD DE CHILE

TESIS PARA OPTAR AL GRADO DE MAGÍSTER EN GESTIÓN DE OPERACIONES

MEMORIA PARA OPTAR AL TÍTULO DE INGENIERO CIVIL INDUSTRIAL

CRISTIÁN ADOLFO VELOSO DORNER

PROFESOR GUÍA:

SR. RAFAEL EPSTEIN NUMHAUSER

PROFESOR CO-GUÍA:

SR. MARCEL GOIC FIGUEROA

MIEMBROS DE LA COMISIÓN:

SR. RODOLFO URRUTIA URIBE

SRA. ANGÉLICA BOSCH CARTAGENA

SANTIAGO DE CHILE
2020

Resumen Ejecutivo

En mayo del 2018 se promulgó la Ley de Educación Superior, que comenzó a implementarse el año 2016 por medio de modificaciones al presupuesto de la nación. Sin embargo, hasta el día de hoy no existe claridad sobre su impacto.

Este trabajo busca aportar a la evaluación del impacto económico de la gratuidad universitaria en la Universidad de Chile y a la modelación del comportamiento estudiantil, por medio del estudio de la trayectoria de una cohorte de ingreso (grupo de estudiantes que comenzaron sus estudios universitarios el mismo año) de alumnos beneficiados con gratuidad de la Facultad de Derecho y la Facultad de Ciencias Físicas y Matemáticas.

De este modo, se utilizó la totalidad de información disponible de la primera cohorte beneficiada de las facultades mencionadas, y se completó el resto de su trayectoria mediante simulaciones con parámetros ajustados con trayectorias reales de cohortes que han cumplido con holgura el tiempo necesario para terminar sus programas.

La trayectoria se modeló mediante transiciones de estado en tiempo discreto. Las probabilidades de transición se calcularon a nivel individual, para así encontrar el impacto en la deserción y el tiempo hasta la completitud del título, de ciertas variables socioeconómicas y de rendimiento académico. Las principales variables predictoras resultaron ser las de rendimiento del año más cercano al periodo a estimar.

A través de la simulación se determinaron los ingresos totales proyectados por gratuidad para las facultades y cohortes estudiadas. Sin embargo, se considera que el mayor aporte corresponde a los resultados de las sensibilizaciones de parámetros de la simulación, que permiten entender los órdenes de magnitud de los impactos que pueden tener cambios en el reglamento de la gratuidad o del comportamiento de los estudiantes.

Así, entre otras sensibilizaciones, se determinó que extender un año el beneficio tendría un aumento de alrededor del 9% en los ingresos. Acortar las carreras en un año reduciría cerca del 11% los ingresos (se puede contrarrestar aumentando matrícula). Y que cambios esperables en los niveles de retención y composición no tendrían efectos considerables.

Agradecimientos

A mi familia, que me ha llenado de amor y apoyo incondicional durante todos los años de mi vida. Faltan las palabras para expresar lo agradecido y afortunado que me siento por tenerlos, demás está decir que este logro (un poco postergado), se debe principalmente a ustedes.

A mi Florencia, que ha sido una guía, una inspiración y una fuente inagotable de alegrías desde el colegio y esperemos que por mucho tiempo más.

A mis amigos, que pase lo que pase, siempre estarán ahí. En particular a los Posos queridos. También a las maravillosas personas que conocí en la U; a Las Precisas y los del fútbol, que hicieron que el paso por la universidad sea una etapa plagada de buenos recuerdos.

Agradezco también, a Luis San Martín, por su desinteresada y fundamental colaboración para la realización de este trabajo.

Por último y no menos importante, al Club Social y Deportivo Colo Colo, por darme tantas alegrías y hacer de la marraqueta el pan más crujiente del mundo por tantos lunes.

Cristián Adolfo Veloso Dorner

Tabla de contenido

Resumen Ejecutivo	I
Agradecimientos	II
1. Introducción	1
1.1. Contexto General y Motivación	1
1.2. La Universidad de Chile	3
1.3. Gratuidad	4
1.3.1. Breve historia	4
1.3.2. Situación actual	5
1.4. Objetivos del trabajo	6
1.4.1. Objetivo general	6
1.4.2. Objetivos específicos	6
2. Trayectoria y gratuidad	8
2.1. Impacto económico de la gratuidad	8
2.1.1. Universo de estudio	9
2.2. Trayectoria en la FCFM y Derecho	10
2.2.1. Facultad de Derecho	10
2.2.2. FCFM	11
3. Análisis Exploratorio de los datos	12
3.1. Descripción de las variables	12
3.1.1. Variable dependiente - Estado por semestre	12
3.1.2. Variables de caracterización de los estudiantes	14
3.1.3. Variables de rendimiento académico	16
4. Modelo Matemático de la Trayectoria Estudiantil	17
4.1. Revisión de literatura	17
4.2. Modelamiento	20

4.2.1. Postergaciones	20
4.2.2. Estados, transiciones y periodos	22
4.2.3. Probabilidades de transición	23
4.2.4. Censura	24
5. Cálculo de probabilidades de transición	25
5.1. Análisis exploratorio con modelos no paramétricos	25
5.1.1. Segmentación por variables socioeconómicas	28
5.1.2. IVE Colegio de egreso	29
5.1.3. Clusterización de rendimientos	29
5.1.4. Estimación no paramétrica para los Clusters de rendimiento	30
5.2. Estimación de probabilidades de transición con modelos Logit	31
5.2.1. Resultados Facultad de Derecho	32
5.2.2. Resultados Facultad de Ciencias Físicas y Matemáticas	35
5.3. Sobre las variables predictoras	37
6. Simulación y evaluación económica	40
6.1. Simulación	40
6.2. Resultados de la simulación	41
6.2.1. Facultad de Derecho	41
6.2.2. Facultad de Ciencias Físicas y Matemáticas	43
7. Sensibilización de los Parámetros Críticos	48
7.1. Arancel Regulado y Duración de Beneficio	48
7.2. Retención	49
7.3. Duración	50
7.4. Composición de la matrícula	51
8. Trabajos Futuros y Conclusiones	53
8.1. Trabajos futuros: Estimación del impacto en el comportamiento de los estudiantes	53
8.2. Conclusión	54
Bibliografía	57
Apéndices	59
A . Cálculo de Arancel de Referencia y Regulado	59
A .1. 1. Constitución de los grupos de universidades	59
A .2. 2. Definición aranceles de referencia originales por carrera.	59

A .3. 3. Definición aranceles regulados	59
B . Aranceles Regulados Facultad de Derecho y FCFM	60
C . Marco Conceptual - Glosario	60
C .1. Estimación por Máxima Verosimilitud	60
C .2. Probabilidades de transición	60
C .3. Heterogeneidad observable y no observable	61
C .4. Regresión logística Multinomial - Logit Multinomial	61
C .5. Modelo de clases latentes	61
C .6. Criterios de selección AIC y BIC	61
C .7. Algoritmo Kmedias	62
D . Algoritmo kmedias - Regla del codo	62
E . Resultados Logit	63
E .1. Clases Latentes Derecho	63
E .2. Elección Programa de Continuidad Ingeniería	63

Índice de cuadros

3.1. Cantidad de Estados Anuales por Estudiante	13
3.2. Cantidad de Alumnos de las cohortes 2009-2011 que continuaron a cada carrera de la FCFM.	13
3.3. Promedio PSU y cantidad alumnos por IVE.	15
3.4. Región de origen de los alumnos, cohortes 2009, 2010 y 2011.	15
3.5. Alumnos con gratuidad.	16
3.6. Ramos aprobados y reprobados al año por sexo y facultad - Cohortes 2009, 2010 y 2011.	16
4.1. Alumnos Postergados de las cohortes 2009, 2010 y 2011, por año de postergación. . .	21
4.2. Cantidad de alumnos de las cohortes 2009, 2010 y 2011 con al menos un proceso de postergación clasificados por la cantidad de años inactivos.	22
5.1. Tabla de vida - FCFM.	26
5.2. Tabla de vida - Facultad de Derecho.	27
5.3. Modelos Logit sin heterogeneidad no observable	33
5.4. Modelos de regresión lineal para predicción de rendimiento académico hasta 3er año	38
5.5. Tabla contingencia Gratuidad vs IVE colegio egreso	38
5.6. Tabla contingencia Gratuidad vs Región colegio de egreso	39
5.7. Puntaje PSU alumnos con Gratuidad y sin Gratuidad	39
6.1. Distribución de Alumnos FCFM por carrera en simulaciones	44
6.2. Promedio y Desviación Estándar de la duración por carrera	46
7.1. Ingresos adicionales al extender cobertura por gratuidad, sin considerar ingresos post-duración	49
7.2. Sensibilización Tasa de Retención Facultad de Derecho	50
7.3. Resultados Sensibilización de Adelantar y Retrasar las probabilidades de Egreso	50
7.4. Sensibilización comparación alumno con y sin gratuidad	51

8.1. Resultados Clases Latentes Derecho	63
8.2. Resultados Clases latentes con heterogeneidad en parámetros temporales e intercepto	64
8.3. Resultados Modelo Seleccionado Para Elección de Carrera 1/2	65
8.4. Resultados Modelo Seleccionado Para Elección de Carrera 2/2	66

Índice de figuras

1.1. % PIB Educación Superior por fuente de aporte Fuente: OECD - 2018	5
3.1. Distribución Temporal de Eliminaciones y Egresos por Mes	14
4.1. Diagrama compacto de transiciones a modelar en este trabajo	22
4.2. Diagrama de transiciones a modelar en este trabajo	23
4.3. Secuencia para ingeniería.	23
5.1. Probabilidades de transición FCFM y Facultad de Derecho	26
5.2. Tasa de Riesgo de Egreso y Eliminación de los estudiantes de regiones vs de la región metro- politana	28
5.3. Tasa de Riesgo de Egreso y Eliminación de los estudiantes por grupo IVE de establecimiento de egreso	29
5.4. Tasa de Riesgo de Egreso y Eliminación de los distintos Clusters	30
5.5. Proporción de Estado promedio Modelo vs Realidad	34
5.6. Probabilidad de Egreso y Eliminación Por Carrera de la FCFM por Semestre	37
6.1. Resultado de la Simulación para la Facultad de Derecho	42
6.2. Distribución Eliminaciones y Egresos y Supervivencia Alumnos Derecho	42
6.3. Distribución Eliminaciones y Egresos Alumnos Con Gratuidad vs Alumnos Sin Gratuidad	43
6.4. Simulación Ingresos Gratuidad Facultad de Derecho	44
6.5. Resultados de la Simulación para la Facultad de Ciencias Físicas y Matemáticas	45
6.6. Distribución Eliminaciones, Egresos y Supervivencia Alumnos FCFM	45
6.7. Distribución Eliminaciones, Egresos y Supervivencia Alumnos FCFM	46
6.8. Simulación Ingresos Gratuidad Facultad de Ciencias Físicas y Matemáticas	47
7.1. Cambios en los ingresos por gratuidad al variar 1 peso el arancel regulado Derecho	48
8.1. Aranceles reales y regulados de las carreras de la FCFM	60
8.2. Regla del codo suma de varianza dentro del cluster, clusterización por rendimiento FCFM	62

Capítulo 1

Introducción

1.1. Contexto General y Motivación

El sistema de educación superior chileno ha experimentado importantes cambios en las últimas décadas, se ha buscado mejorar los niveles de calidad, aumentando las instancias de fiscalización y promoviendo a través de incentivos que las instituciones alcancen mejores estándares. Sin embargo, el mayor esfuerzo ha estado en la cobertura, buscando garantizar el acceso a todas las personas que tengan las capacidades para obtener un título de educación superior, sin importar su capacidad de pago. Si en la década de los '80 el sistema contaba con poco más de 100.000 alumnos repartidos sólo en 8 universidades, para el año 2018 esa cifra casi alcanzó los 1.200.000 alumnos¹, repartidos entre universidades, centros de formación técnica e institutos profesionales.

En concordancia con lo descrito en el párrafo anterior, el año 2015 se promulgó la Ley 20.890 o «Ley corta de gratuidad», con modificaciones en los siguientes años y punto cúlmene el 2018 con la promulgación de la Ley 21.091 o «Ley de Educación Superior», que busca entender la educación superior como un derecho social, al alcance de todas las personas, de acuerdo a sus capacidades y mérito, liberando del pago de aranceles a los estudiantes de los niveles socioeconómicos más bajos, pero a medida que el PIB del país crezca, extenderlo como un derecho universal. Al año 2019 el beneficio puede ser recibido por los alumnos de los primeros tres quintiles (según formulario FUAS² e información estatal pertinente) que estudien en las universidades adscritas a la gratuidad.

Para que una Universidad esté adscrita a la gratuidad debe pertenecer a uno de los siguientes subconjuntos de universidades³:

- Estatales.
- Privadas tradicionales (comprendida en el art. N° 1 del DFL N°4 de 1981).

¹Datos del Consejo Nacional de Educación - <https://www.cned.cl/indices-educacion-superior> - 2019.

²Formulario Único de Acreditación Socioeconómica

³Ley 21.091: Sobre Educación Superior.

- Privadas acreditadas por al menos 4 años y estar constituidas como persona jurídica sin fines de lucro o en caso de tener controlador(es), que estos sean personas naturales o jurídicas sin fines de lucro.

Además, el 80% de sus alumnos deben contar con un promedio simple mayor o igual a los 450 puntos, entre su puntaje Prueba de Selección Universitaria (PSU) de Matemáticas, PSU de Lenguaje, Puntaje de notas de enseñanza media (NEM) y puntaje de ranking de notas⁴. Por último, se exige que las universidades no estatales estén adscritas al Sistema Único de Admisión o tener un sistema de admisión transparente, objetivo y que no implique discriminaciones arbitrarias, basado en el mérito de los estudiantes.

En caso de que un alumno cumpla con los requisitos e ingrese a una universidad adscrita, entonces podrá optar al beneficio. El Estado aporta un monto fijo por alumno determinado para cada carrera y universidad, llamado «Arancel Regulado», que suele ser menor al arancel real cobrado por las instituciones a los alumnos que no reciben este beneficio. Las universidades no pueden realizar cobros adicionales al «Arancel Regulado» a los beneficiados.

Actualmente, no existe total certeza sobre el impacto que tiene esta medida en la situación financiera de las universidades. Algunas no se han adscrito debido a la incertidumbre⁵ y dificultades económicas que pueda producir esta forma de financiamiento, que por un lado limitará la capacidad de las universidades de cobrar libremente, pero por otro aumentará el aporte estatal (antiguamente el monto referencial era menor al arancel regulado) y disminuirá las cuentas incobrables de las instituciones (problemática no abordada en este trabajo).

Se plantea que es de gran utilidad para entender la magnitud de los ingresos, revisar el monto total que ingresa a la universidad por este concepto, o calcular el ingreso asociado a una cohorte⁶ por medio de fórmulas simples como:

$$\text{Ingresos Carrera} = \text{Estudiantes con Gratuidad} \times \text{Duración media} \times \text{Arancel Regulado} \quad (1.1)$$

Sin embargo, al calcular los efectos de esa manera no se consideran los posibles cambios que se pueden haber producido en los estudiantes tras la aplicación de este beneficio. Además, estas fórmulas no se hacen cargo de las diferencias entre los alumnos con gratuidad y los que no la tienen.

La Universidad de Chile, como la principal universidad pública del país, ha tenido un rol activo en las negociaciones con el Estado, pero es fundamental tener una mayor claridad sobre el impacto real para poder tener una postura más sólida.

⁴Puntaje otorgado por posición relativa del rendimiento académico de los estudiantes en su establecimiento escolar

⁵Universidad de los Andes, Universidad del Desarrollo, Universidad Adolfo Ibañez, entre otras

⁶cohorte de ingreso se refiere a todos los estudiantes que ingresaron a la universidad el mismo año

Dado este contexto, en este trabajo se propone una forma de medición del impacto a través de la proyección de la trayectoria de los estudiantes, considerando que las primeras cohortes con gratuidad desde primer año aún no terminan sus estudios.

1.2. La Universidad de Chile

La Universidad de Chile es una institución de educación superior pública. Es la casa de estudios más antigua del país y a lo largo de su historia ha sido un pilar fundamental en el desarrollo del mismo. Está compuesta por 20 unidades académicas que en la práctica funcionan con bastante autonomía; 14 Facultades, 5 Institutos dependientes de Rectoría y el Hospital Clínico de la Universidad de Chile, que cultivan todas las áreas disciplinarias, impartiendo más de 71 programas de pregrado, 107 Magíster, 81 Títulos de Profesional Especialista y 38 Doctorados⁷.

Durante el año 2017 la comunidad universitaria estuvo compuesta por 3.825 académicos, 41.693 estudiantes; 32.422 alumnos de pregrado y 9.271 estudiantes de postgrado. Además de 9.234 funcionarios que apoyan en tareas de servicios, administración y otras tareas especializadas, concentrados principalmente en el IDIEM y Hospital Clínico.

Los resultados del quehacer universitario se pueden agrupar en 4 actividades principales: La docencia de Pregrado, la docencia de Postgrado, la Investigación y la Extensión, además de la Administración y Gestión, que funciona como soporte de las anteriores.

Actualmente, no existe una gestión de costos que permita identificar las actividades sustentadas por los distintos gastos. Por ejemplo, se puede mencionar que el mayor gasto universitario corresponde al salario de los académicos, pero como no se cuenta con un sistema para definir qué fracción de los salarios debería ser asignada como costo de docencia de pregrado, resulta imposible realizar una evaluación económica que tome en consideración los cambios en los costos⁸. Por lo anterior, el análisis que se llevará a cabo en este trabajo, sólo tomará en cuenta los cambios en los ingresos como consecuencia de la aplicación de esta ley.

Con respecto a la formación de pregrado de la Universidad de Chile, se puede decir que es altamente selectiva; para entrar a cualquier carrera se necesita una alta ponderación en la Prueba de Selección Universitaria. Sin embargo, existen cupos destinados a ingresos especiales (cupos para deportistas, cupos para indígenas, cupos por nivel socioeconómico, cupos para extranjeros, etc.) para facilitar el acceso a individuos de grupos, que por diferentes motivos, están subrepresentados o se ven en desventaja al momento de postular a la universidad por la vía regular. Anualmente, ingresan en

⁷Oferta a diciembre de 2017.

⁸Se considera que es importante que la Universidad tenga una actualización en la forma de gestionar los costos, pero escapa al alcance de este proyecto (Ver sistemas de costeo ABC o TDABC).

torno a 6.500 alumnos (considerando todas las vías de ingreso), de los cuales aproximadamente 29 % provienen de colegios municipales, 31 % de colegios particulares subencionados y 41 % de colegios particulares. La matrícula total es de cerca de 32.000 alumnos⁹. Los egresados tienen tasas de empleabilidad y salarios bastante mayores a la media nacional¹⁰.

1.3. Gratuidad

1.3.1. Breve historia

Desde hace bastante tiempo el foco de las políticas públicas con respecto a la educación superior ha estado en la cobertura. De esta forma, se puede mencionar que a principios de los años 80 la matrícula total en el sistema de educación superior era menor a 120.000 estudiantes, mientras que el año 2018 esa cifra alcanzó los 1.176.915 alumnos.

Esto ocurrió debido a diferentes factores, como el aumento de la oferta, la desregulación del sistema, la educación secundaria obligatoria, la aparición de nuevas formas de financiamiento, entre otras. Si, en 1980 existían sólo 8 universidades en todo Chile (la Universidad de Chile, la Universidad Técnica del Estado y la Universidad Católica tenían además 17 sedes regionales), a marzo del 2018 esta cifra ascendía a 150 instituciones vigentes.

El movimiento estudiantil del 2011, logró poner énfasis en la regulación del sistema, donde las principales críticas apuntaban a la mercantilización y los altos intereses de los créditos entregados por la banca e instituciones con bajos niveles de calidad que lucraban gracias a distintas artimañas, pese a que por ley estaba prohibido. Todo esto, en un contexto con aranceles muy altos y sin directa relación con la calidad o costo de la educación entregada, sino más bien con la disposición a pagar de los estudiantes, marcada por tendencias históricas y proyección de salarios futuros. En la imagen 1.1 se puede apreciar la proporción del PIB destinada a educación superior, segmentada según la proveniencia del aporte; privado, público o transferencia pública a los privados. Se observa que Chile tiene el segundo gasto más alto y sobresale la gran proporción financiada por los privados a través de los aranceles.

Gracias a las presiones sociales se decretaron diversas leyes que han mejorado las instancias de fiscalización y regulado las formas de financiamiento.

⁹Promedio años 2016 a 2018

¹⁰Servicio de Información de Educación Superior - www.mifuturo.cl

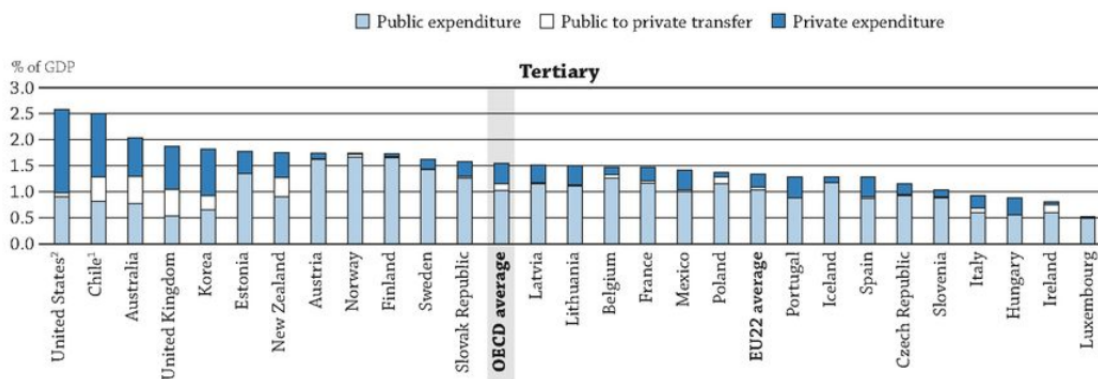


Figura 1.1: % PIB Educación Superior por fuente de aporte Fuente: OECD - 2018

1.3.2. Situación actual

Tramitación ley

La gratuidad entró en funcionamiento por primera vez el año 2016 a través de una modificación de la Ley N°20.882 de Presupuesto del Sector Público 2016. Para el año 2017 y 2018, su inclusión también fue por glosas presupuestarias que debían ser aprobadas año a año en el congreso. Sin embargo, el 2018 se promulgó la Ley 21.091 Sobre Educación Superior, cuyos principales cambios fueron la creación de una Subsecretaría de Educación Superior, el diseño de un Nuevo Sistema de Aseguramiento de calidad y la consagración de la gratuidad como ley. En esta consolidación se establecieron las reglas para su expansión a más deciles, lo cual se irá implementando según el crecimiento económico del país.

