



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

**ESTUDIO DEL EFECTO DEL PRIMER CANJE DE PUNTOS EN EL  
COMPORTAMIENTO DEL CLIENTE MEDIANTE CAUSAL FOREST**

**MEMORIA PARA OPTAR AL TÍTULO DE INGENIERO CIVIL INDUSTRIAL**

**NICOLÁS AGUSTÍN YÁÑEZ MORENA**

PROFESOR GUÍA:  
RICARDO MONTOYA MOREIRA

MIEMBROS DE LA COMISIÓN:  
MARCEL GOIC FIGUEROA  
DANIEL SCHWARTZ PERLROTH

SANTIAGO DE CHILE  
2020

RESUMEN DE LA MEMORIA PARA OPTAR AL  
TÍTULO DE INGENIERO CIVIL INDUSTRIAL  
POR: NICOLÁS AGUSTÍN YÁÑEZ MORENA  
FECHA: 2020  
PROFESOR GUÍA: SR. RICARDO MONTOYA  
MOREIRA

## **ESTUDIO DEL EFECTO DEL PRIMER CANJE DE PUNTOS EN EL COMPORTAMIENTO DEL CLIENTE MEDIANTE CAUSAL FOREST**

Hoy en día, las empresas tienen muchos incentivos para fidelizar a sus clientes. Uno de los métodos más usados actualmente es el de los clubes de lealtad, los que tienen como fin fidelizar a sus clientes mediante recompensas y beneficios exclusivos. Una de las maneras que tienen los clientes para obtener recompensas es mediante la acumulación de puntos, para luego canjearlos por premios. Sin embargo, la efectividad de estos programas de fidelización no está asegurada, ya que no siempre tienen un impacto en la lealtad de los clientes. Es por esa razón que en el presente estudio se analiza el impacto que tiene el primer canje de puntos en el comportamiento de compra del cliente.

La estimación del efecto que tiene el primer canje de puntos en el comportamiento del cliente representa un desafío metodológico interesante, ya que la intervención que se quiere realizar – el primer canje de puntos en este caso – no es asignada aleatoriamente, sino que son los mismos sujetos en estudio quienes deciden canjear o no. Esto conlleva que exista un sesgo por autoselección, debido a que el canje podría deberse a características específicas de los clientes que canjearon y que generan esas diferencias con respecto a los clientes que no canjearon. Es por esta razón que ambos grupos no pueden ser comparados directamente. Para reducir los efectos que tiene el sesgo por autoselección, en este estudio la estimación del efecto se lleva a cabo utilizando un algoritmo de Machine Learning llamado Causal Forest. Este método consiste en agrupar a los clientes similares, mediante una clasificación basada en árboles, para posteriormente definir el efecto del primer canje de puntos dentro de cada hoja de los árboles. Esta metodología se utilizó para estimar el efecto del primer canje de puntos en dos clubes de lealtad: uno con clientes de una empresa bajo régimen contractual y otra con clientes de otra empresa bajo régimen no contractual. Además, se usaron los resultados individuales para estimar una caracterización del efecto encontrado en los clientes bajo régimen contractual.

Para el caso de los clientes bajo régimen contractual, se observó en los clientes que realizaron su primer canje de puntos un aumento estadísticamente significativo de 12,7% en el monto gastado mensualmente en la empresa con respecto a los que no canjearon puntos, mientras que no se observó una diferencia significativa en la cantidad de productos consumidos mensualmente en la empresa. Además, se observó que el impacto del primer canje de puntos en el monto gastado mensualmente es mayor en clientes con un consumo relativamente alto, en hombres y en personas con una edad cercana a los 40 años. Para el caso de los clientes bajo régimen no contractual, se observó un aumento del 25,4% en el monto de compra promedio mensual, un aumento de la frecuencia promedio de compra de 8,85% y un adelantamiento de la primera compra posterior al primer canje de una semana aproximadamente. Sin embargo, no se detectó que este efecto fuese distinto entre los clientes que sí realizaron su primer canje.

En base a estos resultados, se puede concluir que el primer canje de puntos sí tiene un efecto significativo en el comportamiento de compra de los clientes, se encuentren éstos bajo régimen contractual o no.

# DEDICATORIA

A Mara Morena,

Mi mamá. Todavía no tengo claro quién se ha esforzado más para lograr esto durante estos 25 años, si yo o ella. Puede que la memoria de título sea la guinda de la torta al esfuerzo, pero sería lindo que esa guinda lleve su nombre.