Funcionamiento para el año 2019

La gratuidad sólo cubre a los alumnos de los primeros 6 deciles socioeconómicos (en los CFT e IP cubre a los alumnos de los primeros 7 deciles), y que estudien en instituciones adscritas a la gratuidad. Los estudiantes son evaluados con el Formulario Único de Acreditación Socioeconómica (FUAS), además el Ministerio de Educación (Mineduc) tiene las atribuciones para verificar y complementar la información proporcionada por los estudiantes con antecedentes de otras entidades públicas y privadas.

El Estado aporta un monto establecido por el arancel regulado que varía por carrera y universidad. Para su cálculo se utiliza la metodología publicada por el Mineduc [2018]. Por el momento, el principal insumo para el cálculo es el arancel referencial. Sin embargo, el año 2019, a través de la Alta Dirección Pública, se formó una comisión de expertos para definir las bases técnicas con que se calcularán estos montos sumamente críticos.

Desde el 2019 se limitó el arancel cobrable a los alumnos de los deciles 7, 8 y 9, estableciendo como máximo un factor de 1,4 sobre el arancel regulado para los alumnos del decil 7 y 1,6 sobre el arancel regulado para los alumnos de los deciles 8 y 9. Así, las universidades sólo tendrán libertad para fijar los aranceles a los alumnos del decil 10. En otras palabras, si arancel regulado es 100, a los alumnos del decil 7 se les podrá cobrar máximo 140 y a los del decil 8 y 9 máximo 160, mientras que existe libertad en el monto a cobrar a los alumnos del decil 10.

Finalmente, es importante mencionar que la gratuidad acepta sólo un cambio de carrera, y cubre únicamente por la duración formal del programa, lo que para los estudiantes puede ser un gran problema, pues en general los tiempos reales para completar los estudios superan las duraciones formales. Sin embargo, los alumnos pueden optar a otros beneficios estatales cuando ya no sean candidatos a la gratuidad, como el crédito solidario o las becas bicentenario. Para las universidades este punto es muy complejo, ya que una vez finalizado el financiamiento por gratuidad, tienen permitido como máximo cobrar el 50% del arancel regulado.

En el caso de las continuidades de estudio, como las que se dan con los bachilleratos o en carreras de entrada a un plan común, el periodo de cobertura se recalcula al cambiar de carrera. Además se aceptan postergación de estudio, sólo en caso de existir una razón de fuerza mayor aprobada por el Mineduc.

1.4. Objetivos del trabajo

Dado este contexto, en donde se presenta una de las mayores reformas ocurridas en Chile en los últimos años, pero sobre la que no existen consensos sobre sus impactos, se definen los siguientes objetivos:

1.4.1. Objetivo general

Desarrollar una herramienta que permita simular las trayectorias estudiantiles de alumnos de la Universidad de Chile. A partir de esto, estimar los impactos económicos por los ingresos directos de la gratuidad.

1.4.2. Objetivos específicos

1. Implementar la metodología en un subconjunto de estudiantes y calcular los ingresos proyectados.
2. Modelar aspectos específicos: retención/graduación.
3. Aplicar modelos con distintas características y evaluación de su calidad.

4. Sensibilizar los parámetros críticos y medir su impacto económico.

Capítulo 2

Trayectoria y gratuidad

2.1. Impacto económico de la gratuidad

La implementación de gratuidad en la educación superior es una de las grandes reformas de los últimos tiempos en el ámbito educacional. Sin embargo, no existe certeza sobre los múltiples efectos que esta tendrá, pero como se comentó anteriormente, este trabajo pretende ser un aporte al entregar una metodología que cuantifique el impacto en los ingresos institucionales.

Medir el impacto económico de la gratuidad en todas sus dimensiones tiene ciertas dificultades que se resumen a continuación:

1. Su implementación ha sido muy reciente y no existen cohortes que hayan tenido gratuidad por todo su periodo de estudio: La primera cohorte beneficiada por la gratuidad desde primer año, fue la cohorte 2016, por lo que la información puede ser insuficiente para encontrar cambios en el comportamiento de los estudiantes influenciados por la gratuidad.
2. Ya existían otros beneficios que fueron reemplazados y que hoy en día son alternativas complementarias: Si bien es valioso tener una noción de los ingresos totales asociados al beneficio, para analizar cómo afecta las finanzas de las universidades puede ser más importantes las variaciones con respecto al estado anterior. Esto se puede dificultar debido a la dinámica de los sistemas de información, que pueden ser sustituidos o se puede dejar de almacenar información necesaria para hacer la comparación.
3. Siguen existiendo modificaciones a su aplicación y las reglas han variado a lo largo de los años: Este punto puede complicar las mediciones, pues al haber cambios los sistemas de información se pueden adecuar y dejar/comenzar a almacenar información necesaria para medir los efectos de forma consistente.

De esta forma, Se decidió calcular los ingresos totales por el concepto de gratuidad, para una cohorte que haya dispuesto del beneficio desde su ingreso a la universidad. Para hacer esto, es

necesario conocer el tiempo en que los estudiantes comienzan y terminan sus estudios. Sin embargo, la gran mayoría de los estudiantes de la primera cohorte que entró a la universidad con gratuidad (cohorte 2016), siguen siendo alumnos regulares. Es decir, no se conoce el momento en que los estudiantes que aún están activos abandonarán sus estudios o egresarán. Hasta el momento 112 de los 1.279 alumnos que ingresaron ya han abandonado su carrera, sin embargo se espera que en los próximos años más alumnos lo hagan. Por último, tampoco se puede saber el momento exacto de su egreso o abandono, debido a que dependen del cumplimiento de requisitos y esto presenta aleatoriedad. Ante esta situación, se decidió simular el resto de las trayectorias de dichos alumnos.

Para calibrar los parámetros de la simulación, se analizó la trayectoria de estudiantes de cohortes de ingreso que hayan superado con holgura el tiempo de duración formal de sus programas. Luego se estimó cómo las características almacenadas en sistema impactaron sus trayectorias. Es relevante y posible distinguir la probabilidades de egreso y eliminación a nivel estudiante porque se sabe quienes son los alumnos de la cohorte 2016 que tienen gratuidad, por lo que se utilizará su heterogeneidad para tener estimaciones más ajustadas a su realidad.

Un punto central de este trabajo, que justifica la utilización de modelos más sofisticados para estimar las probabilidades de eliminación y egreso por periodo, en vez de ocupar simplemente las proporciones muestrales de la población con gratuidad de las cohortes 2009, 2010 y 2011, reside en que no se conoce con exactitud los alumnos de dichas cohortes, que cumplen o cumplían los requisitos para contar con este beneficio. Si se contará con ese dato, se podría calcular la proporción para sólo ese subconjunto de estudiantes y agregar dicha variable a los modelos de predicción.

Se considera que el gran valor que tiene este enfoque está dado porque permitirá hacer evaluaciones contrafactuales, simulando el efecto de variaciones en parámetros críticos como el valor del arancel regulado, los tiempos de estudios o retención. Así, se podrá simular por ejemplo, el efecto en los ingresos de la universidad al aumentar o disminuir la probabilidad de eliminación de cada estudiante, adelantar o retrasar el momento de egreso y muchas otras alternativas.

Es importante mencionar que, para que esta metodología tenga una mayor validez, es necesario que se cumplan ciertos supuestos que tienen que ver con los efectos de la gratuidad en el comportamiento de los estudiantes, pues cuanto más se parezcan los comportamientos de los alumnos pre y post gratuidad, más robustos serán los resultados encontrados. Esta temática se aborda con mayor detalle en el capítulo 7.

2.1.1. Universo de estudio

Los datos utilizados fueron proporcionados por el Departamento de Pregrado, dependiente de la Vicerrectoría de Asuntos Académicos de la Universidad de Chile. Además, el procesamiento y

desarrollo de algoritmos fueron realizados en Microsoft Excel y RStudio.

Se decidió acotar el estudio a los alumnos de la Facultad de Ciencias Físicas y Matemáticas (FCFM) y la Facultad de Derecho (se le tratará indistintamente como carrera de Derecho). La FCFM presenta una de las mallas más flexibles de la Universidad, por lo que los modelos aplicados en esta facultad serán fácilmente adaptables a otras, además el tesista maneja sus procesos. La carrera de Derecho fue elegida porque es uno de los programas en donde los alumnos demoran más tiempo en titularse y parece relevante estudiar cómo esto se condice con el pago de aranceles.

La carrera de derecho tiene una duración de 5 años y los programas dictados en la FCFM duran 6 años o menos. De esta forma, parece razonable seleccionar como tope a los alumnos que ingresaron el año 2011. Porque mientras más antiguos los datos, presentan más inconsistencia, pero una fracción importante de los alumnos de la cohorte 2012 no han terminado los estudios.

Para evitar variaciones debidas a particularidades de cada año se seleccionará más de una cohorte. De esta forma se analizarán las cohortes 2009, 2010 y 2011 para calcular los parámetros que serán utilizados para simular las trayectorias de la cohorte 2016, correspondiente a la primera cohorte con gratuidad desde primer año.

2.2. Trayectoria en la FCFM y Derecho

Es importante entender el objeto a modelar, por lo que se detallarán las particularidades de los programas de ambas facultades.

2.2.1. Facultad de Derecho

Todos los años ingresan cerca de 500 estudiantes¹ a la carrera de Derecho cuyo nombre formal es Licenciatura en Ciencias Jurídicas y Sociales. Cerca del 50% corresponden a mujeres y el 50% a hombres. La carrera dura 5 años y no existe elección de especialidades intermedias. Sin embargo, el proceso de titulación es bastante complejo y debe ser entendido para saber los hitos que marcan el fin del deber de pago de aranceles a la universidad.

El plazo máximo para egresar es de 8 años. Para hacerlo se debe cumplir con la aprobación de una cantidad de créditos establecidos en el reglamento de la carrera. El título de licenciado (título entregado por la facultad), se obtiene tras aprobar el examen de licenciatura (conocido como examen de grado).

El título de abogado es entregado por la Corte Suprema de Justicia y requiere cumplir todos los requerimientos de la universidad, además de la realización de una práctica profesional, que para desarrollarla exige que los estudiantes estén egresados o cuenten con el título de licenciatura.

¹Desde el año 2018, antes entraban en torno a 400.

Los estudiantes están obligados a pagar aranceles sólo hasta el egreso, por lo que el período necesario hasta la obtención de la licenciatura o del título no será considerado en este trabajo.

2.2.2. FCFM

Anualmente ingresan cerca de 800 estudiantes de Pregrado, de los cuales cerca del 70 % son hombres y 30 % mujeres. El 87% de ellos accede a la universidad por la vía de ingreso regular mediante la ponderación de la PSU (Prueba de Selección Universitaria) y rendimiento escolar en la enseñanza media, además existen vías de acceso especiales, entre las que se encuentra el acceso SIPEE, PACE, cupos para deportistas, extranjeros, indígenas y equidad de género. También, pueden acceder estudiantes provenientes del programa de Bachillerato.

Los estudiantes entran a un Plan Común que tiene dos años de duración formal, tras lo cual pueden optar por seguir cualquier carrera de la Facultad, correspondientes a tres Licenciaturas terminales de 4 años de duración, dos carreras profesionales de 5,5 años de duración y nueve carreras profesionales de 6 años de duración.

Los primeros dos años de Plan Común tienen los mismos requerimientos para todas las carreras sin importar la continuidad que se siga. No obstante, después del plan común cada carrera tiene sus propias exigencias necesarias para completar cada programa.

Los ramos son semestrales, con diferentes formas de evaluación. Las mallas son muy flexibles, permitiendo a los estudiantes adelantar asignaturas que formalmente deberían tomar en años futuros y también tienen libertad de tomar ramos de cualquier especialidad mientras cumplan con los requisitos mínimos.

Según el reglamento de la Facultad, la permanencia máxima como alumno activo dentro de una carrera corresponde a 10 años, sin embargo pueden haber excepciones. Los periodos de postergación académica (inactividad), no son considerados en el cálculo de este plazo.

Capítulo 3

Análisis Exploratorio de los datos

Las cohortes utilizadas para ajustar los modelos fueron las 2009, 2010 y 2011, con 2189 trayectorias diferentes para la FCFM y 1139 para Derecho. Se eliminaron los registros de cerca de 60 individuos de cada facultad por falta de información (principalmente de ingresos especiales).

3.1. Descripción de las variables

3.1.1. Variable dependiente - Estado por semestre

Esta variable se generó a partir de la «Situación Académica» de los estudiantes almacenada en sistema. Dentro de los estados de Situación Académica recogidos por el sistema informático, podemos encontrar 27 categorías diferentes que fueron agrupados en 4 grandes estados:

- Regular/Activos (A): Corresponde a los estudiantes matriculados, que aún no han egresado.
- Inactivo (C): Corresponde a los alumnos que no se matricularon, pero no han terminado su vínculo de estudiantes con la universidad, ni tampoco han egresado.
- Egresado (Eg): Corresponde a los estudiantes con el título de egresado, dado que el egreso marca el punto en que los estudiantes dejan de pagar arancel, es el hito que interesa (no la titulación).
- Eliminado (El): Corresponde a los estudiantes que han terminado su vínculo con la universidad sin haber alcanzado el título/licenciatura terminal de su carrera, no se distingue el motivo de la eliminación.

Por cada cambio en el estado de los alumnos se genera un registro en la base de datos institucional; la mayoría de los estudiantes cambia de estado a lo más una vez al año, sin embargo existen estudiantes con varios cambios de estado en un año como producto de transferencias de carrera,

Estados Anuales	Cantidad de casos
1	21700
2	8755
3	2993
4 o más	300

Cuadro 3.1: Cantidad de Estados Anuales por Estudiante

postergaciones y eliminaciones ocurridas en dicho periodo de tiempo. Los alumnos que continúan sus estudios en programas de postgrado con doble titulación, son considerados como egresados.

De esta forma los alumnos sólo pueden aparecer como matriculados, eliminados, egresados. Una vez alcanzan un estado terminal, dejan de tener registros. Los estados egresados y eliminados, se asignan al semestre posterior al último semestre que el estudiante tuvo que pagar (según el reglamento de la universidad). Por ejemplo, si un estudiante elimina a comienzos del semestre (31 de marzo para los semestre de otoño y 31 de julio para semestre de primavera), entonces el registro de eliminación será en el semestre en curso, si elimina una vez pasada esa fecha de corte, entonces su eliminación será asignada al siguiente semestre. Esta variable se analizará en mayor detalle con la evaluación del modelo no paramétrico.

Para el caso de la FCFM, además de guardar el estado de la matrícula, se almacena la carrera en que está inscrito cada estudiante. En este proyecto, sólo se tratarán los casos con un cambio de carrera, además los cambios a programas de otras facultades se consideran como eliminados. En el set de datos se registran los siguientes cambios de carreras:

Programa	Número de Estudiantes
GEOLOGÍA	192
INGENIERÍA CIVIL	264
INGENIERÍA CIVIL ELÉCTRICA	194
INGENIERÍA CIVIL INDUSTRIAL	423
INGENIERÍA CIVIL MATEMÁTICA	80
INGENIERÍA CIVIL MECÁNICA	154
INGENIERÍA CIVIL DE MINAS	182
INGENIERÍA CIVIL QUÍMICA	75
INGENIERÍA CIVIL EN BIOTECNOLOGÍA	38
LICENCIATURA EN CIENCIAS	67
INGENIERÍA CIVIL EN COMPUTACIÓN	130
Total	1799

Cuadro 3.2: Cantidad de Alumnos de las cohortes 2009-2011 que continuaron a cada carrera de la FCFM.

Las ingenierías en Biotecnología y Química se consideraron como una carrera, debido al parecido de sus mallas y para no tener tan pocas observaciones en cada una.

Por último se revisó la agrupación temporal de las ocurrencias de las eliminaciones y egresos, pues esto influye en el modelo que se debe elegir para estudiar la duración, los resultados se observan en la siguiente figura:

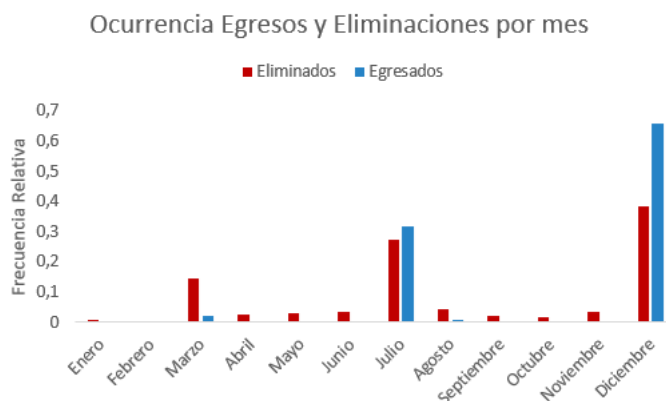


Figura 3.1: Distribución Temporal de Eliminaciones y Egresos por Mes

Se observa que ambas se concentran casi totalmente en los meses finales de cada semestre, especialmente los egresos. Se considera que esto no refleja de manera fidedigna la realidad, pues los egresos también ocurren a lo largo de los semestres, sin embargo esta es la fecha en que se registran en sistema.

3.1.2. Variables de caracterización de los estudiantes

Con respecto a las variables predictoras, se considerará la región de origen (Como variable proxy se utilizó la región del colegio de egreso de enseñanza media, pues la información sobre la dirección de los estudiantes es inconsistente), el sexo y el Índice de Vulnerabilidad Escolar (IVE) del colegio de egreso de los estudiantes. El IVE es un indicador del nivel de vulnerabilidad de cada colegio que se calcula como la suma de los alumnos en primera, segunda y tercera prioridad (estas prioridades corresponden a alumnos vulnerables), partido la matrícula total. Los colegios particulares no tienen IVE. Para las cohortes 2009, 2010 y 2011, este indicador fue calculado extrayendo el IVE del colegio en un año conocido.

Según la literatura analizada, las tres variables pueden tener poder predictivo; los estudiantes de regiones suelen comportarse de forma diferente a los alumnos de Santiago, esto hace sentido debido al control que pueden ejercer las familias sobre la vida de los alumnos. Además, en general es más caro desde el punto de vista económico y del bienestar para los estudiantes de regiones.

El IVE será el indicador utilizado como proxy del nivel socioeconómico de los estudiantes. Además, se utilizará el puntaje ponderado en la Prueba de Selección Universitaria y la vía de

acceso.

Es importante mencionar que se acotó el número de variables socioeconómicas a estudiar, debido a que se utilizarán como predictoras las variables de rendimiento hasta tercer año, que se espera tengan un poder predictivo mayor.

		IVE A	IVE B	IVE C	Sin IVE	Total
FCFM	Cantidad	37	207	724	1221	2189
	Promedios PSU	737	742	748	749	748
Derecho	Cantidad	37	117	357	628	1139
	Promedios PSU	713	717	720	725	722

Cuadro 3.3: Promedio PSU y cantidad alumnos por IVE.

Como es de esperar, los estudiantes de los colegios más vulnerables promedian peores puntajes PSU (inclusive en este caso, que existe un sesgo de selección pues estamos promediando entre los estudiantes que lograron ingresar a estas carreras). Resalta la baja proporción de estudiantes de los grupo A y B, lo que demuestra que por más que se considere que la Universidad de Chile es un espacio republicano con representación de todos los sectores socioeconómicos, sigue siendo un espacio muy elitista (al menos la FCFM y Facultad de Derecho).

En caso de agregar el IVE como variable predictora categórica, los grupos A y B serán tratados como uno.

En la siguiente tabla se observa la proporción de los estudiantes de la Región Metropolitana y regiones por facultad. Se observa que en ambas facultades la proporción de estudiantes de Santiago

Facultad	Metropolitana	Otras
FCFM	72,8 %	27,2 %
Derecho	74,2 %	25,8 %

Cuadro 3.4: Región de origen de los alumnos, cohortes 2009, 2010 y 2011.

casi alcanza los 3/4 del total los casos.

En la siguiente tabla se muestra la proporción de alumnos con gratuidad por año. Esta variable no será utilizada como predictora, ya que no aplica para las cohortes 2009-2011. Sin embargo, es central en este trabajo y será usada para la valorización de la simulación.

Se observa que, en torno a un tercio de los estudiantes de ambas facultades estudian con gratuidad.

		2016	2017	2018
FCFM	Gratuidad	0,33	0,30	0,31
	Sin Gratuidad	0,67	0,70	0,69
Derecho	Gratuidad	0,35	0,30	0,34
	Sin Gratuidad	0,65	0,70	0,66

Cuadro 3.5: Alumnos con gratuidad.

3.1.3. Variables de rendimiento académico

Se cuenta con la proporción de ramos aprobados, inscritos y sus promedios para cada año. Dado que el objetivo es simular a partir del tercer año, sólo se utilizará la información hasta dicho periodo. En la siguiente tabla se muestra el promedio y desviación estándar de dos de estas variable, clasificadas por sexo y facultad.

Facultad	Sexo	Promedio Aprobados	Promedio Reprobados	$\sigma_{Aprobados}$
Ingeniería	Mujeres	10,2	0,9	3,2
	Hombres	9,8	1,1	3,4
Derecho	Mujeres	11,6	0,8	3,1
	Hombres	11,2	1,0	3,4

Cuadro 3.6: Ramos aprobados y reprobados al año por sexo y facultad - Cohortes 2009, 2010 y 2011.

Para confeccionar esta tabla se eliminaron los años finales, en donde en general se inscriben menos ramos. Se aprecia una considerable diferencia de rendimiento a favor de las estudiantes mujeres, también se observa que la cantidad de ramos aprobados tiene una alta dispersión en ambas facultades, tanto para hombres como para mujeres.

Capítulo 4

Modelo Matemático de la Trayectoria Estudiantil

Como la variable de interés corresponde al tiempo en que los estudiantes están matriculados en la universidad hasta el egreso o eliminación, es necesario conocer el momento de entrada del estudiante y el de graduación o de abandono. También se deben analizar los periodos de postergación, ya que estos representan una fracción considerable del tiempo en que los estudiantes están vinculados con la universidad, y en ellos los alumnos no deben pagar arancel.

Dado que el objetivo es obtener las duraciones por medio de simulaciones, el foco estará puesto en el cálculo de las probabilidades de transición entre estados para los distintos periodos y no en la obtención analítica de dicha duración. No obstante, se estudió literatura de los modelos de supervivencia (también llamados de duración o riesgo), para estudiar distintas formas de calcular las probabilidades de transición (tasas de riesgo).

El tiempo es una variable continua, pero dependiendo de las aplicaciones o el tipo de datos disponible se puede discretizar. En este trabajo se optó por medir la duración de forma discreta, como el número de semestres en que los estudiantes están matriculados. Esto, debido a las condiciones de pago de la matrícula y aranceles (importan los estados intermedios de la trayectoria) y a que gran parte de los eventos se concentran en períodos particulares del año (ver 3.1), y en la literatura recomiendan que cuando existen estas concentraciones es mejor proceder con métodos discretos ([Singer and Willett \[1993\]](#), [Scott and Kennedy \[2005\]](#), [Rodríguez \[2007\]](#)), por lo que se exploró literatura con dicha especificación.

4.1. Revisión de literatura

Se estudiaron aplicaciones en el ámbito educativo que modelan la trayectoria como procesos de duración discretos, en donde los estudiantes se ven a expuestos a uno o varios riesgos. El objetivo de este análisis, es obtener los estados o riesgos típicamente modelados y las formas en que se calculan

las probabilidades de transición. Es importante notar que, dada la simulación que se realizará en este trabajo, resultados de gran interés como la duración o tiempo hasta llegar a los estados terminales no serán calculados de forma analítica.

Los modelos de duración están especialmente diseñados para analizar el tiempo transcurrido hasta la ocurrencia de un evento. Fundamentalmente, estos procesos se destacan porque entregan herramientas para tratar la censura (observaciones que no han vivido el evento en el periodo de evaluación, pero quizás sí lo hagan en el futuro) y permiten agregar covariables y efectos que cambien en el tiempo [Rodríguez \[2007\]](#), [Bhat \[2007\]](#). A continuación, se analizarán una serie de trabajos que han aplicado modelos de duración en tiempo discreto al ámbito de la educación.

Uno de los primeros artículos en esta área es el de [Singer and Willett \[1993\]](#) en donde utilizan un modelo logit para medir la probabilidad de que los profesores dejen la enseñanza. Los profesores sólo se enfrentan a un «riesgo»(dejar de enseñar), que además ocurre sólo una vez. En [Singer and Willett \[1995\]](#) se presenta una generalización del problema anterior, al considerar la recurrencia de eventos. En cada fase los profesores se siguen enfrentando a un sólo riesgo y modelan el mismo riesgo que en [Singer and Willett \[1993\]](#), pero una vez ocurre el evento, modelan la probabilidad de que vuelvan a hacer clases.

[DesJardins et al. \[1999\]](#) son de los primeros autores en utilizar estos modelos para modelar las trayectorias estudiantiles, este trabajo continúa con [DesJardins et al. \[2002a\]](#), [DesJardins et al. \[2002b\]](#) y [DesJardins et al. \[2006\]](#), dónde utilizan el enfoque más completo revisado, ya que consideran que los estudiantes se enfrentan a diferentes riesgos; matrícula regular, congelación, abandono y graduación. Además, por modelar el proceso de congelación («stop out»), deben utilizar técnicas para modelar la recurrencia en los riesgos. Además, agregan parámetros para considerar heterogeneidad no observable.

[Calcagno et al. \[2007\]](#) lo utilizan para ver el efecto en la trayectoria de estudiantes de Community College el lograr ciertos hitos en la carrera (por ejemplo, completar el 50% de los créditos) y para medir la diferencia en los impactos según el grupo etario de los estudiantes. Este trabajo tiene bastante similitud con lo que se hará en este proyecto, en el sentido de que se usarán las variables de rendimiento hasta cierto momento para predecir la duración de las carreras, sin embargo, en [Calcagno et al. \[2007\]](#) consideran a los estudiantes eliminados como censurados. En esta tesis sí se modelara dicha transición.

[Bowers \[2010\]](#) Implementa un modelo de duración para calcular el riesgo de abandono de estudiantes de secundaria («highschool»), en donde los estudiante se enfrentan a un sólo riesgo correspondiente al abandono, los estudiantes que se gradúan o transfieren a otros establecimientos

se consideran como censurados desde el momento en que eso sucede. Al igual que en la mayoría de los otros trabajos, para calcular las tasas de riesgo utilizó la regresión logística. Las variables temporales, entran como variables dummies. No obstante, también ajusta un modelo en donde no entran variables temporales, es decir, la tasa de riesgo es constante en el tiempo y depende del promedio de notas de los estudiantes (GPA). Se menciona este caso porque los resultados obtenidos son aceptables, considerando lo sencillo del modelo.

Se destacan los artículos de [Scott and Kennedy \[2005\]](#) y [Glocker \[2011\]](#), pues son los artículos externos que funcionaron como base para este trabajo por su parecido. Los dos artículos proponen un enfoque de riesgos en competencia para medir los resultados estudiantiles. El foco del primero es medir el impacto de ciertas conductas como trabajar o postergar, mientras que en el segundo se desea medir diferencias producto de beneficios estudiantiles. Los alumnos se enfrentan a las mismas alternativas que se modelaron en este trabajo; egresar y eliminar (en [Scott and Kennedy \[2005\]](#) se modela además un tercer riesgo correspondiente a la posibilidad de transferencia a otro programa). Ambos modelan las probabilidades con modelos logit multinomiales. Además, el primero presenta una metodología de análisis exploratorio que fue desarrollada en esta tesis y en el segundo se agrega heterogeneidad no observada de la misma forma que en este trabajo para la Facultad de Derecho. La principal diferencia de este trabajo con estos dos artículos, es que la duración no se calculará analíticamente, sino que por medio de simulaciones.

Además, se estudiaron publicaciones académicas y tesis de estudiantes de la Universidad de Chile (en particular de la FCFM) para encontrar qué variables son las de mayor valor predictivo sobre el rendimiento de los alumnos¹.

En primer lugar, se menciona el trabajo de [Celis \[2012\]](#), un trabajo cuyo objetivo y metodología es muy similar a lo pretendido en esta investigación, por lo que sirvió como punto de comparación y validación, para hipótesis encontradas.

El autor ajusta dos modelos, el primero para obtener el tiempo necesario hasta la primera “congelación” y el segundo para calcular el tiempo necesario hasta la graduación. Para el primer modelo utiliza un logit simple y un modelo log-log, optando por este último por entregar resultados más consistentes. Mientras que para el segundo, ocupa una regresión de Cox. Incluye heterogeneidad no observable por medio de la inclusión de parámetros con distribución normal y efecto cambiante en el tiempo para ciertas variables (mediante la interacción con una dummy temporal). Sin embargo, al igual que en este trabajo, no incluye variables cambiantes en el tiempo. Las principales diferencias, se dan porque en el presente trabajo se modelan las probabilidades para tres eventos al mismo

¹Es importante recalcar que las variables para predecir el rendimiento de los estudiantes en distintas universidades puede ser diferente.

tiempo (continuar como alumno regular, abandonar y egresar), mientras que los modelos de [Celis \[2012\]](#), sólo modela un riesgo simultaneo. No especifica como trata la censura, ni a los alumnos que sufren otro evento y que por lo tanto no podrán sufrir el evento modelado.

En [Celis et al. \[2015\]](#) aplican una técnica de minería de datos para estimar la probabilidad de doble reprobación de los alumnos de primer año de la FCFM, con variables socioeconómicas y la información académica de hasta el primer control del segundo semestre. Seleccionan las variables con distintos métodos y aplican un modelo de regresión logística. Las variables con mayor poder predictivo resultaron ser las del rendimiento mostrado por los alumnos, mientras más «cercana» temporalmente la variable, mejor la precisión. De esta forma, la de mayor valor fue la nota en el control 1 del segundo semestre.

El trabajo de [Aguirre \[2012\]](#) busca medir la capacidad predictora de la Prueba de Selección Universitaria en el rendimiento de los alumnos de la FCFM, mediante modelos de mínimos cuadrados ordinarios (MCO), encontrando que el mejor predictor es el puntaje en la prueba de ciencias.

El trabajo de [Grau \[2015\]](#) busca medir el poder predictivo del puntaje Ranking PSU², mediante MCO y modelos de duración. Al igual que en el trabajo anterior, la variable con mayor influencia corresponde al puntaje en la PSU de ciencias (al menos para los primeros años).

El trabajo de [Aguirre \[2016\]](#) tiene mayor relación con este trabajo, pues utiliza información del desempeño en la universidad hasta el primer semestre para predecir el rendimiento del segundo semestre. En su metodología clusteriza por rendimiento con el algoritmo KMedias y luego encuentra las variables socioeconómicas que están relacionadas con cada cluster (siendo la variable región la única con diferencia significativa entre las distintas agrupaciones).

4.2. Modelamiento

En primer lugar, se indica que en el anexo C, se presentan brevemente los herramientas utilizadas para lectores ajenos a las metodologías de estimación, con el fin de facilitar la lectura y entregar las bases conceptuales.

A continuación se detallarán las transiciones que se modelarán en este trabajo. Se tratará la omisión de un proceso considerable en la trayectoria de una gran cantidad de los estudiantes, correspondiente a los procesos de inactividad o postergación.

4.2.1. Postergaciones

En este trabajo no se modelarán los periodos de postergación, debido a que se considera que proyectar estos periodos con información de hasta tercer año tendría demasiadas fuentes de error

²Puntaje asociado al desempeño en enseñanza media de los alumnos en relación a sus compañeros de colegio.

(no hay que perder de vista que sólo se utilizará información hasta tercer año porque el objetivo es simular la trayectoria restante de los estudiantes de la cohorte 2016), además en dichos periodos los estudiantes no tienen una obligación financiera con la universidad.

Modelar este proceso es complejo y requiere de muchas observaciones para obtener estimaciones sin simplificarlo demasiado³, considerando que factores como su duración, el momento en que esta ocurre y la cantidad de postergaciones que ha tenido el alumno influyen en las probabilidades de transición (ver tablas 4.1 y 4.2). Por ejemplo, la probabilidad de volver a ser alumno regular después de congelar en primero o en sexto es diferente, lo mismo ocurre para las probabilidades de volver a ser alumno regular después de llevar un semestre postergado versus diez.

En la tabla 4.1 se muestra la proporción de duraciones de 1 año o menos con respecto al total de postergaciones según el momento de la carrera en que ocurre. Se observa que a medida que la postergación ocurre en etapas más tardías, la proporción de postergaciones cortas es mayor, sin embargo este efecto se revierte desde el sexto año en adelante, debido a que a partir de ese año comienzan las postergaciones largas en ingeniería causadas por la dificultad de egresar.

Años desde Ingreso	Postergaciones	Menores a un año
1	95	38,9 %
2	90	48,9 %
3	97	61,9 %
4	106	72,6 %
5	118	65,3 %
6	147	47,6 %
7 o más	362	25 %

Cuadro 4.1: Alumnos Postergados de las cohortes 2009, 2010 y 2011, por año de postergación.

En la tabla 4.2 se muestra la duración del proceso de postergación de los alumnos reincorporados y la cantidad de alumnos que hasta la fecha de corte no se han reincorporado, en dónde se observa que mientras más años se prolonga la inactividad, menor es la proporción de estudiantes que retoman los estudios.

Se observa que muy pocos alumnos con periodos extensos de postergación retomarán sus estudios, por lo que se indica que los alumnos que en la fecha de obtención de la data (Otoño 2019) lleven más de 4 semestres no matriculados, serán considerados como eliminados. Esto se debe realizar, porque considerarlos como censurados subestimaría demasiado el riesgo de eliminación. Una regla similar es aplicada por [Scott and Kennedy \[2005\]](#).

³Se decidió no considerar más cohortes para agrandar la muestra, porque las cohortes más antiguas presentan más inconsistencias y en la FCFM hubo un importante cambio de malla el año 2008.

Tiempo Postergados (Años)	Alumnos que se reincorporaron en periodo de evaluación	Alumnos postergados a fecha de obtención data
[0;1)	646	40
[1;2)	19	88
[2;3)	4	110
[3;4)	2	83
[4;5)	1	42
[5;6)	0	22
[6;7)	0	35
[7;8)	0	40
[8;9)	0	26
[9;10)	0	21

Cuadro 4.2: Cantidad de alumnos de las cohortes 2009, 2010 y 2011 con al menos un proceso de postergación clasificados por la cantidad de años inactivos.

4.2.2. Estados, transiciones y periodos

A continuación se definirán los componentes del modelo y se muestra de forma gráfica las transiciones a modelar en este trabajo.

Estados: Regular, Eliminado, Egresado

Periodos: Los periodos corresponden a los semestres.

P_{kijt} : La probabilidad de que el alumno k se cambie del estado i al estado j en t . Se muestran dos esquemas para graficar el mismo proceso, la figura 4.1 corresponde a un diagrama más compacto, y se puede calificar como una cadena de markov discreta no homogénea, pues las probabilidades de transición dependen únicamente del último estado, pero varían según la etapa.

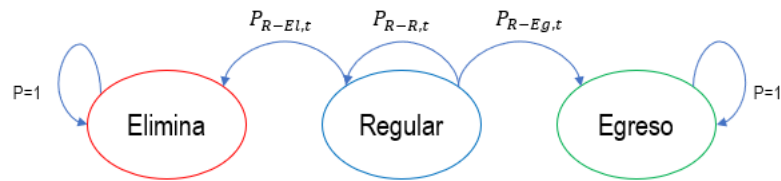


Figura 4.1: Diagrama compacto de transiciones a modelar en este trabajo

En el diagrama 4.2 se expande dicha cadena, logrando una cadena homogénea. Para esto se enumera la cantidad de etapas en el estado Regular. Además, se distinguen las probabilidades de transición que serán calculadas por los modelos Logit presentados más adelante.

Es importante recalcar que estas probabilidades se estimarán a nivel individual. Además en la Facultad de Ciencias Físicas y Matemáticas, las tasas dependerán de la carrera de cada estudiante, pero muchos de los estudiantes de la cohorte 2016 aún no eligen especialidad, por lo que la elección de carrera se modelará cómo un proceso previo. En otras palabras, se necesita simular la carrera de

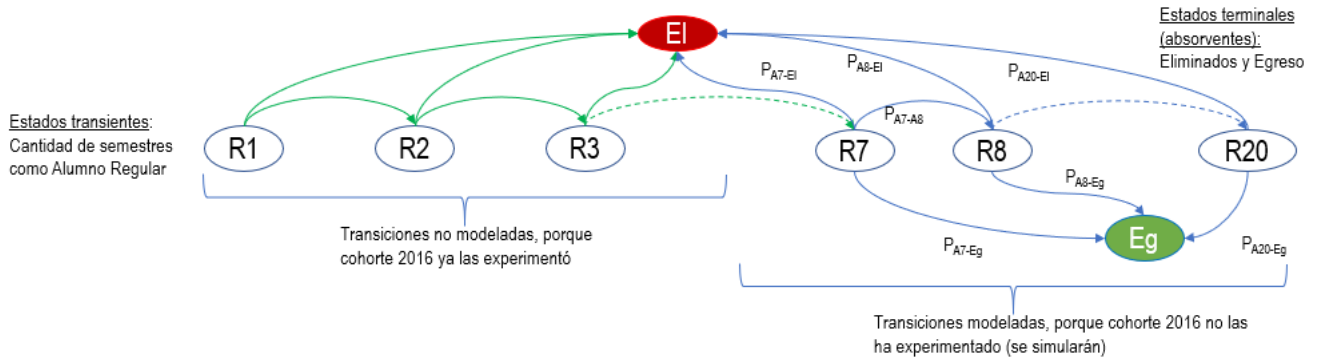


Figura 4.2: Diagrama de transiciones a modelar en este trabajo

los estudiantes que aún no han elegido su programa de continuidad, debido a que esta sirve como entrada para el cálculo de las probabilidades mostradas en los diagramas anteriores. Finalmente se indica que se optó por modelar esta elección de forma atemporal. En la imagen 4.3 se grafica la modelación de este proceso.

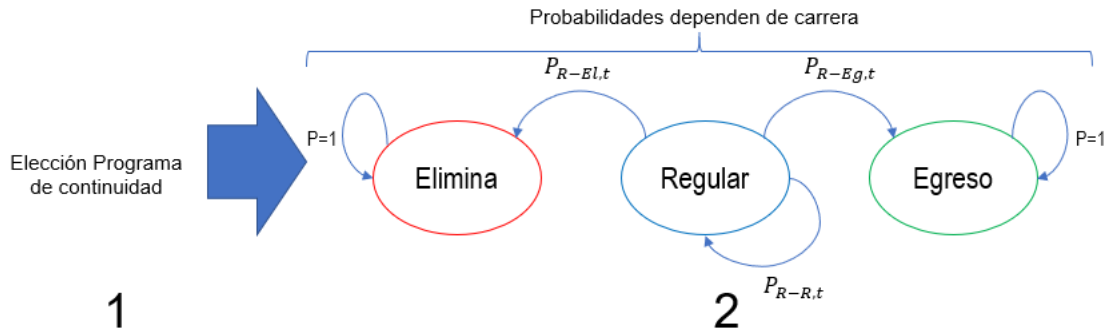


Figura 4.3: Secuencia para ingeniería.

4.2.3. Probabilidades de transición

La probabilidad de transición P_{kijt} , corresponde a la probabilidad del alumno k de pasar del estado i al estado j en el periodo t (la indexación en t sólo tiene sentido para las matrices de transición no homogéneas). Será el insumo clave de la simulación.

A esta probabilidad se le puede imponer estructura o calcular de forma no paramétrica. En la siguiente ecuación, se puede observar una formulación genérica, en donde la probabilidad de transferencia del estado i al j para el alumno k en t, puede depender del periodo, variables del alumno k constantes o dinámicas en el tiempo, heterogeneidad no observable, entre otras.

$$P_{kijt} = P(i \rightarrow j | t, x_k(t), \theta_k) \quad (4.1)$$

Con:

i, j : Estados $\in \{\text{Egresar, Eliminar, Regular}\}$.

t : Semestre.

$x_k(t)$: variables del alumno k .

θ_{kt} : parámetros de comportamiento a ser estimados, que se asumen heterogéneos en la población de estudiantes.

4.2.4. Censura

En el caso de este trabajo, existen observaciones censuradas, correspondiente a los alumnos de las cohortes 2009, 2010 y 2011, que hasta 2019 no han egresado o eliminado y no llevan más de dos años postergados (recordar que los alumnos con más de dos años de postgrados a la fecha de obtención de la información son considerados como eliminados).

Esta censura será considerada como independiente (censura no informativa). Por lo que, las observaciones censuradas serán ocupadas hasta la ocurrencia de la censura, [Scott and Kennedy \[2005\]](#) llaman a esto “ignorar la censura”.

Que la censura sea no informativa o independiente, puede ser entendido como que los individuos censurados presentan la misma probabilidad de tener un evento que los no censurados. Matemáticamente, esto se cumple si la siguiente igualdad es cierta:

$$\frac{P(\text{Evento en } t)}{P(\text{Sin evento antes de } t)} = \frac{P(\text{Evento en } t \text{ y no censurado en } t)}{P(\text{sin evento hasta } t \text{ y no censurado en } t)} \quad (4.2)$$

De otra manera los estudiantes censurados tendrían más o menos probabilidad de sufrir el evento y al ignorar la censura se estaría sub o sobre estimando las tasas de riesgo respectivamente.

Capítulo 5

Cálculo de probabilidades de transición

Como una fase previa al ajuste de los modelos se realizó un análisis exploratorio inspirado en la estrategia presentada por [Scott and Kennedy \[2005\]](#), en donde se calcula el estimador no paramétrico de las probabilidades de transición para las observaciones agrupadas en distintas variables candidatas a ser variables predictoras (se calcularían sólo de este modo si no interesara calcular las probabilidades a nivel individual), un enfoque similar es usado por [Marcel Goic \[2019\]](#).

5.1. Análisis exploratorio con modelos no paramétricos

El objetivo de este análisis es encontrar gráficamente las variables que hacen cambiar el valor del estimador. Además, la temporalidad de estas diferencias, para ver de qué modo deberían entrar en el modelo. Las observaciones son agrupadas por los valores de la variable explicativa que se analizará, por ejemplo, si se desea estudiar las diferencias entre los sexos, se agrupan las observaciones femeninas y masculinas para calcular el estimador de forma independiente. En caso de que la variable no sea categórica, se debe agrupar en intervalos contruídos a juicio del investigador.