# TABLA DE CONTENIDO

DEDICATORIA.....	ii
TABLA DE CONTENIDO .....	iii
ÍNDICE DE TABLAS.....	v
ÍNDICE DE ILUSTRACIONES.....	vi
1. INTRODUCCIÓN.....	1
2. LITERATURA PREVIA.....	3
3. OBJETIVOS.....	6
3.1. OBJETIVO GENERAL .....	6
3.2. OBJETIVOS ESPECÍFICOS .....	6
4. ALCANCES.....	6
5. MARCO TEÓRICO.....	7
5.1. CAUSALIDAD.....	7
5.2. PROPENSITY SCORE.....	9
5.3. CAUSAL FOREST.....	10
6. DESARROLLO METODOLÓGICO.....	15
6.1. APLICACIÓN N° 1: RÉGIMEN CONTRACTUAL.....	15
6.1.1. DATOS DISPONIBLES .....	15
6.1.2. ASIGNACIÓN DE GRUPOS EXPERIMENTALES Y VARIABLES A ESTUDIAR.....	16
6.1.3. ESTIMACIÓN DEL EFECTO MEDIANTE CAUSAL FOREST.....	18
6.1.4. ESTIMACIÓN DEL EFECTO MEDIANTE METODOLOGÍAS ALTERNATIVAS.....	20
6.1.5. CARACTERIZACIÓN INDIVIDUAL DEL EFECTO ESTIMADO MEDIANTE CAUSAL FOREST.....	21
6.1.5.1. Caracterización del efecto en el ARPU mensual.....	21
6.1.5.2. Caracterización del efecto en los productos comprados mensualmente.....	26
6.1.6. CONCLUSIONES DE LA PRIMERA APLICACIÓN.....	27
6.2. APLICACIÓN N° 2: RÉGIMEN NO CONTRACTUAL.....	28
6.2.1. DATOS DISPONIBLES .....	28
6.2.2. DETERMINACIÓN DE GRUPOS EXPERIMENTALES Y VARIABLES A ESTUDIAR.....	29
6.2.3. ESTIMACIÓN DEL EFECTO MEDIANTE CAUSAL FOREST.....	32
6.2.4. ESTIMACIÓN DEL EFECTO MEDIANTE METODOLOGÍAS ALTERNATIVAS.....	36

6.2.5. CARACTERIZACIÓN INDIVIDUAL DEL EFECTO ESTIMADO MEDIANTE CAUSAL FOREST .....	37
6.2.6. CONCLUSIONES DE LA SEGUNDA APLICACIÓN.....	38
7. CONCLUSIONES.....	40
8. TRABAJOS FUTUROS.....	42
9. BIBLIOGRAFÍA .....	43
10. ANEXOS .....	45

# ÍNDICE DE TABLAS

Tabla 1: Distribución de los canjes por mes dentro del período de tratamiento para la primera aplicación.....	16
Tabla 2: Tipos de variables de la primera aplicación .....	17
Tabla 3: Caracterización de grupos experimentales según variables disponibles para la primera aplicación.....	18
Tabla 4: Resultados de la estimación del efecto mediante Causal Forest para la primera aplicación .....	19
Tabla 5: Importancia relativa de las covariables en la construcción de los bosques de la primera aplicación.....	19
Tabla 6: Resultados de la estimación del mejor estimador lineal para la primera aplicación .....	20
Tabla 7: Comparación de las estimaciones del efecto mediante las tres metodologías para la primera aplicación .....	21
Tabla 8: Diferencias por género del efecto estimado mediante Causal Forest de la primera aplicación .....	22
Tabla 9: Diferencias por zona geográfica del efecto estimado mediante Causal Forest de la primera aplicación .....	22
Tabla 10: Resultados de la regresión lineal del CATE estimado mediante Causal Forest en función de las covariables para la primera aplicación.....	23
Tabla 11: Caracterización de los segmentos de clientes de la primera aplicación.....	25
Tabla 12: Caracterización de los clientes con efecto estadísticamente significativo para el segundo caso de la primera aplicación .....	26
Tabla 13: Niveles de canje.....	28
Tabla 14: Distribución de los canjes por mes dentro del período de canje para la segunda aplicación.....	30
Tabla 15: Tipos de variables para la segunda aplicación .....	30
Tabla 16: Caracterización de grupos experimentales según variables disponibles para la segunda aplicación.....	31
Tabla 17: Covariables utilizadas para cada caso de la segunda aplicación .....	32
Tabla 18: Hiperparámetros para cada caso de la segunda aplicación.....	33
Tabla 19: Resultados de la estimación mediante Causal Forest para la segunda aplicación.....	34
Tabla 20: Importancia relativa de las covariables en la construcción de los bosques para el efecto en monto de compra de la segunda aplicación .....	34
Tabla 21: Importancia relativa de las covariables en la construcción de los bosques para el efecto en frecuencia y recencia de compra de la segunda aplicación .....	34
Tabla 22: Resultados de la estimación del mejor estimador lineal para la segunda aplicación ....	35
Tabla 23: Resultados de la estimación del efecto al mes siguiente mediante Causal Forest para la segunda aplicación.....	35
Tabla 24: Comparación de las estimaciones del efecto mediante las tres metodologías para la primera aplicación .....	36
Tabla 26: Correlación entre el CATE estimado mediante Causal Forest y las covariables más importantes para la segunda aplicación .....	38