Sea N_t el número de estudiantes activos en t y N_{kt} el número de estudiantes que sufren evento k en t , entonces el estimador máximo verosímil no paramétrico de la tasa de riesgo para el evento k en t , se calcula como ([Scott and Kennedy \[2005\]](#)):

$$\hat{P}_{kt} = \frac{N_{\text{Evento } k \text{ en } t}}{N_t} \quad (5.1)$$

Este estimador es fácil de obtener, es muy intuitivo y servirá como base para la comparación de los modelos más complejos. En las tablas 5.1 y 5.2 se observan las «tablas de vida» para la FCFM y Facultad de Derecho respectivamente que muestran la cantidad de alumnos activos, eliminaciones, egresos, observaciones censuradas y las tasas de abandono y egreso por semestre. Por último, en la figura 5.1 se grafican dichas tasas.

Como se mencionó en la sección 3.1.1, la variable dependiente toma el valor de Egresado y Eliminado el primer semestre en que el estudiante no tiene obligación financiera con la universidad. Al observar la siguiente tabla, los 12 estudiante que eliminaron durante el primer semestre, corresponden a alumnos que eliminaron las primeras semanas del semestre (actualmente el reglamento establece que los alumnos que eliminan hasta el 31 de marzo están excentos del pago anual). Así, los estudiantes que eliminaron posterior a la quinta semana, son consideradas como eliminaciones del segundo semestre.

Semestre	Total	Regular	Eliminado	Egresado	Censurados	\widehat{P}_{eg}	\widehat{P}_{el}
1	2254	2242	12	0	0	0,00	0,01
2	2242	2158	84	0	0	0,00	0,04
3	2158	2076	82	0	0	0,00	0,04
4	2076	2034	42	0	0	0,00	0,02
5	2034	1989	45	0	0	0,00	0,02
6	1989	1966	23	0	0	0,00	0,01
7	1966	1936	30	0	0	0,00	0,02
8	1936	1908	28	0	0	0,00	0,01
9	1905	1866	33	6	3	0,00	0,02
10	1866	1831	34	1	0	0,00	0,02
11	1829	1759	54	16	2	0,01	0,03
12	1746	1660	54	32	13	0,02	0,03
13	1607	1188	194	225	53	0,14	0,12
14	1158	984	31	143	30	0,12	0,03
15	877	529	131	217	107	0,25	0,15
16	484	389	10	85	45	0,18	0,02
17	224	142	18	64	165	0,29	0,08
18	118	87	2	29	24	0,25	0,02
19	35	17	0	18	52	0,51	0,00
20	9	5	0	4	8	0,44	0,00

Cuadro 5.1: Tabla de vida - FCFM.

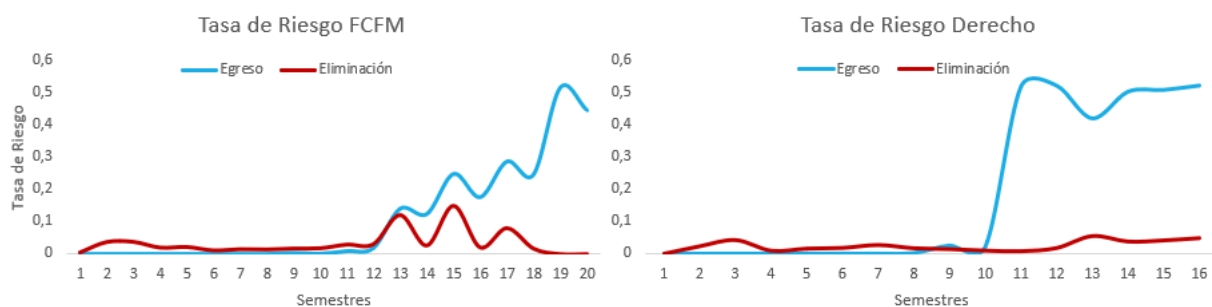


Figura 5.1: Probabilidades de transición FCFM y Facultad de Derecho

Como era de esperar, los egresos se concentran hacia el decimotercer-decimoquinto semestre

Semestre	Total	Regular	Eliminado	Egresado	Censurados	\widehat{P}_{eg}	\widehat{P}_{el}
1	1199	1198	1	0	0	0,00	0,00
2	1198	1170	28	0	0	0,00	0,02
3	1170	1121	49	0	0	0,00	0,04
4	1121	1110	11	0	0	0,00	0,01
5	1110	1092	18	0	0	0,00	0,02
6	1092	1072	20	0	0	0,00	0,02
7	1072	1043	29	0	0	0,00	0,03
8	1042	1022	18	2	1	0,00	0,02
9	1022	981	15	26	0	0,03	0,01
10	981	950	10	21	0	0,02	0,01
11	949	447	8	494	1	0,52	0,01
12	447	205	8	234	0	0,52	0,02
13	204	107	11	86	1	0,42	0,05
14	107	49	4	54	0	0,50	0,04
15	49	22	2	25	0	0,51	0,04
16	21	9	1	11	1	0,52	0,05

Cuadro 5.2: Tabla de vida - Facultad de Derecho.

para la FCFM y del undécimo en adelante para Derecho. Esta tasa de riesgo es ascendente (con pequeñas oscilaciones) durante todo el periodo de observación. Esto representa el hecho de que la mayor proporción de los estudiantes egresa en la duración formal de los programas más uno o dos años, pero tras cada año es más probable que una mayor proporción de los estudiantes que aún no han egresado, egresen.

De esta manera, el mayor probabilidad ocurre en los semestre 17 y 19 para la FCFM y desde el 14 para Derecho (hacia los últimos semestres estos estimadores pierden significancia porque quedan muy pocas observaciones).

Los programas de ambas facultades tienen problemas con la titulación oportuna ¹, Pero esto no necesariamente se condice con el egreso. Como se puede observar, la tasa de egreso en derecho es más alta que en la FCFM, ya que el problema con la titulación en esa facultad ocurre post egreso (examen de grado), en cambio en la FCFM, egresar prácticamente equivale a titularse.

Se puede notar que en la FCFM a partir de los semestres 9 y 10 hay estudiantes que egresan. Los primeros en hacerlo corresponden a los estudiantes de las licenciaturas que tienen una duración formal de 8 semestres.

Con respecto al proceso de Eliminación, es más probable que los alumnos abandonen en una etapa temprana de los estudios (principalmente primer y segundo año). Sin embargo, en la FCFM los valores más altos se dan en los años posteriores a la duración formal, esto se puede deber al problema que tienen los estudiantes con el trabajo de título. Casi la totalidad de dichos «Eliminados»

¹Indicador que mide la proporción de estudiantes que se titulan en la duración formal+1 año del programa.

corresponde a estudiantes que a la fecha de corte de la información llevaban más de dos años inactivos (no eliminaron formalmente). Existe la probabilidad de que esos estudiantes vuelvan a terminar sus estudios, pero como se comentó en una sección anterior, considerarlos como censurados disminuiría demasiado el riesgo de eliminación y se presume que la mayoría de dichos estudiantes no volverá. Además, un porcentaje no menor corresponde a estudiantes cursando un programa de articulación de magíster, y no deberán pagar otro semestre de pregrado para egresar.

Se debe tener en cuenta que en este trabajo no se hace distinción entre los distintos tipos de deserción que pueden existir, por ejemplo deserciones vocacionales, expulsiones/deserciones por rendimiento, etc.

5.1.1. Segmentación por variables socioeconómicas

Región de origen

A continuación se muestran las diferencias en las tasas de egreso y eliminación para los estudiantes de regiones vs los de la Región Metropolitana.

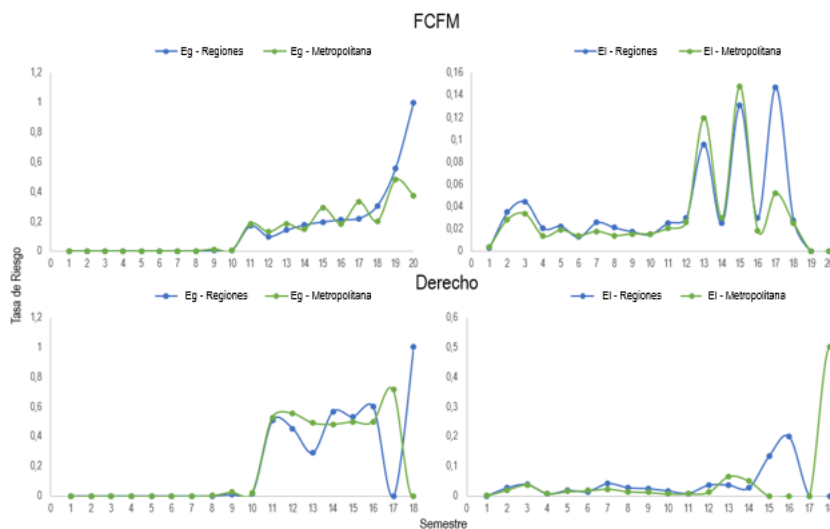


Figura 5.2: Tasa de Riesgo de Egreso y Eliminación de los estudiantes de regiones vs de la región metropolitana

Si bien no hay una diferencia clara, los estudiantes de la Región Metropolitana parecen tener mayor probabilidad de egreso y menor de eliminación en la mayoría de los periodos. Esto es consistente con lo encontrado en otros trabajos de modelos de predicción de rendimiento aplicados en la FCFM mencionados en la revisión de literatura.

5.1.2. IVE Colegio de egreso

En el siguiente gráfico se muestran las tasas de egreso y eliminación para los estudiantes de cada grupo IVE. Para el análisis se unieron los grupos A y B se consideraron como uno debido a que existen muy pocos estudiantes en esos grupos (incluso con la unión parece ser un número insuficiente de observaciones).

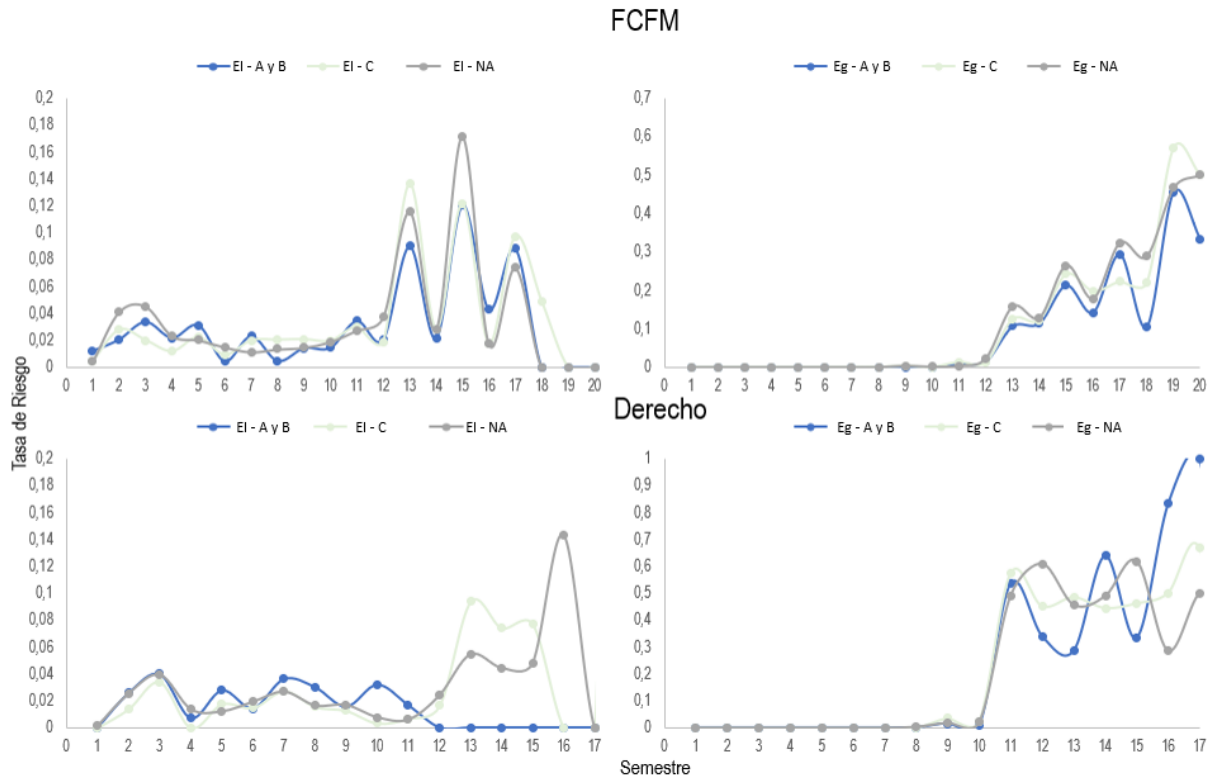


Figura 5.3: Tasa de Riesgo de Egreso y Eliminación de los estudiantes por grupo IVE de establecimiento de egreso

En el gráfico no se aprecian tendencias clara, aunque pareciera que los alumnos de los colegios sin IVE suelen graduarse antes que los de los otros grupos. Además, estos mismos presentan una pequeña diferencia positiva en la eliminación para los primeros semestres. No obstante, el grupo que parece más vulnerable a la eliminación son los alumnos del grupo IVE C.

5.1.3. Clusterización de rendimientos

Cómo se comentó previamente, el análisis de esta sección requiere segmentar las variables continuas, para poder mostrar las curvas de cada segmento de la población. Como las variables de rendimiento son continuas, se optó por agrupar con el algoritmo KMedias, en vez de usar el juicio del investigador.

La clusterización se realizó con las observaciones de cada facultad por separado. Además se consideraron las siguientes variables de rendimiento de los tres primeros años: Número de ramos aprobados de cada año, porcentaje de aprobación de cada año, promedio de cada año. Para la elección de la cantidad de clusters se utilizó la regla del codo con la suma de la varianza total dentro de los grupos, siendo 3 clusters la opción escogida para ambas facultades (el gráfico se puede ver en las imágenes 8.2 y 8.3 del anexo).

5.1.4. Estimación no paramétrica para los Clusters de rendimiento

Las tasas de riesgo se pueden apreciar en el siguiente gráfico.

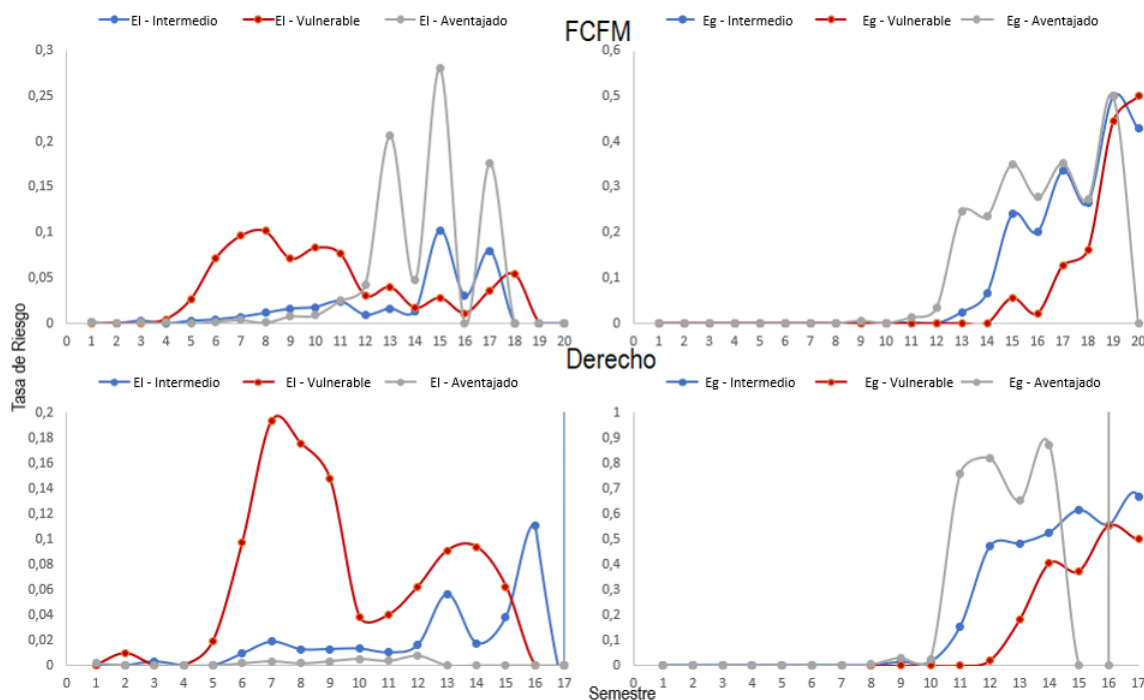


Figura 5.4: Tasa de Riesgo de Egreso y Eliminación de los distintos Clusters

Se observa una gran diferencia en las tasas de cada cluster, los alumnos del cluster «Gris» fueron bautizados como Aventajados. Estos alumnos casi no tienen riesgos de eliminación y se gradúan antes que los otros grupos. Sin embargo, hay un fenómeno extraordinario en la FCFM, en donde después del doceavo semestre la tasa de eliminación explota incluso superando las demás, esto según Celis [2012], se debe a problemas con el trabajo de título, posibles ramos críticos y al alto costo oportunidad, dado que gran parte de las empresas están dispuestas a contratar a alumnos sin que estén graduados.

El cluster «Rojo» corresponde a los alumnos más vulnerables, son los que corren más riesgo de eliminar y su tasa de graduación es menor. Finalmente los intermedios («Gris»), presentan una

tasa intermedia de eliminación y una alta tasa de titulación desde el semestre 14 en adelante (un poco más tardía en comparación con los Aventajados).

5.2. Estimación de probabilidades de transición con modelos Logit

Para la estimación de los modelos se usaron variables predictoras de hasta el tercer año y se predijo sólo desde el cuarto año, dado que el objetivo es predecir las trayectorias de los estudiantes que entraron el 2016, que actualmente cursan dicho año.

De esta forma, para el estudiante i la utilidad de la alternativa $j \in \{\text{Regular, Eliminar, Egresar}\}$ en t , queda dada por la siguiente fórmula.

$$U_{ijt} = \alpha_j + \beta_{tj} + \gamma_{tj}X_i \quad (5.2)$$

Con:

α_j : Un intercepto para cada alternativa.

β_{tj} : Utilidad Alternativa j en periodo t (primer periodo normalizado).

γ_{tj} : Efecto de covariables dependiente de t .

X_i : Variables propias de alumno i , constantes en t .

La alternativa de ser estudiante regular se utilizó como base (utilidad 0). No se utilizarán variables que cambien en el tiempo, porque el objetivo es simular con la información disponible hasta tercer año. Sin embargo, sí se explorarán efectos variables en el tiempo (γ_{jt} va indexado en t), pues es esperable que el efecto de ciertas variables sea diferente en distintas etapas de la carrera. Por ejemplo, se espera que el rendimiento previo sea un excelente predictor de los rendimientos futuros en periodos consecutivos, pero también es esperable que esta correlación vaya desapareciendo en el tiempo. Además, se indica que no se incluirá la variable “Gratuidad” como variable explicativa, porque como se mencionó anteriormente, no se conoce con exactitud que alumnos de las cohortes 2009, 2010 y 2011 cumplían o cumplen las condiciones para acceder al beneficio.

Una recomendación común en los trabajos revisados, es utilizar variables dummies para la inclusión de la dimensión temporal en las especificaciones de los modelos (mientras la cantidad de datos sea suficiente y no se corra el riesgo de sobre ajustar). De este modo se da mayor libertad a la tasa de riesgo base (entendida como la tasa de riesgo común a todos los individuos por periodo) e imponer condiciones como una dependencia lineal o polinómica pueden traer consigo sesgo en los estimadores. Además en [Dolton and von der Klaauw \[1995\]](#) se demuestra que utilizando este enfoque, los problemas de no especificar heterogeneidad no observada se minimizan. Para no sobreajustar

los modelos, se decidió restringir levemente dicho efecto temporal a que sea el mismo durante todo el año (no diferenciar por semestre), además desde el séptimo año en adelante en Derecho y octavo año en adelante en Ingeniería, se utilizó el mismo estimador temporal.

Por último, el proceso de elección de carrera en ingeniería se modeló con un Logit Multinomial, donde la utilidad por cada carrera $j \in \{\text{Carreras FCFM, Sin Carrera}\}$, se modeló de la siguiente manera (se considera la alternativa “Sin Carrera”, por que no todos los estudiantes alcanzan a elegir una especialidad):

$$U_{ij} = \alpha_j + \beta_j X_i \quad (5.3)$$

Con:

U_{ij} : Utilidad alumno i por carrera j .

α_j : Intercepto carrera j .

β_j : Coeficientes alternativa j .

X_i : Variables alumno i .

El proceso de elección de carrera, no se modeló temporalmente, por lo que no se estimará el semestre en que los alumnos realizan el cambio.

5.2.1. Resultados Facultad de Derecho

En la tabla 5.3 se pueden ver los resultados de parte de los distintos modelos de trayectoria estimados a través del software estadístico R. Con:

SAP: Suma aprobados hasta tercer año.

PA3: Promedio tercer año.

Region: Dummy con valor 1 si alumno es de la región metropolitana.

Sexo: Dummy con valor 1 si alumno es hombre.

IVE-C: Dummy con valor 1 si alumno proviene de colegio con IVE grupo C.

IVE-NA: Dummy con valor 1 si alumno proviene de colegio sin clasificación IVE.

Para el criterio de selección, se decidió priorizar los modelos con parámetros significativos, también se calcularon los indicadores AIC y BIC. Se descartaron los modelos con variables sociodemográficas, porque no presentaron parámetros significativos, excepto la variable sexo.

Los mejores modelos evaluados fueron el tercero y cuarto, que ocupan como variables predictoras, aparte de las dummies temporales, la cantidad de ramos aprobados durante los primeros tres años, el promedio en tercer año y el sexo de los estudiante (sólo el cuarto modelo).

Se observa que la probabilidad de egresar en todos los modelos (excepto el 5to), aumenta explosivamente a partir de sexto año. En cuarto y quinto es muy baja (su intercepto es muy negativo),

Parámetro\Modelo	1	2	3	4	5
α_{Eg}	-6,93***	-15,8***	-15,76***	-15.60***	-3,34***
α_{El}	-3,84***	2,15***	0,94*	0.83	1,00*
<i>PERIODO</i> _{5Eg}	3,18***	3,16***	3,17***	3.10	
<i>PERIODO</i> _{6Eg}	7,05***	7,63***	7,70***	7.67***	
<i>PERIODO</i> _{7Eg}	6,88***	8,76***	8,85***	8.68***	
<i>PERIODO</i> _{5El}	-0,50*	-0,36	-0,29	-0.32	
<i>PERIODO</i> _{6El}	0,15	-0,24	-0,19	-0.25	
<i>PERIODO</i> _{7El}	1,54***	0,38	0,35	0.52	
<i>SAP</i> _{Eg}		0,23***	0,10***	0.10***	0,02**
<i>SAP</i> _{El}		-0,18***	-0,04*	-0.03*	-0,041*
<i>PA3</i> _{Eg}			0,27***	0.28***	0,07***
<i>PA3</i> _{El}			-0,32***	-0.32***	-0,33***
<i>Region</i> _{Eg}					0,02
<i>Region</i> _{El}					0,23
<i>Sexo</i> _{Eg}				-0.22*	0,06
<i>Sexo</i> _{El}				0.19	-0,15
<i>IVE - C</i> _{Eg}					-0,01
<i>IVE - C</i> _{El}					0,38
<i>IVE - NA</i> _{Eg}					-0,02
<i>IVE - NA</i> _{El}					0,22
LL	-1.942	-1.689	-1.607	-1.605	-2.861
AIC	3.900	3.399	3.254	3.253	5.753
BIC	3.953	3.465	3.335	3.347	5.860

Niveles de Significancia: '***' $\leq 0,001$, '**' $\leq 0,01$, '*' $\leq 0,05$, '.' $\leq 0,1$

Cuadro 5.3: Modelos Logit sin heterogeneidad no observable

debido a que en esos años prácticamente no existen egresos. La probabilidad de eliminar se mantiene baja y constante durante todo el periodo, aunque aumenta un poco hacia los años finales. Al agregar las variables de rendimiento, las dummies de eliminación pierden significancia, se considera que sucede debido a que son pocos los casos de eliminación previo al egreso (se debería contar con una muestra más grande).

La cantidad de ramos aprobados y el promedio anual de 3er año resultan significativos en todos los modelos, y como era de esperar, tienen signos contrarios para ambas alternativas. Esto parece ir de acuerdo a la intuición, pues mientras mejor promedio y ramos aprobados, mayor es la probabilidad de egresar, y menor es la probabilidad de eliminar. Por último, se observa que los hombres presentan una menor probabilidad de egresar (significativo al 95 %) y una mayor probabilidad de eliminar.

Además, se calcularon las probabilidades entregadas por los modelos para todas las etapas que vivió cada estudiante, se calculó un promedio por periodo y se comparó con las tasas reales, para revisar que los modelos que utilizan más heterogeneidad, sigan obteniendo resultados similares a la proporción real. En la gráfica 5.5 se muestra esta comparación para el modelo que incorpora

únicamente el efecto temporal, el que agrega las variables de rendimiento y el que considera variables temporales, de rendimiento y sexo.

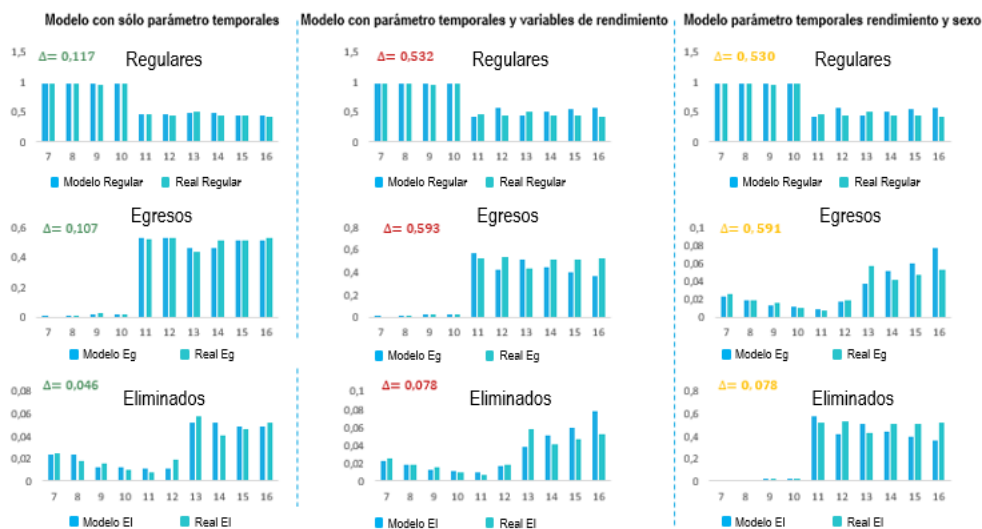


Figura 5.5: Proporción de Estado promedio Modelo vs Realidad

Además, se muestra la suma de la diferencia de cada tasa con la proporción real por todos los periodos (el valor que acompaña a los Δ) y se marca con los colores del semáforo dónde verde es el que tuvo mejores resultado y rojo el de peor resultado.

Se observa que el modelo que incorpora sólo las dummies temporales logra una proporciones casi idénticas que los datos reales, esto es esperable pues la especificación de un modelo con sólo dichas variables no obtiene exactamente los mismos proporciones sólo por las restricciones del logit. Por otro lado, se observa que ambos modelos con variables explicativas obtienen resultados muy similares, pero considerando que la variable dummy sexo permitirá utilizar más heterogeneidad y sus resultados son un poco más cercanos a la proporción real, se decide que es valioso considerarla en el modelo final.

A los modelos anteriores, se decidió agregar heterogeneidad no observada, pues es probable que exista una heterogeneidad no capturada por las variables de los modelos y que sea persistente en el tiempo. Dado que no se puede agregar un parámetro por persona porque significaría añadir miles de parámetros al modelo, se utilizó el enfoque de clases latentes [Kamakura \[1989\]](#), y para no aumentar demasiado el número de parámetros, se decidió agregar sólo en un subconjunto de parámetros. Así, se estudió el efecto de agregar heterogeneidad en sólo el intercepto y en los parámetros temporales. En la siguiente ecuación se muestra la probabilidad para el modelo con heterogeneidad en intercepto y parámetros temporales (intercepto y efecto temporal indexado por clase):

$$P_{ijt} = \sum_c P_c \frac{e^{(\alpha_{ijc} + \beta_{jtc} + \beta_{ij} x_{ijt})}}{\sum_j e^{\alpha_{ijc} + \beta_{jtc} + \beta_{ij} x_{ijt}}} \quad (5.4)$$

En donde, además de los parámetros de los modelos previos, se estiman las probabilidades de pertenecer a cada clase c . En la primera tabla de sección ?? del anexo, se muestran los resultados con 2, 3 y 4 clases² Con las variables del mejor modelo encontrado en el paso anterior.

Se observa que agregar una cuarta clase aporta muy poco a la verosimilitud, pues escasas observaciones pertenecerían a dicha clase. Además, se observa que los modelos con heterogeneidad en los parámetros temporales, tienen un considerable aporte a la verosimilitud, sin embargo, a costa de agregar bastantes parámetros. Debido a esto y a que las diferencias de los parámetros relacionados a la eliminación no eran mayores, se exploró un modelo que sólo agregara heterogeneidad a los parámetros temporales asociados al egreso (modelo “3 Clases B” en segunda tabla de anexo ??).

Con respecto a las diferencias de los parámetros de las clases agregadas (ver tabla ?? en anexo D). En el caso de los modelos con heterogeneidad en sólo los interceptos, se observa que se crea una clase con casi nula probabilidad de eliminación y alta de egreso, otra clase con una menor probabilidad de egreso y alta de eliminación y otra clase intermedia. En el caso de los modelos con heterogeneidad en los parámetros temporales, se observa un comportamiento similar, y además se aprecia cómo pertenecer a cada clase adelanta o retrasa la probabilidad de egreso (la probabilidad de eliminación también tiene diferencias temporales, pero no muy importantes).

Finalmente, se indica que se optó por utilizar el modelo con heterogeneidad en los interceptos y parámetros temporales de egreso, por ser el de mejores resultados, según los criterios BIC y AIC.

5.2.2. Resultados Facultad de Ciencias Físicas y Matemáticas

Modelo elección programa

Como se mencionó anteriormente, el primer paso es modelar el proceso de elección de programa de continuidad al Plan Común en la FCFM, para luego estimar las probabilidades para los estudiantes de cada carrera por separado. Esta metodología es bautizada como Logit Secuencial en Rodríguez [2007] (recordar que esto se realiza porque gran parte de los estudiantes de la cohorte 2016 aún no eligen su programa de continuidad y las probabilidades de transición dependerán, entre otras variables, de cada programa).

Si bien, no se cuenta con suficiente información como para tener un modelo que prediga con alta precisión qué carrera elegirán los estudiantes, se considera que sí puede existir correlación entre la elección de especialidad y alguna de las variables socioeconómicas o de rendimiento presentadas

²No se consideraron los modelos con más clases porque agregaban muy poco a la maximización de la verosimilitud.

anteriormente. Por ejemplo, existen carreras que se caracterizan por tener una baja proporción de estudiantes mujeres, mientras que otras tienen una proporción mucho mayor. Del mismo modo, se considera que puede haber una relación entre los rendimientos y la elección de la especialidad. De esta forma, se cree que es mejor tener una estimación paramétrica a ocupar simplemente las proporciones totales de la muestra.

Los resultados del modelo de elección de carrera se pueden observar en la tabla E .2 del anexo. Tal como se esperaba, las predicciones tendrán baja precisión. Sin embargo los modelos son significativos según la prueba de bondad de ajuste(χ^2).

Los mejores predictores resultaron ser las variables de rendimiento y el sexo, además la variable sexo apunta en el sentido de los juicios previos a la aplicación del modelo; las carreras del departamento IQBT y Geología tienen una mayor probabilidad de ser elegidas por mujeres que el resto de las carreras, mientras que Eléctrica, Matemática, Mecánica y Minas son significativamente más elegidas por hombres. Es llamativo cómo la variable región es significativa (al 99,9 %) para la carrera de minas, en donde los estudiantes de la Región Metropolitana tienen una mucho menor probabilidad de elegir esta carrera, quizás porque los alumnos de regiones ya se han enfrentado a estudiar o trabajar fuera de la ciudad natal y el ejercicio de esa profesión generalmente así lo requiere.

Al ocupar los resultados académicos de tercer año, se puede estar incurriendo en un problema de causalidad inversa. No obstante, para el fin de este trabajo que se quiere predecir la elección y no hacer inferencia sobre los factores de la elección de carrera, esto no tiene mayor importancia.

Modelo de Trayectoria Estudiantil

Tras modelar la elección de carrera, se ajusta un logit multinomial por programa con la misma estructura que los ocupados para Derecho (esto es equivalente a ajustar sólo un logit, pero con todos los parámetros indexados por programa). No obstante, se tuvo que tener en consideración que serían menos observaciones por carrera, por lo que se ajustaron menos parámetros y no se agregaron parámetros para capturar heterogeneidad no observable.

En la figura 5.6 se observan las probabilidades de eliminación y egreso para el alumno promedio de cada año para cada carrera. Las licenciaturas comienzan a tener probabilidad de egreso desde octavo semestre, Ingeniería Matemática y en Computación, también tienen probabilidad de egreso levemente anticipada (tienen una duración de un semestre menos que las demás carreras). Los alumnos del resto de las carreras comienzan a tener probabilidades de egresar a partir del duodécimo y decimotercer semestre.

Los modelos tienen pocos parámetros significativos en comparación al modelo ajustado en derecho, además con respecto a las tasas de titulación, se ve un problema que aún no tiene solución

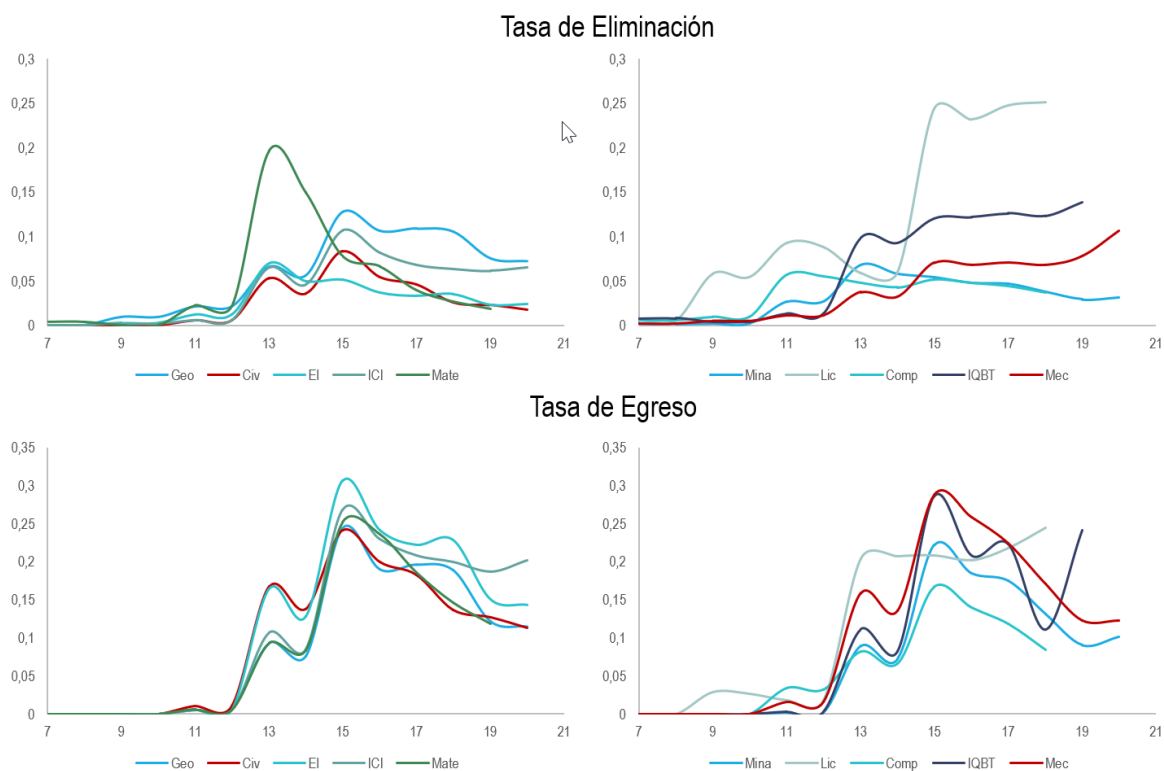


Figura 5.6: Probabilidad de Egreso y Eliminación Por Carrera de la FCFM por Semestre

en la FCFM, correspondiente al alto porcentaje de alumnos que posponen su egreso, debido al alto costo de oportunidad que presenta realizar el trabajo de título a esa altura de la carrera, lo que se traduce en que con las simplificaciones tomadas en este trabajo, sean considerados como alumnos eliminados. De tal forma, en los últimos años de las carreras las tasas de eliminación suben considerablemente.

5.3. Sobre las variables predictoras

Dado el objetivo de este trabajo, se justifica utilizar las variables de rendimiento en tercer año como principal predictor de las trayectorias de los estudiantes. Sin embargo, es de sumo interés entender cómo las variables sociodemográficas impactan los resultados estudiantiles y sería especialmente útil encontrar el efecto del programa de gratuidad.

Para medir el efecto de las variables socioeconómicas se utilizará un enfoque de regresión lineal, utilizando como variable dependiente la cantidad de asignaturas aprobadas hasta tercer año, dado que resultó ser la variable con mayor poder predictivo de los modelos anteriores. Con este objetivo se utilizarán las observaciones de la cohorte 2016. Los resultados se muestran en la tabla 5.4.

Se observa que los predictores estadísticamente significativos corresponden al puntaje promedio

Coefficiente	Facultad Ingeniería	Facultad Derecho
α_0	-21.04***	-22.45**
$PUNTAJE_{POND}$	0.07***	0.076***
Masculino	-0.67	0.0899
IVE_C	1.53**	-0.1813
$IVE_{NoAplica}$	1.25*	1.51**
Gratuidad	0.60	0.55
$ViaAcceso_{NoPsu}$	0.79	-0.10
Regiones	-1.34**	-1.32**
R^2	0,14	0,17
pvalor estadístico F	$< 2 * 10^{-16}$	$< 7 * 10^{-11}$

Cuadro 5.4: Modelos de regresión lineal para predicción de rendimiento académico hasta 3er año

PSU, al IVE del colegio de egreso y al efecto de ser de regiones (egresar de un establecimiento de fuera de la región Metropolitana).

Es importante notar que el estimador del efecto de la gratuidad no es significativo, esto se puede deber a que el efecto del programa sea heterogeneo entre los estudiantes beneficiados y no se haya hecho un correcto control con el resto de las variables.

Se encuentra que el tipo de colegio de egreso es un buen predictor del rendimiento estudiantil y se espera que los alumnos de los establecimientos más vulnerables tengan un peor resultado académico. En los modelos anteriores no se había encontrado este efecto, sin embargo sí aparecía en la literatura estudiada. Es lógico esperar que los alumnos con gratuidad provengan de colegios más vulnerables como se puede ver en la tabla de contingencia de a continuación (se muestra la tabla de ingeniería pero en Derecho se aprecia el mismo efecto).

	Grupos A y B	Grupo C	No aplica	Total
Con Gratuidad	81	98	41	220
Sin Gratuidad	40	134	315	489
Total	121	232	356	709

Cuadro 5.5: Tabla contingencia Gratuidad vs IVE colegio egreso

Con respecto a la región de origen de los estudiantes. Se observa que al igual que en los resultados de los modelos anteriores (modelos de duración), los estudiantes de regiones suelen tener peores resultados que los estudiantes de la Región Metropolitana (algunas posibles razones fueron expuestas en la sección anterior). Es importante ver en tabla 5.6 que los estudiantes con gratuidad parecen tener una mayor proporción de estudiantes de regiones. Esto a primer instante parece contraintuitivo, porque se puede pensar que los estudiantes de regiones más vulnerables deberían tener mayor dificultad a estudiar fuera de su ciudad de origen, sin embargo, se puede deber a que para el cálculo del NSE los ingresos no son normalizados por el costo de vida de cada región.

	Región Metropolitana	Otras regiones	Total
Con Gratuidad	132	88	220
Sin Gratuidad	375	114	489
Total	507	202	709

Cuadro 5.6: Tabla contingencia Gratuidad vs Región colegio de egreso

Por último, sobre el puntaje PSU, se observa una correlación postiva entre el puntaje PSU y el rendimiento académico, además, como se observa en la siguiente tabla, los alumnos con gratuidad tienen peores puntajes que los estudiantes sin gratuidad.

	Puntaje PSU
Con Gratuidad	715
Sin Gratuidad	724
Total	721

Cuadro 5.7: Puntaje PSU alumnos con Gratuidad y sin Gratuidad

Dados los resultados anteriores, se concluye que si bien no se encuentra que la variable Gratuidad sea significativa y pese a que tenga signo postivo, se espera que los estudiantes con el beneficio tengan un peor resultado, en promedio aprobando 0,6 menos ramos hasta tercer año (considerando todas las variables del modelo), en caso de considerar sólo las significativas esta diferencia sube a 1,2 asignaturas.

Capítulo 6

Simulación y evaluación económica

6.1. Simulación

Para cada estudiante de la cohorte 2016 que no ha llegado a un estado terminal, el sistema se inicializa con el último estado conocido:

$E(6) =$ último estado conocido.

Luego, las trayectorias son calculadas con el siguiente algoritmo, que usa los resultados de los modelos logit presentados en el capítulo anterior para calcular las probabilidades de elección de carrera y transición entre estados:

- 1: Iniciar semilla aleatoria en valor predeterminado.
- 2: **for** n in Estudiantes 2016 **do**
- 3: **if** n \in Estudiantes Ingeniería y no ha elegido Carrera **then**
- 4: Se calcula $P_{nc} \quad \forall c \in Carreras$
- 5: Carrera: Simulación Distribucion Categórica¹(P_{nc})
- 6: **end if**
- 7: **for** t in (6:Semestre límite carrera) **do**
- 8: Se calcula $P_{njt} \quad j \in \{Eg, El, R\}$
- 9: $E(t + 1) =$ Simulación Distribucion Categórica(P_{njt})
- 10: **if** $E(t + 1) \in \{Eg, El\}$ **then**
- 11: Exit For (Pasar a siguiente estudiante)
- 12: **end if**
- 13: **end for**
- 14: **end for**

Para la simulación se utilizaron las medias de los parámetros encontrados. No obstante, estos también se podrían haber simulado, utilizando la información de su media y la matriz de varianza-

¹La Distribución Categórica es el caso de la Distribución Multinomial con 1 sólo realización.

covarianza. La simulación se realizó 1000 veces para analizar una amplia cantidad de posibles escenarios.

6.2. Resultados de la simulación

A continuación se muestran los resultados de la simulación y la valorización de las trayectorias de los estudiantes con gratuidad. Para esto último, no se considera el valor del dinero en el tiempo. Sin embargo, cabe decir que los aranceles regulados se reajustan año a año por IPC (3%), por lo que considerar aranceles fijos, es equivalente a considerar una tasa de descuento igual a la inflación.

De este modo, se calcula el monto total que se espera ingrese por gratuidad con los aranceles regulados correspondientes al año 2019 según la siguiente fórmula, considerando que post duración formal los estudiantes deben pagar a lo más la mitad del arancel regulado ²:

$$\text{Ingresos Gratuidad} = \sum_{i \in E} \sum_{t=1}^{D_i} \mathbb{1}_{t \leq D_{ci}} \frac{A}{1,03^t} + \mathbb{1}_{t > D_{ci}} \frac{A}{2 \times 1,03^t} \quad (6.1)$$

con:

E : Estudiantes con Gratuidad.