# ÍNDICE DE ILUSTRACIONES

Ilustración 1: Ejemplo de árbol de clasificación .....	11
Ilustración 2: Algoritmo Double-Sample Trees .....	12
Ilustración 3: Ventana temporal del tratamiento .....	16
Ilustración 3: Estimación mediante Causal Forest del efecto del primer canje en el ARPU mensual en función de la edad.....	23
Ilustración 4: Ventana temporal del tratamiento para la segunda aplicación .....	29

# 1. INTRODUCCIÓN

Hoy en día, las personas se ven expuestas a una gran variedad de alternativas para satisfacer sus necesidades de consumo. Esto ha incentivado a las empresas a retener a sus clientes, usando mecanismos que fomenten su lealtad. Ésta se define como la fuerza de la relación cliente-firma y tiene dos dimensiones: la primera, de carácter conductual, es la decisión de volver a comprar un producto en el tiempo. La segunda tiene que ver más con la actitud del cliente y es el apego que tiene éste con la marca o la empresa. La lealtad conductual está relacionada con los patrones de compra, como la frecuencia o el volumen, mientras que la lealtad actitudinal se relaciona con el nivel de compromiso del cliente o el afecto positivo (Dick & Basu, 1994).

Tener clientes leales puede significar un beneficio adicional a capturar nuevos clientes, dado que el costo de ofrecerles el producto o servicio es menor, aumentan sus compras con el paso del tiempo, realizan recomendaciones a sus cercanos y suelen ser menos sensibles al precio (Reichheld & Sasser, 1990).

Uno de los métodos más utilizados por las empresas para retener a sus clientes son los programas de lealtad o clubes de recompensas. Estos programas son aquellos que permiten a los consumidores obtener recompensas o premios de manera gratuita tras realizar compras repetitivamente en la firma, con el fin de promover la lealtad de sus participantes (Liu, 2007). En la misma línea, el diccionario de la Asociación Americana de Marketing (American Marketing Association) se refiere a los programas de lealtad como “programas de incentivos continuos ofrecidos por un minorista (retailer) para recompensar a sus clientes e incentivar compras repetitivas”.

Si bien cada programa tiene sus propias reglas para recompensar la lealtad de sus clientes, generalmente éstos son diseñados en base a una de dos estructuras. La primera, recompensa al cliente leal por la frecuencia de sus compras. Por compras de cierto monto el cliente suele recibir un premio. También recompensa mediante descuentos o regalos a aquellos clientes que cumplan ciertos requisitos, diferenciando muy poco entre clientes. La segunda estructura asigna a los clientes a distintos segmentos, dependiendo de qué tan valiosos sean para la empresa. Generalmente obtienen mejores beneficios los clientes que significan más para la empresa (Dorotic et. al., 2012).

Independiente de la estructura de cada programa, el principal objetivo de un club de fidelización es el de fomentar y recompensar la lealtad de los clientes. La lealtad conductual no implica necesariamente que el cliente tenga una actitud leal, ya sus compras repetitivas podrían estar siendo generadas solamente por una falta de alternativas. Para inducir efectos sostenibles en el tiempo, el club debe potenciar “*lealtad verdadera*”, es decir, un comportamiento leal, reforzado por una actitud positiva de apego (Dorotic et. al., 2011).

Esta confusión entre lealtad conductual y lealtad verdadera se puede dar a menudo en la industria de las telecomunicaciones, en la cual, por norma general, el cliente contrata un servicio por un período de tiempo fijo, ya sea de telefonía o servicios para hogar, haciendo muy poco probable que durante ese período que el cliente pueda adquirir productos de la competencia. En cambio, en la industria del retail, debido a la amplia variedad de alternativas y los bajos costos de cambio entre firmas, la lealtad de un cliente puede resultar mucho más atractiva de fomentar, ya que el cliente no necesariamente se encuentra obligado a consumir por un contrato. Por esta razón, este estudio



se enfoca en analizar el efecto del programa de lealtad para clientes bajo régimen contractual y no contractual.