D_i : Duración alumno i .

D_{ci} : Duración carrera alumno i .

A : Arancel Regulado.

6.2.1. Facultad de Derecho

Trayectoria estudiantil

En las gráficas 6.1 y 6.4, se muestran los resultados de la simulación para la Facultad de Derecho. Dado el objetivo de este trabajo, el resultado de mayor importancia es la duración proyectada de las trayectorias y su consiguiente valorización. Adicionalmente se muestra la proporción de egresos y eliminaciones simuladas. Además, en el histograma de duración promedio, se marca el promedio de las cohortes 2009-2011. En el gráfico de valorización económica se dividen los ingresos asociados a la duración formal y los asociados al sobretiempo, pues como se ha mencionado, el beneficio sólo cubre por la duración formal de los programas y posterior a eso, son los estudiantes los que deberían pagar los aranceles correspondientes.

Se observa que los alumnos con gratuidad presentan proporcionalmente más eliminaciones y menos egresos, es decir, se proyecta que tengan un peor resultado académico. Además, los resultados para la población con gratuidad presentan una mayor dispersión, lo que en parte se puede deber a que sea una muestra considerablemente más pequeña.

²En la tabla 8.1 se presentan los aranceles reales y regulados de las carreras de la FCFM y Derecho

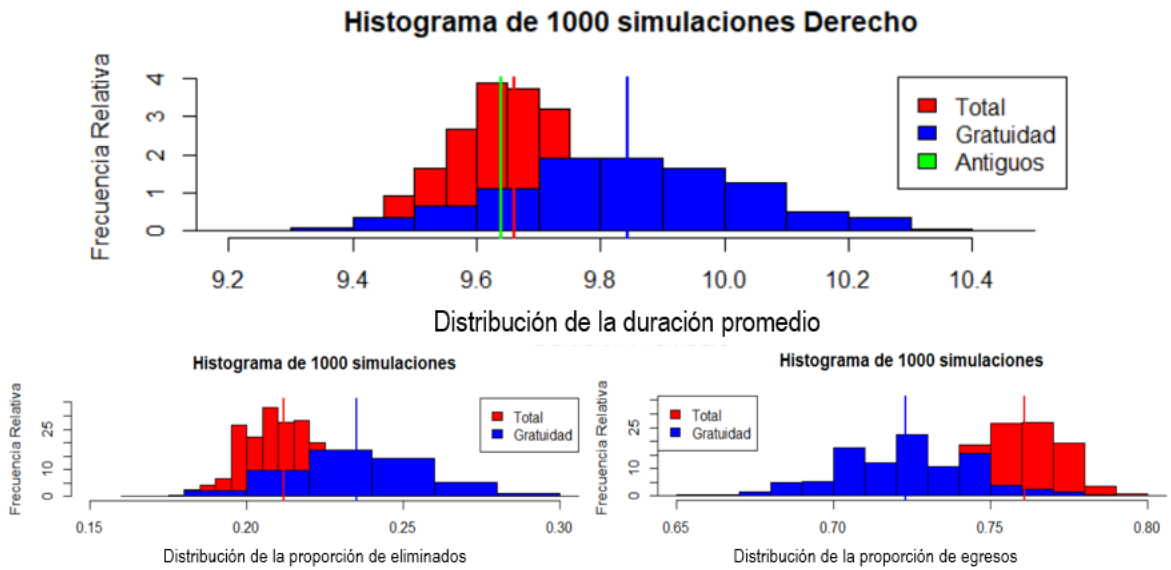


Figura 6.1: Resultado de la Simulación para la Facultad de Derecho

Con respecto a la duración promedio del total de los alumnos, se espera un breve aumento en comparación con las cohortes 2009-2011, una explicación podría ser que con la gratuidad existe una matrícula más diversa que se demora más en egresar. De esta forma, las trayectorias de los alumnos con gratuidad son en promedio casi 0,2 semestres más largas.

A continuación en la figura 6.2, se presentan las distribuciones de los egresos y eliminaciones por semestre, además de la proporción de estudiantes regulares (Se considera que la cantidad de alumnos iniciales corresponde al 100%). En la figura 6.3 se diferencian las distribuciones para los alumnos con gratuidad y sin gratuidad.

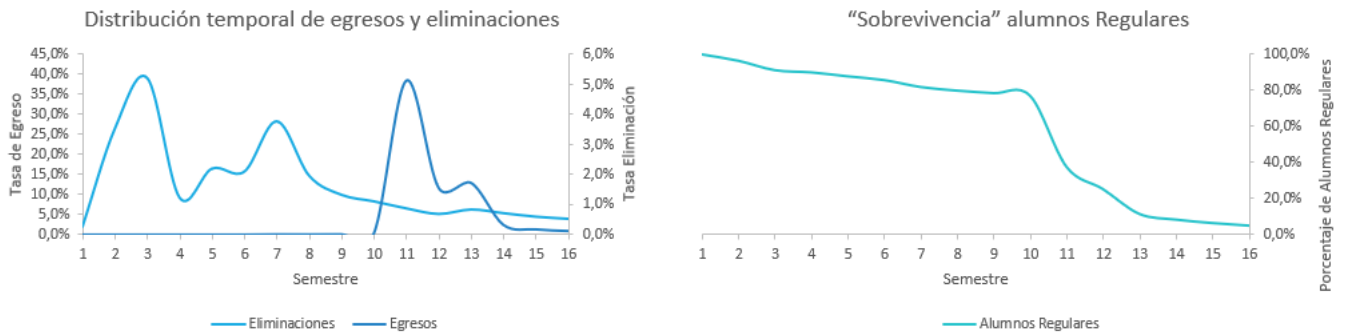


Figura 6.2: Distribución Eliminaciones y Egresos y Supervivencia Alumnos Derecho

Sobre la temporalidad de las eliminaciones, llama la atención la elevada proporción encontrada a los tres años y que posterior a eso descende considerablemente, esto se puede deber a que posterior a tercero, ya se han ido los estudiantes con mayores dificultades y que tras cada año el costo de

oportunidad para alcanzar el egreso es menor. Con respecto al egreso, se observa que el grueso de los estudiantes logra egresar en el tiempo de la duración formal del programa, típicamente el indicador de graduación oportuna³, considera como adecuado un retraso de hasta un año sobre la duración formal, según esa regla, en este caso cerca del 90 % de los egresos se considerarían como oportunos.

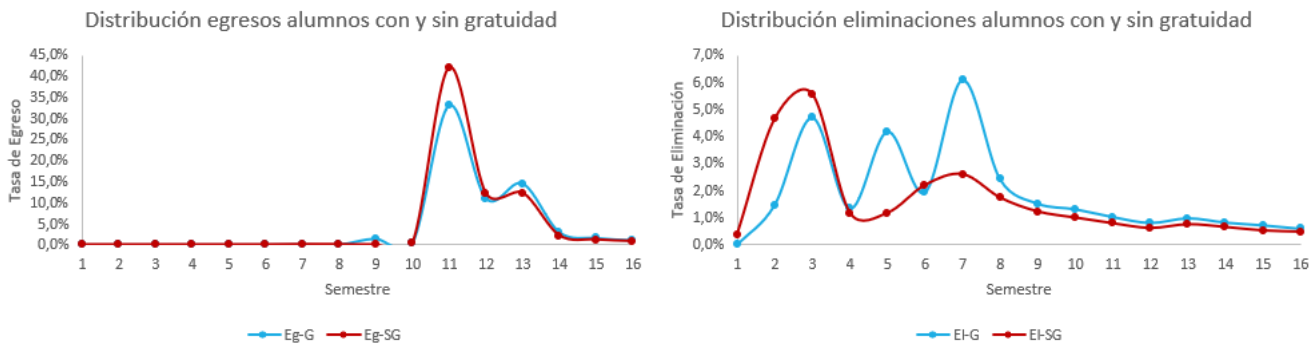


Figura 6.3: Distribución Eliminaciones y Egresos Alumnos Con Gratuidad vs Alumnos Sin Gratuidad

Sobre las proporciones de los estudiantes con y sin gratuidad, se establece que no se observan grandes diferencias entre ambos grupos de estudiantes. Se proyecta que los alumnos sin gratuidad tengan una mayor proporción de egreso temprano (al décimo semestre correspondiente a la duración formal del programa), además que en los primeros años la proporción de abandono de los alumnos sin gratuidad es ligeramente mayor, pero después de segundo año esta situación se revierte (recordar que los 6 primeros semestres no se simulan, porque ya se cuenta con dicha historia).

Evaluación económica

Los ingresos totales para la Facultad de Derecho están en torno a los \$2.700 millones de pesos, mientras que los ingresos post duración formal, se acercan a los \$202 millones de pesos. Como se comentó anteriormente, la gratuidad sólo cubre por la duración formal, por lo que teóricamente este valor debe ser pagado por los estudiantes y no el Estado.

No existe demasiada variación entre las simulaciones, la desviación estándar es de \$42 millones. Lo anterior se explica debido a que son más de 100 las trayectorias que se analizan, y es esperable que su duración promedio no varíe demasiado.

6.2.2. Facultad de Ciencias Físicas y Matemáticas

A continuación, se muestran los resultados para la FCFM, en primer lugar la distribución de elección de carreras y luego histogramas similares a los presentados para la Facultad de Derecho.

³Indicador utilizado para medir la eficiencia de la docencia en las universidades por los entes supervisores como SIES o CNA.

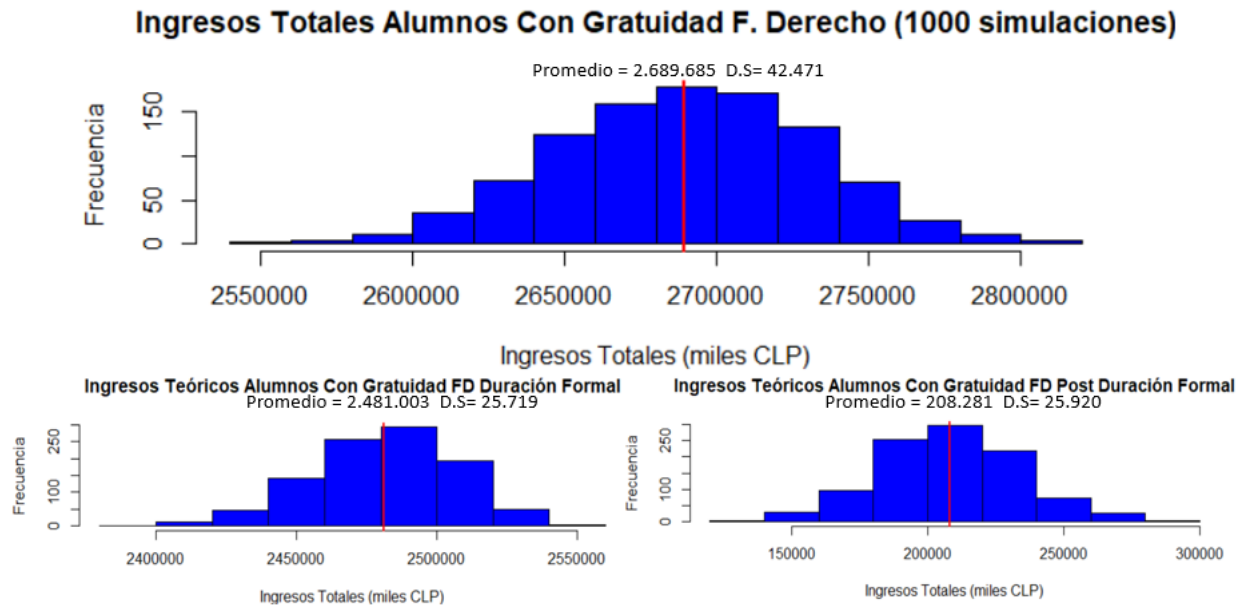


Figura 6.4: Simulación Ingresos Gratuidad Facultad de Derecho

Se compararon las proporciones simuladas, con las proporciones históricas y se mantienen similares.

En los siguientes histogramas (figura 6.5) se muestra la proyección de Duración, proporción de Egresos y Eliminados para todas las carreras de la FCFM en conjunto. Se observa que la proporción de estudiantes que logra egresar, es muy similar a la proporción de estudiantes que eliminan. Se observa que la proporción de egresos y eliminaciones pese a ser mejores que las históricas (encontradas para la cohortes 2009, 2010 y 2011), todavía presentan muchas oportunidades para mejorar. Este problema es un poco menos grave considerando que las observaciones censuradas que llevaban más de dos años inactivas a la fecha de obtención de la información, se consideraron como eliminadas

	N Estudiantes Promedio	Desviación Estándar
Geología	81,2	8,78
Ing. Civil	101,8	8,92
Ing. Eléctrica	68,8	7,77
Ing. Matemática	40	5,31
Ing. Industrial	166,5	11,15
Ing. Computación	50,6	6,24
Ing. Mecánica	48,5	6,64
Ing. Minas	65,7	7,66
Licenciaturas	30,1	4,95
IQBT	44,7	6,31
Sin Carrera	158,1	8,24

Cuadro 6.1: Distribución de Alumnos FCFM por carrera en simulaciones

pese a tener oportunidades de egresar, sin embargo, las cifras no dejan de ser preocupantes.

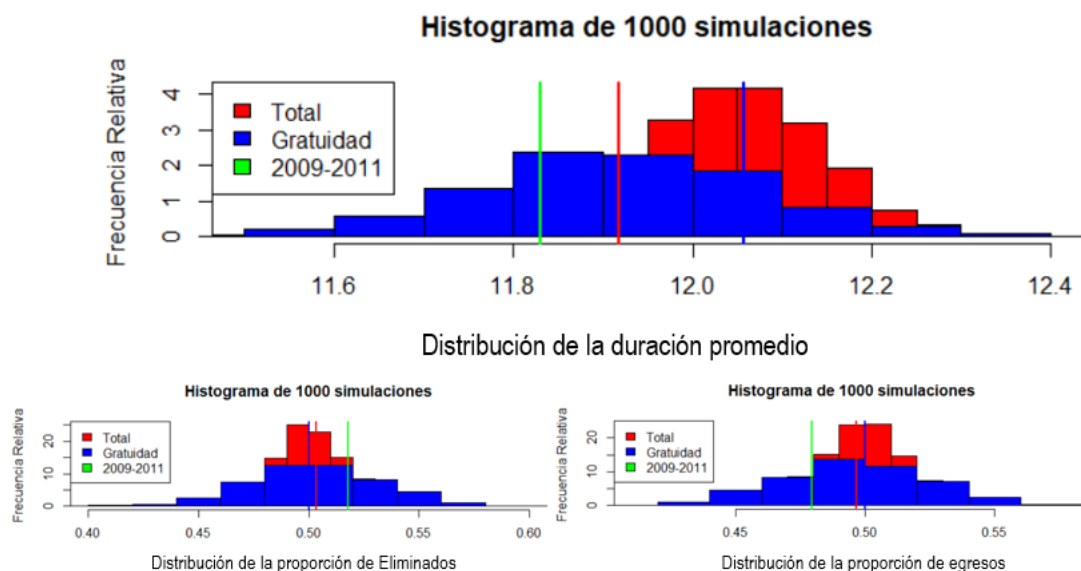


Figura 6.5: Resultados de la Simulación para la Facultad de Ciencias Físicas y Matemáticas

En la figura 6.6, se aprecia la proporción de estudiantes que egresan y eliminan cada semestre y la proporción de alumnos regulares restantes por semestre (se considerará que el 100% corresponde a la cantidad de alumnos activos el primer semestre). Se observa que la proporción de estudiantes eliminados sube demasiado en los últimos semestres, producto a problemas con el proceso de egreso. Además, que la proporción de estudiantes activos tras el octavo año comienza a ser muy baja, por lo que luchar por conseguir que el financiamiento por gratuidad se prolongue sin límite tras la duración formal, no es tan relevante (desde el punto de vista de la universidad), porque con uno o dos años sobre la duración formal es suficiente para capturar el grueso de los ingresos.

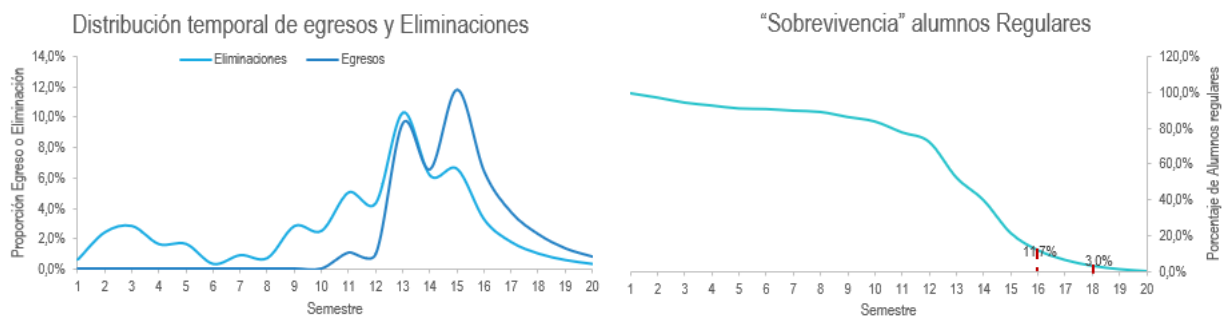


Figura 6.6: Distribución Eliminaciones, Egresos y Supervivencia Alumnos FCFM

Finalmente, en la figura 6.7, se comparan las proporciones de egresos y eliminaciones por semestre para los alumnos con y sin gratuidad. Se aprecia que las curvas para ambos grupos de alumnos son

casi idénticas (principalmente la de egresos). Esto se debe a que entre los estimadores de los modelos no existen variables sociodemográficas (sólo el sexo) y a que los resultados académicos en tercer año de los alumnos con y sin gratuidad no presentan grandes diferencias. Sin embargo, en los primeros años (que no fueron simulados), se aprecia que los estudiantes con gratuidad fueron más propensos a eliminar.

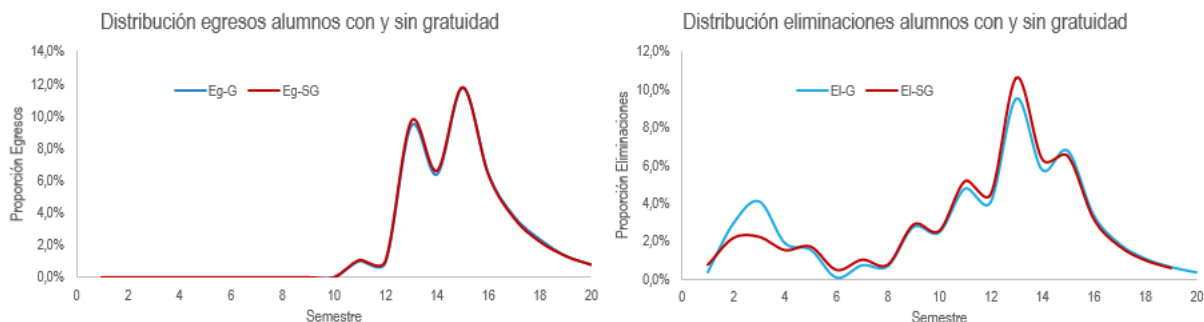


Figura 6.7: Distribución Eliminaciones, Egresos y Supervivencia Alumnos FCFM

En la tabla 6.2 se muestra la duración promedio y desviación estándar por carrera. Se aprecia que la duración promedio es al menos un semestre máyor a la duración formal de los programas (la duración promedio considera a los estudiantes eliminados).

	Duración promedio	Desviación Estándar
Sin Carrera	5,69	0,242
Geología	13,7	0,261
Ing. Civil	13,57	0,232
Ing. Eléctrica	13,45	0,277
Ing. Matemática	13,1	0,358
Ing. Industrial	13,74	0,151
Ing. Computación	13,37	0,472
Ing. Mecánica	13,65	0,369
Ing. Minas	13,7	0,347
Licenciaturas	11,82	0,492
Ing. Química y BT	13,34	0,382

Cuadro 6.2: Promedio y Desviación Estándar de la duración por carrera

La permanencia de los estudiantes de los programas más cortos, efectivamente es menor, sin embargo la diferencia real es menor a la diferencia reglamentaria. Este fenómeno puede parecer extraño en el caso de los estudiantes de licenciaturas, ya que para alcanzar el título no deben hacer un trabajo de título (se considera que es uno de los principales factores que alargan la duración en el resto de las carreras). Sin embargo, las licenciaturas presentan una menor tasa de eliminación y las eliminaciones en el caso de las otras carreras disminuyen los tiempo promedios.

Con respecto a los ingresos, se observa que están en torno a los \$6.500 millones de pesos con una desviación estándar de aproximadamente \$71 millones, de esta forma el percentil 10 y 90 de los resultados tienen una diferencia de menos de \$220 millones. Del mismo modo que en la Facultad de Derecho, no existe una dispersión considerable pese a que en la FCFM se aumenta la variabilidad del proceso al agregar la elección de carrera.

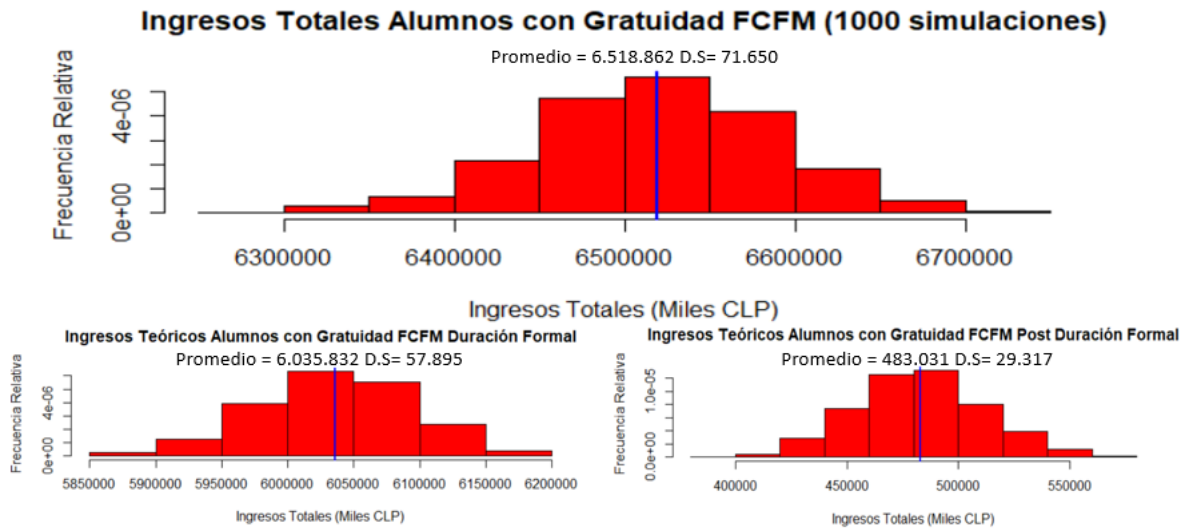


Figura 6.8: Simulación Ingresos Gratuidad Facultad de Ciencias Físicas y Matemáticas

Se considera que se podría disminuir aún más la dispersión de los ingresos con un arancel regulado único para todas las carreras.

Capítulo 7

Sensibilización de los Parámetros Críticos

Se sensibilizarán 4 parámetros que a juicio del autor son de suma relevancia por el alto impacto que tienen y porque pueden sufrir variaciones en el mediano plazo. Se realizó la sensibilización sólo para la Facultad de Derecho.

7.1. Arancel Regulado y Duración de Beneficio

El valor del arancel regulado afecta directamente los ingresos de la universidad y es una de las variables que pueden cambiar fuertemente en el corto plazo, ya que se espera que las nuevas definiciones de la comisión de expertos creadas para definir una nueva metodología de cálculo del arancel regulado, considere componentes que hasta el momento no habían sido tomados en cuenta. Hoy en día las definiciones de los aranceles de referencia y regulados sólo consideran los precios de mercado, más una ponderación por el nivel de calidad asociado a cada carrera, medido como el nivel de los académicos, su productividad científica y acreditación de cada institución.

En el siguiente histograma se grafica el efecto de variar un peso el arancel regulado para la Facultad de Derecho. Se separa en ingresos de la duración formal y post duración formal. Se

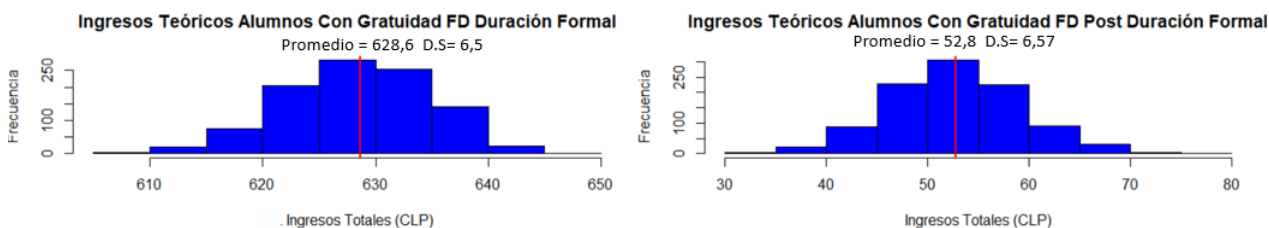


Figura 7.1: Cambios en los ingresos por gratuidad al variar 1 peso el arancel regulado Derecho

observa que una variación de un peso puede significar hasta casi 690 pesos menos de ingresos para la facultad, lo que parece irrelevante. Sin embargo, los montos podrían variar en mucho más de un

peso, de esta manera, un cambio de 100.000 en el arancel regulado, significará un cambio en los ingresos de aproximadamente 69 millones de pesos (el impacto es lineal).

Además, se calculó el efecto de aumentar el tiempo de cobertura de la gratuidad. En la tabla 7.1 se muestran los ingresos extras dependiendo del número de años por los que se extienda el beneficio.

Cobertura	Ingresos (Millones CLP)	Diferencia (Millones CLP)
Duración Formal + 0 Años	2.481	0
Duración Formal + 1 Año	2.691	210
Duración Formal + 2 Años	2.776	295
Sin restricción	2.897	416

Cuadro 7.1: Ingresos adicionales al extender cobertura por gratuidad, sin considerar ingresos post-duración

Se observa que el aumento de ingresos al sensibilizar esta restricción es bastante considerable. Con un aumento de cobertura de un año, ingresarían cerca de 210 millones adicionales, correspondientes a casi el 10 % de los ingresos totales por este concepto. Además, se observa que aumentar la cobertura por más tiempo, tiene un impacto marginal decreciente, y que en caso de negociar esta restricción, se debería abogar por al menos aumentar en un año la cobertura.

7.2. Retención

En general las facultades de la Universidad de Chile tienen altos niveles de retención en comparación a la media nacional. Existe consenso en que el objetivo debe ser mantenerla alto, pues la deserción y no completitud de los estudios significan altos costos para los estudiantes, el Estado y las instituciones. Sin embargo, siguen existiendo oportunidades de mejora y riesgos de empeorar al tratar de lograr una matrícula más diversa.

Teniendo en consideración lo descrito en el párrafo anterior, la probabilidad de eliminación se sensibilizó en un rango de $\pm 10\%$.

En esta sensibilización, a diferencia de las otras simulaciones, también se proyectarán los tres primeros años, para los cuales la probabilidad de eliminación de cada estudiante con gratuidad será la tasa muestral de dichos estudiantes. De esta forma, la probabilidad base para cada semestre $t \leq 6$ será:

$$P_{El-t} = \frac{N_{El-t}}{N_{Reg-t}} \quad (7.1)$$

Así, se obtuvieron los siguientes resultados.

Como se puede apreciar, los cambios no son cuantiosos, dado que las tasas sensibilizadas en $\pm 10\%$ son muy pequeñas como para tener un efecto significativo. En ingeniería se encontraría un cambio más grande en la proporción de estudiantes eliminados, pero sin mayor impacto en la

Indicador\Factor Retención	0,9	1	1,1
Eliminados promedio	32,14	34,13	37,01
Duración total promedio	10,75	10,66	10,55
Valorización (Millones CLP)	2.705	2.689	2.680

Cuadro 7.2: Sensibilización Tasa de Retención Facultad de Derecho

valorización, ya que las tasas de eliminación son altas en los últimos periodos, en donde las tasas de egreso son aún más altas.

7.3. Duración

Se considera que la duración es otro factor clave que debe ser sensibilizado, principalmente por dos motivos:

1. Dado que la gratuidad beneficia a los estudiantes sólo por la duración formal de los programas, los estudiantes con el beneficio podrían verse motivados a terminar más rápido que antiguamente. Por otro lado, se espera que la composición de la matrícula cambie paulatinamente y esto puede tener asociado cambios en el comportamiento de los estudiantes.
2. A juicio del tesista, la universidad se debería adaptar a las tendencias mundiales en cuanto a la duración de sus carreras, en donde las carreras de pregrado suelen durar entre 4 y 5 años (Estados Unidos y Europa). En este sentido, se destaca que los alumnos que ingresaron el 2019 a la FCFM, cursarán una nueva malla que reduce de 12 a 11 semestres el plan de estudios.

Esta sensibilización se llevará a cabo desplazando en un rango de ± 1 y ± 2 semestres las tasas de egreso base.

Las simulaciones que adelantan las tasas de egreso deberían reflejar el efecto de acortar las carreras en 1 y 2 semestres respectivamente. Mientras que los casos en que se retrasa el egreso, se podrían dar si los nuevos estudiantes tienen un peor desempeño académico.

Atraso/Adelanto	Prop. Egresos	Duración media (Semestres)	Ingresos (MMClp)
+ 2 semestres	70,60 %	8,46	2.401
+ 1 semestre	70,20 %	9,09	2.552
Situación Actual	69,60 %	9,73	2.689
- 1 semestre	69,20 %	10,37	2.781
- 2 semestres	68,95 %	10,98	2.868

Cuadro 7.3: Resultados Sensibilización de Adelantar y Retrasar las probabilidades de Egreso

Se observa que acortar o alargar la carrera de derecho en un año, tendría un impacto cercano a los 200 millones de pesos. Es obvio que mientras más tiempo dure la relación de los estudiantes con

la universidad, mayores serán los ingresos. Desde este punto de vista, se puede concluir que existe un incentivo perverso para que las universidades extiendan el tiempo de estadía de los estudiantes, por lo que es importante regular y fiscalizar las duraciones e incentivar a las universidades a maximizar la titulación oportuna. La diferencia de ingresos al acortar la carrera se puede compensar mediante el aumento de los de cupos de las carreras.

Finalmente se observa que mientras más se retrasa la probabilidad de egreso, es menor la proporción de estudiantes que logran egresar, pues se alarga el tiempo que están expuestos al riesgo de eliminar.

7.4. Composición de la matrícula

En el caso que la cantidad de estudiantes de cada decil socioeconómico cambie, los ingresos de la universidad también variarán. A priori, se pensaría que el efecto en la composición económica sería un aumento en la diversidad y representación de los quintiles menos favorecidos. Sin embargo [Bucarey \[2018\]](#) plantea que esta política podría estar desplazando a los alumnos de menos recursos, que ya contaban con beneficios, en favor de los estudiantes de sectores medios.

Para que la comparación con los resultados anteriores tenga sentido se deben ocupar los mismos modelos. Sin embargo, estos no incluyen variables socioeconómicas para la estimación de la trayectoria (sólo el sexo), por lo que la sensibilización estará dada por la valoración de las mismas trayectorias y no el cambio de estas.

Calcular la diferencia de ingresos por los cambios en la composición estudiantil no es trivial. Se debe considerar que si aumenta la proporción de estudiantes con el beneficio, disminuirá la cantidad de estudiantes sin acceso a este. Sobre estos últimos estudiantes, no existe la restricción que permite cobrar máximo la mitad del arancel después de la duración formal.

Para calcular el efecto de aumentar la proporción de estudiantes con gratuidad, se compararon los ingresos de un alumno con gratuidad y otro sin gratuidad, los que fueron elegidos aleatoriamente entre las trayectorias de los alumnos con y sin gratuidad respectivamente. Este ejercicio se realizó 1000 veces por tipo de alumno, obteniendo los resultados de la tabla 7.4.

	%Elimina	% Egreso	Tiempo estudio (semestres)	Ingresos beneficio (\$MM)	Ingresos no cubiertos (\$MM)	σ tiempo estudio
Alumno con gratuidad	31,0 %	69,0 %	11,05	17,3	2,7	5,23
Alumno sin gratuidad	25,1 %	74,9 %	11,23	22,7	NA	4,91

Cuadro 7.4: Sensibilización comparación alumno con y sin gratuidad

De esta forma, se puede notar que los estudiantes con gratuidad presentan peores resultados y más heterogeneidad en la duración de sus estudios. Esto, junto al hecho de que post duración formal se puede cobrar máximo el 50 % del arancel regulado a los alumnos con gratuidad, tiene un impacto de 2,7 millones en la diferencia de ingresos por alumno, que aumenta a 5,4 si no se consideran los ingresos post duración formal de los alumnos con gratuidad (como se ha repetido, es esperable que no los deban pagar). Así, si se aumenta la proporción de alumnos con gratuidad de forma significativa, el impacto sería considerable (impacto lineal sobre 2,7 o 5,4 millones dependiendo si se considera el ingreso post duración formal).

El beneficio aplica para el 60 % de los estudiantes más vulnerables del país, por lo que si la matrícula de la Universidad de Chile representará de igual forma a todas las clases sociales, de los 420 estudiantes de la cohorte 2016, 252 deberían ser de aquellos quintiles y contar con el beneficio (108 más que en la situación actual). De este modo, el impacto sería de 291 o 583 millones de pesos (dependiendo si se contabiliza o no los montos no cubiertos por el beneficio), con respecto a la situación actual.

Finalmente, se piensa que este factor no será de alto impacto a corto plazo, pues la composición de la matrícula no se diversificará de forma repentina. Además, si se cambiarán las metodologías de selección con el fin de garantizar una matrícula más diversa, es probable que los nuevos alumnos no obtengan los resultados deseados, pues arrastran problemas de formación que deben ser solucionados en primer lugar. En conclusión, el autor considera que la universidad debe apuntar a lograr la matrícula más diversa posible, pero que esto se debe lograr de modo responsable.

Capítulo 8

Trabajos Futuros y Conclusiones

8.1. Trabajos futuros: Estimación del impacto en el comportamiento de los estudiantes

Se considera que existe bastante margen para mejorar los modelos de estimación de probabilidades de transición, pues el foco de este trabajo estuvo en tener probabilidades de transición coherentes y que capturaran lo posible de heterogeneidad para la simulación, pero sería deseable contar con mejores modelos para tener probabilidades de transición más precisas.

Con lo anterior, se podrá realizar inferencia y un mayor análisis del impacto de las variables (lo que en este trabajo no se hizo en profundidad debido a que se considera que los estimadores no eran lo suficientemente insesgados).

Además del mejoramiento de los modelos, se plantea que un trabajo futuro fundamental, es entender cómo la gratuidad afecta el comportamiento de los estudiantes. En este trabajo fueron sensibilizadas las variables en que se consideró que podría haber más impacto, pero se plantea que con diseños estadísticos cuasiexperimentales se podría obtener mayor certeza sobre cual fue y cual será el impacto real en dichos comportamientos.

Se considera que debido a la forma en que fue implementada la gratuidad, se pueden aprovechar métodos cuasiexperimentales como Difference in Difference (dado el aumento que ha existido en la cobertura) o de regresión discontinua (dado el score necesario para acceder al beneficio).

Esta última, se ha popularizado para medir el impacto de distintos tipos de políticas a las que se puede acceder por cumplir cierto criterio (medible y continuo). En Chile, en el ámbito de la educación ver [Bucarey et al. \[2018\]](#), [Montoya et al. \[2017\]](#), que miden los efectos del CAE (Crédito con Aval del Estado) o el ya citado artículo de [Bucarey \[2018\]](#) que lo utiliza para medir los efectos de la gratuidad.

Por último, se indica que además de medir estos impactos de la forma más precisa posible, es necesario aplicar políticas que ayuden a reducir posibles efectos negativos y permitir que los

estudiantes que accedan a la universidad, puedan terminarla de un modo satisfactorio.

8.2. Conclusión

La gratuidad ha significado una inversión de miles de millones de pesos para el Estado y puede implicar importantes desafíos para las instituciones, que aún no conocen el impacto total que tendrá en sus finanzas. Por lo tanto, se plantea que es necesario contar con métodos técnicos objetivos y con la mayor precisión posible para analizar el éxito de esta forma de financiamiento. El presente trabajo busca ser un aporte a la medición de algunos efectos de la gratuidad desde la perspectiva de la gestión universitaria, entregando metodologías para estimar los ingresos asociados a esta política.

Se considera que pese a que no exista mucha literatura sobre los modelos de duración discretos (en comparación a los continuos), estos presentan muchas herramientas y flexibilidad para modelar el proceso educativo. Además su elección parece la correcta considerando que los eventos están bastante agrupados en ciertos intervalos de tiempo. Se considera que se presenta una herramienta muy potente para entender los ordenes de magnitud en que distintos factores impactan los ingresos que entrarán a la Universidad por la Gratuidad.

Con respecto a los estimadores de las trayectorias de los estudiantes, se debe mencionar que el mejor predictor resultó ser el rendimiento en el semestre más cercano del periodo a estimar, siendo el de mayor relevancia la cantidad de asignaturas aprobadas. Las variables de rendimiento concentraron casi todo el poder explicativo, haciendo que las variables socioeconómicas perdieran significancia, similar a lo ocurrido en el trabajo de [Celis et al. \[2015\]](#).

Dado el objetivo del trabajo, se justifica la elección de variables explicativas, sin embargo se considera que sería un aporte conocer el impacto de las variables socioeconómicas. Por último, se destaca la inclusión de heterogeneidad no observable, permitiendo efectos no observables persistentes en el tiempo (a través del intercepto) y diversidad en los efectos temporales (a través de las variables temporales).

Adicionalmente, se considera que sería deseable contar con más observaciones (en ciertos semestres hay muy pocas eliminaciones o egresos, por lo que es difícil obtener parámetros significativos en dichos periodos), para poder ajustar modelos de mayor complejidad y que permita hacer inferencia sobre el efecto de las variables. De esta forma se podrían agregar más variables socioeconómicas, variables que cambien en el tiempo como las calificaciones¹ y variables cuyo efecto sea diferente en el tiempo, pues se consideraron sólo variables con efecto homogéneo en el tiempo.

¹dado que el foco de este trabajo era predecir la trayectoria de una cohorte que todavía no se completa, no tenía sentido utilizar variables que cambiaran en el tiempo.

A juicio del autor, el principal aporte de este trabajo corresponde a la simulación y la posterior sensibilización de sus parámetros. Se encontró que sin alterar las reglas de la gratuidad, los ingresos por este concepto pueden cambiar considerablemente dependiendo de las trayectorias de los estudiantes. De esta forma para la FCFM, ordenando de peor a mejor las 1000 simulaciones (peor a mejor sólo en el sentido de los ingresos), en el percentil 10 ingresan \$6.450 millones, mientras en el percentil 90 ingresan \$6.600 millones, es decir una diferencia cercana a los \$150 millones de pesos, que se convierte en más de \$300 millones si se compara el percentil 1 con el 99. Para la Facultad de Derecho la diferencia entre el percentil 10 y 90 es menor, pero de igual forma es considerable, superando los \$100 millones de pesos.

Sobre la sensibilización de los parámetros de la simulación, se observa que pequeños cambios en los parámetros pueden tener un significativo impacto en los ingresos de las facultades. A juicio del autor, los parámetros que presentan la mayor oportunidad o riesgo son el arancel regulado y la duración del beneficio. Según este estudio una variación de un peso del arancel regulado en la Facultad de Derecho tiene un impacto de aproximadamente \$700 pesos, por lo que por ejemplo, una variación de 1 millón de pesos en el arancel regulado tendría un efecto de \$700 millones de pesos en los ingresos de la Facultad. En este momento el arancel regulado y el arancel real para la facultad de derecho presenta una diferencia de aproximadamente 700 mil pesos, mientras que para las carreras de ingeniería es de cerca de 1 millón 500 mil. Es decir, esta brecha puede significar diferencias de ingresos muy significativas.

Del mismo modo, se calculó la diferencia en los ingresos al extender en un año la cobertura del beneficio, encontrando un aumento cercano a los 200 millones de pesos y 400 millones si se elimina esta restricción.

Otro efecto relevante encontrado, corresponde al que puede producir acortar o alargar las carreras². El autor considera que es necesario que las carreras de la Universidad de Chile se adapten a los estándares internacionales y que en ese sentido se deberían reducir las duraciones, pero existe un incentivo perverso, pues los ingresos son proporcionales a la duración de las carreras, aunque esto se puede revertir agrandando la matrícula. En este trabajo se encontró que acortar la carrera de derecho en un año significaría una disminución de aproximadamente 280 millones en el concepto de gratuidad correspondiente a más del 10% del ingreso proyectado.

Es importante mencionar que los resultados encontrados no son extensibles a otras facultades o universidades. Sin embargo, se puede aplicar la misma metodología, mientras tengan la información necesaria para ajustar los modelos. Además se debe recalcar que para tener una visión completa

²Se considera un riesgo controlable, dado que el riesgo de que se aumenten considerablemente los tiempos de graduación es pequeño, y los tiempos formales de los programas son controlados por la universidad.

de la rentabilidad asociada a la docencia de pregrado, es necesario contar con una mejor gestión de costos que permita definir los gastos asociados a los distintos macroproductos universitarios. Un análisis que queda propuesto, es comparar la situación post ley de gratuidad con la situación previa, en donde los estudiantes de los niveles socioeconómicos más bajos ya contaban con beneficios.

Por último, se considera que es de vital importancia que la Universidad de Chile esté preparada de la mejor manera posible para enfrentar las amenazas externas, como la incertidumbre respecto a las reglas de financiamiento externo y la mejora en calidad del resto de las universidades del sistema educacional chileno. Para cumplir con lo anterior es necesario que eleve al máximo posible los estándares de gestión. Frente a esto, se destacan dos medidas: La implementación de un sistema ERP que está llevando a cabo la Vicerrectoría de Asuntos Económicos y Gestión Institucional, y la construcción de un datawarehouse institucional que permitirá centralizar los datos de toda la universidad.

Bibliografía

Mineduc. Metodología de determinación de arancel de referencia 2019 para universidades acreditadas - ministerio de educación. 2018.

Judith D. Singer and John B. Willett. It's about time: Using discrete-time survival analysis to study duration and the timing of events. *Journal of Educational Statistics*, 1993.

Marc A. Scott and Benjamin B. Kennedy. Pitfalls in Pathways: Some Perspectives on Competing Risks Event History Analysis in Education Research. *Journal of Educational and Behavioral Statistics*, 2005.

Germán Rodríguez. Lecture notes on generalized linear models., 2007. URL <http://data.princeton.edu/wws509/notes/>.

Chandra Bhat. Duration models. In DAVID A. HENSHER and KENNETH J. BUTTON, editors, *Handbook of Transport Modelling*, chapter 6, pages 91–110. PERGAMON, 2007.

Judith D. Singer and John B. Willett. It's déjà vu all over again: Using multiple-spell discrete-time survival analysis. *Journal of Educational and Behavioral Statistics*, 1995.

Stephen L. DesJardins, Dennis A. Ahlburg, and Brian P. McCall. An event history model of student departure. *Economics of Education Review*, 1999.