Se encuentren los clientes de una empresa bajo un régimen contractual o no, las firmas esperan lograr una modificación en el comportamiento de compra de ellos a través de los programas de lealtad. En particular, para el caso de las relaciones no contractuales, el éxito de un programa de lealtad debiese conllevar un aumento en la frecuencia de compra del cliente, un aumento de los montos de las compras en el tiempo y un aumento en su cuota de cartera<sup>1</sup> (Dorotic et. al., 2012). Por otro lado, para el caso de las relaciones contractuales, se busca una renovación de los contratos y posibles ascensos de categoría.

La relación causa-efecto entre la participación de un cliente en un programa de lealtad y el cambio en su comportamiento de compra puede ser estudiada usando inferencia causal (Morgan & Winship, 2007). Si bien existen investigaciones previas que estudian el impacto de los programas de lealtad en el comportamiento del cliente, pareciera no haber un consenso en las conclusiones. Carboni (2012) desarrolló una investigación en la cual se encuentra un impacto positivo del programa. En cambio, Sharp & Sharp (1997) muestran evidencia de que no existe ningún efecto significativo. Algunas investigaciones, como la de Ríos (2013), demuestran que existe evidencia suficiente de un impacto positivo sólo para algunos segmentos del total de clientes.

Esta ambigüedad en los resultados puede ser explicada por dificultades metodológicas que existen al estimar el efecto, ya que el investigador no controla qué individuos son sometidos a la intervención – en este caso, la participación en un programa de lealtad – y cuáles no, dado que son los clientes quienes eligen participar del programa. Dado esto, puede ser que la participación en el programa sea determinada por ciertas características que tienen los clientes y que los motivan a tomar la decisión de participar. Por ejemplo, un programa de lealtad podría resultar más atractivo para algunos clientes que para otros, dadas sus características, sus comportamientos de compra o los posibles beneficios que podrían obtener a partir del canje de puntos. De esta manera, diferencias en los comportamientos de los clientes podrían estar siendo explicadas por diferencias entre sus características y no por el impacto que haya tenido participar de un programa de lealtad. Este fenómeno se conoce estadísticamente como sesgo por autoselección (Heckman, 1979).

Existen diversos métodos para determinar causalidad bajo condiciones de autoselección. En esta memoria se utilizará un algoritmo de Machine Learning llamado Causal Forest (Athey & Wager, 2018). Este método se utilizará para determinar el impacto que tiene el primer canje de puntos en el comportamiento de compra del cliente. Se escogió el primer canje de puntos dado que es una instancia en la que efectivamente se puede determinar que el cliente participa del programa de lealtad. No así la acumulación de puntos, ya que los clientes acumulan puntos cada vez que consumen, pero consumir no significa necesariamente que se encuentren conscientes del programa o involucrados con él.

Este efecto será medido en dos aplicaciones: una con datos de clientes de una empresa de telecomunicaciones bajo régimen contractual y una con clientes de una firma de retail bajo régimen no contractual.

---

<sup>1</sup> La cuota de cartera (share of wallet) es el porcentaje del gasto que los clientes destinan al consumo en la firma con respecto al total del gasto en tiendas del mismo sector.

## 2. LITERATURA PREVIA

Los clubes de lealtad han sido de interés para distintos investigadores, los cuales han medido de diversas maneras el efecto que tienen dichos clubes en el comportamiento del cliente, sin embargo, existen diversas conclusiones disímiles entre sí. Por ejemplo, Sharp & Sharp (1997) realizaron una investigación enfocada en estimar la capacidad que tienen los clubes de lealtad para alterar los patrones de recompra de los clientes, usando el modelo de Dirichlet. Este modelo entrega una línea base sobre la cual es posible medir los cambios que posiblemente pueden deberse al impacto del club de lealtad. Los autores estudiaron 6 marcas en su investigación, de las cuales solamente 2 mostraron efectos significativos, sin embargo, los mismos efectos fueron observados en clientes que no pertenecían al club de lealtad, por lo que el efecto podría no deberse al club.

Como se mencionó en la introducción, existe un sesgo por autoselección que debe tenerse en cuenta para estimar el efecto que tiene un club de lealtad en el comportamiento de los clientes. Un estudio que considera este sesgo es el de Leenheer et. al. (2007), en el cual estudiaron el efecto de pertenecer a un club de lealtad en la decisión de pertenecer al club de lealtad y el efecto que tiene pertenecer a éste en la cuota de cartera de los clientes. A través de variables instrumentales, los investigadores desarrollaron una metodología que tiene en cuenta el sesgo por autoselección y lograron encontrar un efecto positivo y pequeño, significativo estadísticamente, del programa de lealtad en la cuota de cartera de los clientes.