Stephen L. DesJardins, Dennis A. Ahlburg, and Brian P. McCall. A temporal investigation of factors related to timely degree completion. *The Journal of Higher Education*, 2002a.

Stephen L. DesJardins, Dennis A. Ahlburg, and Brian P. McCall. Simulating the longitudinal effects of changes in financial aid on student departure from college. *The Journal of Higher Education*, 2002b.

Stephen L. DesJardins, Dennis A. Ahlburg, and Brian P. McCall. The effects of interrupted enrollment on graduation from college: Racial, income, and ability differences. *Economics of Education Review*, 2006.

- Juan Carlos Calcagno, Peter Crosta, Thomas Bailey, and Davis Jenkins. Stepping stones to a degree: The impact of enrollment pathways and milestones on community college student outcomes. *Research in Higher Education*, 2007.
- Alex J. Bowers. Grades and graduation: A longitudinal risk perspective to identify student dropouts. *The Journal of Educational Research*, 2010.
- Daniela Glocker. The effect of student aid on the duration of study. 2011.
- S. Celis. Student Attrition and Student Time-to-Degree at a Selective Engineering School in Chile. 2012.
- S. Celis, P. Poblete, J. Villanueva, and R. Weber. Un modelo analítico para la predicción del rendimiento académico de estudiantes de Ingeniería. *Revista Ingeniería de Sistemas*, 2015.
- Nelson Aguirre. Factores que predicen el rendimiento académico en la escuela de ingeniería de la universidad de Chile, 2012.
- Matías Grau. Análisis de la capacidad del ranking de notas para predecir el resultado académico de los estudiantes de la fcfm, 2015.
- Camila Aguirre. Superación académica en primer año de ingeniería y ciencias: Mecanismos de permanencia y mejoramiento académico, 2016.
- Isamar Troncoso Marcel Goic, Ricardo Montoya. Estimating the breakage rate of an airline's frequent-flyer program using individual level data. 2019.
- Peter Dolton and Wilbert von der Klaauw. Leaving teaching in the UK: A duration analysis. *The Economic Journal*, 105(429):431–444, 1995.
- Russell G. J. Kamakura, W. A. A probabilistic choice model for market segmentation and elasticity structure. *Journal of marketing research*, pages 379–390, 1989.
- Alonso Bucarey. Who Pays for Free College? Crowding Out on Campus. 2018.
- Alonso Bucarey, Dante Contreras, and Pablo Muñoz. Labor market returns to student loans. 2018.
- Ana Maria Montoya, Carlos Noton, and Alex Solis. Returns to higher education: Vocational education vs college. 2017.
- Ley núm. 21.091: Sobre educación superior - ministerio de educación, 2018. 29 de Mayo de 2018.

Apéndices

A . Cálculo de Arancel de Referencia y Regulado

A .1. 1. Constitución de los grupos de universidades

Se evalúan los siguientes indicadores³: Grado académico de los profesores, productividad científica (publicaciones y proyectos), retención primer año y titulación oportuna. Se estandarizan y promedian para rankear a las universidades y segmentarlas en 4 grupos.

A .2. 2. Definición aranceles de referencia originales por carrera.

Se agrupan todas las carreras con nombres similares y pertenezcan al mismo tipo de programa (profesional con licenciatura, licenciatura no confluente a título profesional, etc.), carreras nuevas que no tengan otras similares se agruparan bajo nombres genéricos similares o categorías genéricas similares.

En cada grupo regirá como el arancel de referencia el de la carrera acreditada por más años, en caso de empates se considerará el de menor valor (si es menor que el promedio del grupo, se elige dicho promedio).

El arancel que rige en el grupo se indexa de acuerdo al reajuste de remuneraciones del sector público.

A .3. 3. Definición aranceles regulados

Para el cálculo del arancel regulado, el año 2019 se creó una comisión que fijará los parámetros y condiciones para determinar que arancel asignar a cada carrera. Además de criterios de calidad, incorporará nociones de costo.

Temporalmente, hasta que dicha comisión entregue resultados, se utilizarán los aranceles referenciales como base del cálculo de los aranceles regulados.

³Mineduc [2018] Ley [2018]

B . Aranceles Regulados Facultad de Derecho y FCFM

En la siguiente tabla se aprecian los aranceles regulados para las carreras de la FCFM y Facultad de Derecho:

Carrera	Arancel Regulado	Arancel Real	Matricula	Deficit
Derecho	3.946.359	4.691.400	139.200	884.241
Ingeniería Civil	4.254.789	5.490.200	139.200	1.374.611
Ingeniería Civil Eléctrica	4.254.789	5.490.200	139.200	1.374.611
Ingeniería Civil en Computación	4.202.943	5.490.200	139.200	1.426.457
Ingeniería Civil de Minas	4.046.312	5.490.200	139.200	1.583.088
Ingeniería Civil Industrial	4.046.312	5.490.200	139.200	1.583.088
Ingeniería Civil Matemática	3.526.901	5.490.200	139.200	2.102.499
Ingeniería Civil Química	4.254.789	5.490.200	139.200	1.374.611
Ingeniería Civil Mecánica	4.254.789	5.490.200	139.200	1.374.611
Ingeniería Y Ciencias - Plan Común	4.430.416	5.490.200	139.200	1.198.984
Ingeniería Y Ciencias - Plan Común	4.254.789	5.490.200	139.200	1.374.611
Geología	3.413.010	5.490.200	139.200	2.216.390
Astronomía	3.296.317	5.490.200	139.200	2.333.083
Geofísica	3.296.317	5.490.200	139.200	2.333.083
Física	3.296.317	5.490.200	139.200	2.333.083

Elaboración Propia en base a decreto exento N°1 (2019) del Mineduc

Figura 8.1: Aranceles reales y regulados de las carreras de la FCFM

C . Marco Conceptual - Glosario

Para lectores ajenos a las metodologías utilizadas en este trabajo, se presenta esta sección con una breve descripción de los conceptos/herramientas técnicas aplicadas, la explicación es sólo conceptual, sin embargo, existe abundante literatura sobre todos los temas tratados para profundizar.

C .1. Estimación por Máxima Verosimilitud

Es un método de estimación que se basa en encontrar los parámetros que harían más probables los valores de la muestra estudiada. Se deben hacer supuestos sobre la distribución de la variable estudiada y, a partir de esta distribución y los valores de la muestra, se crea la función de verosimilitud dependiente de los parámetros $L(\theta|y)$.

En ocasiones la función es diferenciable y su solución se puede obtener analíticamente (caso logit), pero en otros casos es necesarios recurrir a otros métodos de optimización.

C .2. Probabilidades de transición

Corresponden a las probabilidades de cambiar de un estado i a otro j en un periodo t . Las probabilidades modeladas en este trabajo corresponden a la probabilidad de seguir siendo alumno regular, egresar o eliminar en un semestre.

C .3. Heterogeneidad observable y no observable

La heterogeneidad corresponde a las diferencias de las observaciones. En el contexto de este trabajo a las diferencias de los estudiantes. Existen diferencias que pueden ser observadas y capturadas por los sistemas de información (heterogeneidad observable), y otras que no son observables, o no pueden ser capturada por los sistemas de información, por lo que el investigador debe trabajar con esta limitante y aplicar métodos para asegurar la validez de las conclusiones.

C .4. Regresión logística Multinomial - Logit Multinomial

Es una extensión del modelo logit simple de elección discreta para problemas con más de dos resultados posibles.

El modelo logit se puede desprender al considerar que la utilidad para un sujeto n por una alternativa i , está dada por $u_{ni} = v_{ni} + \epsilon_{ni}$, en donde el primer término corresponde a una componente sistemática/observable y la segunda a una componente aleatoria que distribuye valor extremo tipo 1. Al asumir esto, las probabilidades de elección de cada alternativa se calcula como la probabilidad de que su utilidad sea mayor a la otra. De esta forma, se obtiene la siguiente fórmula para la probabilidad de elección de i por parte de individuo n en tiempo t (se normaliza la otra alternativa a utilidad 0): $P_{nit} = \frac{e^{v_{nit}}}{1 + e^{v_{nit}}}$.

Para el modelo multinomial también se asume que la componente aleatoria de la función de utilidad tiene una distribución valor extremo tipo 1 o gumbel e independientes para todas las alternativas. De esta forma, la fórmula para calcular la probabilidad de elegir la alternativa i se convierte en: $P_{ni} = \frac{e^{v_{ni}}}{\sum_j e^{v_{nj}}}$

C .5. Modelo de clases latentes

Es una metodología para hacerse cargo de la heterogeneidad no observada. En esta, se asume una cantidad discretas de clases o tipos de sujetos con diferentes comportamientos. Para cada clase, se calculan las probabilidades de pertenencia y sus parámetros asociados.

C .6. Criterios de selección AIC y BIC

Los criterios de selección AIC y BIC, son criterios utilizados para la selección de modelos que se basan en la verosimilitud, pero que penalizan por la cantidad de parámetros. Por cada parámetro agregado la verosimilitud debería subir o al menos mantenerse constante, por lo que castigando por la cantidad de parámetros se evitan sobreajustes y no se incorporan los parámetros que no contribuyen de modo considerable a la verosimilitud.

C .7. Algoritmo Kmedias

Es un método utilizado con el objetivo de agrupar observaciones en k grupos. En dónde, cada observación se asigna al grupo con centro más cercano. El centro es el punto medio de los componentes del grupo, pero existen variantes en donde en vez de la media se calcula la mediana, en esta aplicación se utiliza la distancia euclidiana, pero otras medidas pueden ser usadas.

D . Algoritmo kmedias - Regla del codo

A continuación se observa la suma de la varianza dentro de los cluster por el número de cluster para la FCFM y Facultad de Derecho.

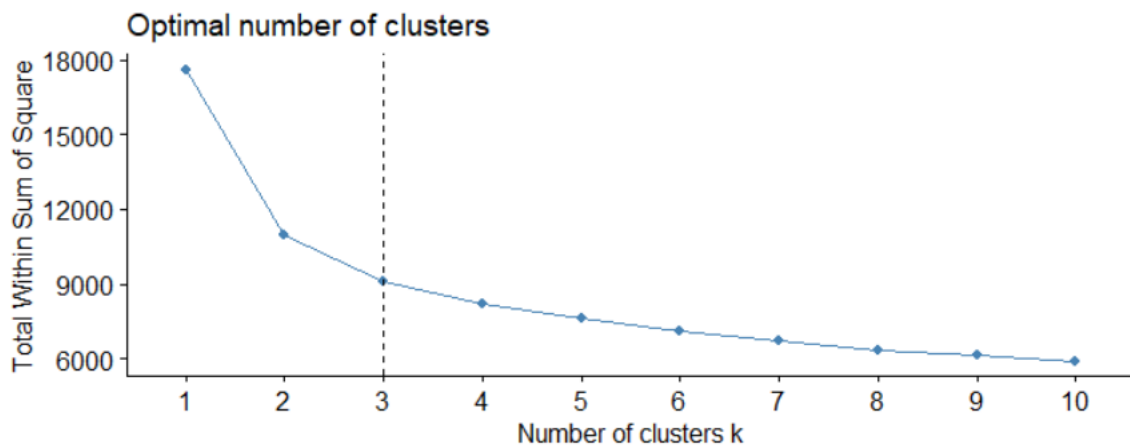


Figura 8.2: Regla del codo suma de varianza dentro del cluster, clusterización por rendimiento FCFM

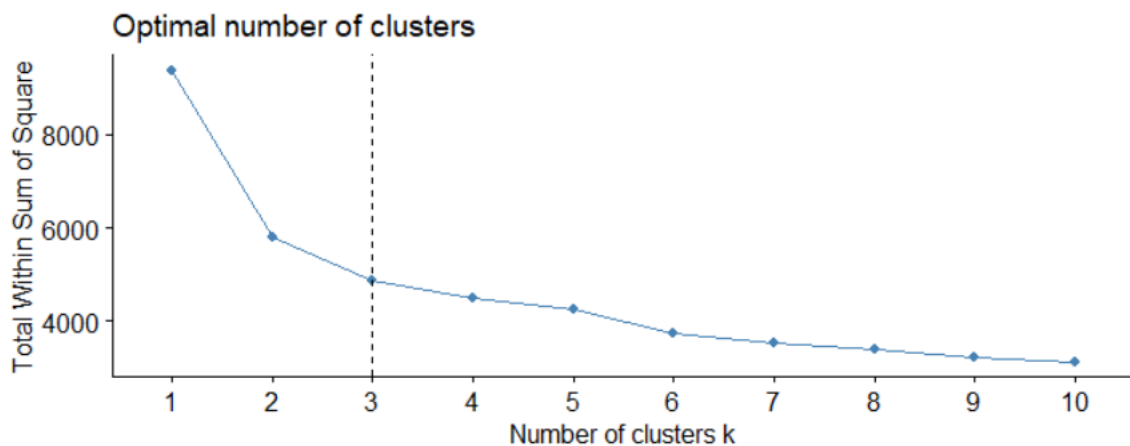


Figura 8.3: Regla del codo suma de varianza dentro del cluster, clusterización por rendimiento Derecho

E . Resultados Logit

E .1. Clases Latentes Derecho

En la siguiente tabla se muestran los resultados de los modelos con 2, 3 y 4 clases latentes con heterogeneidad en solamente los interceptos.

	Modelo 2 Clases	Modelo 3 Clases	Modelo 4 Clases
α_{EIC1}	6.39	6.06	1.12
α_{EgC1}	-17.43	-25.28	-23.06
α_{EIC2}	0.46	0.45	8.20
α_{EgC2}	-16.18	-23.39	-24.20
α_{EIC3}		-36.06	13.55
α_{EgC3}		-14.12	-26.65
α_{EIC4}			-41.09
α_{EgC4}			-13.88
$SexFem_{El}$	-0.51	-0.49	-0.67
$PA3_{El}$	-0.08	-0.07	-0.10
$AP3_{El}$	-0.46	-0.46	-0.75
5_{El}	-0.10	-0.14	-0.39
6_{El}	-0.13	-0.23	-0.81
7_{El}	0.96	0.81	0.28
$SexFem_{Eg}$	0.22	0.20	0.19
$PA3_{Eg}$	0.11	0.13	0.12
$AP3_{Eg}$	0.29	0.35	0.35
5_{Eg}	3.18	7.16	7.11
6_{Eg}	7.83	13.37	13.24
7_{Eg}	9.06	14.81	14.68
PC1	14,8 %	16.18 %	75.30 %
PC2	85.2 %	81.96 %	20.65 %
PC3		1.85 %	2,23 %
PC4			1.83 %
LL	1604	1591	1590
AIC	3.243	3.223	3.225
BIC	3.356	3.356	3.379

Cuadro 8.1: Resultados Clases Latentes Derecho

En la siguiente tabla se muestran los resultados de los modelos con 2 y 3 clases latentes con heterogeneidad en los interceptos y parámetros temporales. Además, se agrega un tercer modelo llamado “3 Clases B”, en donde los parámetros temporales de la eliminación son los mismos para las 3 clases.

E .2. Elección Programa de Continuidad Ingeniería

En la siguiente tabla se muestra los resultados del modelo para simular la elección de carrera de las carreras de continuidad de la FCFM, con el nivel de referencia puesto en ingeniería industrial,

	2 Clases	3 Clases	3 Clases B*
α_{ElC1}	-17,37*	0,31	2,47*
α_{EgC1}	-20,64	-23,15	-23,85*
α_{ElC2}	2,51*	2,50*	-8,31
α_{EgC2}	-17,61*	-23,77	23,97*
α_{ElC3}		2,57	2,14
α_{EgC2}		-15,47*	-17,23*
$SexFem_{El}$	-0,34	-0,26	-0,34
$Promedio_{El}$	-0,04*	-0,03	-0,04*
$Aprobado_{El}$	-0,38*	-0,43*	-0,38*
$Sexo_{Eg}$	0,27*	0,23	0,25
$Promedio_{Eg}$	0,12*	0,14*	0,15*
$Aprobado_{Eg}$	0,37*	0,42*	0,48*
$5C1_{El}$	-8,07	-5,92	-0,25
$6C1_{El}$	-6,45	-5,40	-0,27
$7C1_{El}$	1,39	-7,71	0,14
$5C1_{Eg}$	6,29	5,39	3,00
$6C1_{Eg}$	10,44	11,66	10,05
$7C1_{Eg}$	14,74	16,19	9,13
$5C2_{El}$	-0,26	-0,03	
$6C2_{El}$	-0,18	-1,72	
$7C2_{El}$	0,14	0,35	
$5C2_{Eg}$	-9,08	6,00	4,99
$6C2_{Eg}$	7,21*	12,17	11,57
$7C2_{Eg}$	5,75*	10,28	16,42
$5C3_{El}$		-1,86	
$6C3_{El}$		9,16	
$7C3_{El}$		-3,97	
$5C3_{Eg}$		7,95*	8,95*
$6C3_{Eg}$		8,80	8,52*
$7C3_{Eg}$		18,15	21,78
PC1	49,5 %	50,64 %*	50,79 %*
PC2	50,1 %	47,58 %*	47,49 %*
PC3		1,77 %*	1,72 %*
LL	1.577	1.555	1.559
AIC	3.200	3.174	3.171
BIC	3.353	3.388	3.344

Cuadro 8.2: Resultados Clases latentes con heterogeneidad en parámetros temporales e intercepto

que fue elegida por ser la carrera con más estudiantes.

	Estimate	Std. Error
SinCarrera:(intercept)*	3.70	0.59
046 GEOLOGÍA:(intercept)	-0.17	0.47
047 INGENIERÍA CIVIL:(intercept)*	-1.74	0.46
048 INGENIERÍA CIVIL ELÉCTRICA:(intercept)*	-4.23	0.61
050 INGENIERÍA CIVIL MATEMÁTICA:(intercept)*	-5.50	0.91
051 INGENIERÍA CIVIL MECÁNICA:(intercept)*	-2.31	0.59
052 INGENIERÍA CIVIL DE MINAS:(intercept)*	-1.26	0.52
053 INGENIERÍA CIVIL QUÍMICA:(intercept)*	-1.93	0.63
058 LICENCIATURA EN CIENCIAS:(intercept)*	-2.60	0.77
059 INGENIERÍA CIVIL EN COMPUTACIÓN:(intercept)	-0.85	0.58
SinCarrera:AP3*	-0.45	0.05
046 GEOLOGÍA:AP3	0.05	0.04
047 INGENIERÍA CIVIL:AP3	0.05	0.04
048 INGENIERÍA CIVIL ELÉCTRICA:AP3	0.07	0.04
050 INGENIERÍA CIVIL MATEMÁTICA:AP3*	-0.33	0.05
051 INGENIERÍA CIVIL MECÁNICA:AP3	0.06	0.05
052 INGENIERÍA CIVIL DE MINAS:AP3	0.06	0.04
053 INGENIERÍA CIVIL QUÍMICA:AP3*	0.18	0.05
058 LICENCIATURA EN CIENCIAS:AP3	0.02	0.06
059 INGENIERÍA CIVIL EN COMPUTACIÓN:AP3*	-0.12	0.04
SinCarrera:AP2*	-0.23	0.06
046 GEOLOGÍA:AP2	-0.05	0.04
047 INGENIERÍA CIVIL:AP2	0.06	0.04
048 INGENIERÍA CIVIL ELÉCTRICA:AP2*	0.12	0.05
050 INGENIERÍA CIVIL MATEMÁTICA:AP2*	0.51	0.07
051 INGENIERÍA CIVIL MECÁNICA:AP2	0.02	0.05
052 INGENIERÍA CIVIL DE MINAS:AP2	-0.00	0.05
053 INGENIERÍA CIVIL QUÍMICA:AP2	-0.04	0.06
058 LICENCIATURA EN CIENCIAS:AP2	0.09	0.07
059 INGENIERÍA CIVIL EN COMPUTACIÓN:AP2	0.02	0.05

Cuadro 8.3: Resultados Modelo Seleccionado Para Elección de Carrera 1/2

	Estimate	Std. Error
SinCarrera:Region	-0.07	0.28
046 GEOLOGÍA:Region	-0.28	0.21
047 INGENIERÍA CIVIL:Region	-0.01	0.20
048 INGENIERÍA CIVIL ELÉCTRICA:Region	0.26	0.25
050 INGENIERÍA CIVIL MATEMÁTICA:Region	0.23	0.35
051 INGENIERÍA CIVIL MECÁNICA:Region	-0.38	0.24
052 INGENIERÍA CIVIL DE MINAS:Region*	-1.06	0.21
053 INGENIERÍA CIVIL QUÍMICA:Region	-0.38	0.26
058 LICENCIATURA EN CIENCIAS:Region	-0.37	0.32
059 INGENIERÍA CIVIL EN COMPUTACIÓN:Region	-0.01	0.26
SinCarrera:SexoMASCULINO	0.23	0.35
046 GEOLOGÍA:SexoMASCULINO*	-0.51	0.21
047 INGENIERÍA CIVIL:SexoMASCULINO	0.18	0.21
048 INGENIERÍA CIVIL ELÉCTRICA:SexoMASCULINO*	1.40	0.34
050 INGENIERÍA CIVIL MATEMÁTICA:SexoMASCULINO*	0.97	0.46
051 INGENIERÍA CIVIL MECÁNICA:SexoMASCULINO*	0.78	0.31
052 INGENIERÍA CIVIL DE MINAS:SexoMASCULINO*	0.65	0.27
053 INGENIERÍA CIVIL QUÍMICA:SexoMASCULINO*	-0.71	0.25
058 LICENCIATURA EN CIENCIAS:SexoMASCULINO	-0.07	0.33
059 INGENIERÍA CIVIL EN COMPUTACIÓN:SexoMASCULINO	0.57	0.32
*Marca para los estimadores con significancia al 95 %		Log Verosimilitud -3381

Cuadro 8.4: Resultados Modelo Seleccionado Para Elección de Carrera 2/2