Más allá del simple hecho de pertenecer al club, es posible estudiar el esfuerzo que invierten los clientes en lograr una meta dentro del programa de lealtad. En esta línea, Drèze & Nunes (2011) examinaron el impacto de lograr un objetivo satisfactoriamente en el esfuerzo invertido posteriormente en lograr el mismo objetivo, a través del estudio empírico de un programa de lealtad de una aerolínea. El objetivo en este estudio era pertenecer a una cierta categoría (tier) de la membresía. Este estudio logra demostrar satisfactoriamente, a través de una metodología basada en el modelo de Hazard proporcional, que el esfuerzo invertido en lograr la misma meta es mayor que en la ocasión anterior, si es que el cliente efectivamente logró esta última. En el mismo estudio, los investigadores replicaron los efectos en un estudio de laboratorio, en el cual demuestran que los efectos solamente son significativos cuando el objetivo a cumplir representa una meta desafiante.

En la misma línea, Carboni (2012) estudió el efecto de un programa de lealtad de una cadena de retail chilena en el comportamiento de compra y canje de sus clientes. El autor investigó el impacto del club de lealtad en distintas variables como, por ejemplo: tiempo entre compras sucesivas a medida que los clientes se acercaban al canje de un premio, esfuerzo invertido por los clientes en cada compra al acercarse al canje de un premio y si es que se repetía el patrón de compras del cliente posterior al canje de un premio, entre otras. Utilizando el modelo de Hazard proporcional en tiempo discreto, e incorporando heterogeneidad mediante el uso de clases latentes, el autor pudo observar una aceleración en las compras a medida que los clientes se acercaban al premio, que existía un reinicio parcial del patrón de compra posterior al canje y que el esfuerzo era creciente mientras más se acercaban los clientes al canje. También encontró una relación directamente proporcional entre el esfuerzo invertido y la probabilidad de preferir productos hedónicos<sup>2</sup>. Dado

---

<sup>2</sup> Un producto hedónico es uno que no tiene un fin utilitarista, sino que es comprado solamente por placer. Por ejemplo, un perfume o cigarrillos son productos hedónicos, mientras que una lavadora es un producto con un fin utilitarista.

esto, Carboni pudo concluir que el programa de lealtad estudiado modificaba favorablemente el comportamiento de sus clientes.

Por otra parte, Wang et. al. (2016) investigaron qué pasaba con el esfuerzo invertido en un programa de lealtad si es que no se cumplía la meta. En el experimento realizado, observaron a más de 95.000 clientes del programa de lealtad de una cadena de hoteles, por 8 meses antes del experimento, 8 meses durante el experimento y 8 meses posteriores al experimento. El experimento consistió en ofrecerles a los clientes dentro del grupo de tratamiento puntos de bonificación si cumplían ciertos requisitos del programa durante el período de la promoción. En promedio, la promoción llevó a un aumento en las compras durante el período posterior a la promoción, sin embargo, solamente el 20% de los clientes lograron el objetivo, mientras que el 80% restante lo falló. A través de una metodología basada en el propensity score, los autores examinaron si existían efectos distintos entre los clientes que lograban el objetivo y los que fallaban. Los resultados mostraron que fallar el objetivo reducía significativamente las compras post-promoción.

Por otra parte, es posible estudiar el impacto que tiene el canje de puntos en el comportamiento de compra de los clientes. En esta línea, Montoya & Flores (2019) realizaron un estudio utilizando datos de una tienda por departamentos para estudiar el impacto de una promoción de canje de premios por una combinación de puntos y dinero en el comportamiento de compra y el canje de premios de los clientes. Como resultados, los autores encontraron que la promoción atraía a los clientes menos activos. Además, concluyeron que los clientes que aprovechaban la promoción incrementaban su preferencia por premios hedónicos en comparación a su comportamiento observado anteriormente y que este cambio persistía en el tiempo. Si bien, la promoción aumentaba la cantidad de canjes posteriores, disminuía en variables relacionadas a las compras.

El primer canje de puntos es interesante de estudiar particularmente, ya que es la primera instancia en la que se puede observar que un cliente es consciente del programa de lealtad y está involucrado con él. En esta línea, Ríos (2013) realizó un estudio del impacto del primer canje de puntos para los clientes que participaban del club de lealtad de una empresa de telecomunicaciones chilena, los cuales se encuentran bajo régimen contractual. En su estudio, Ríos segmentó a los clientes en 6 grupos de acuerdo con sus características individuales, con el fin de medir el efecto en cada grupo, a través de una metodología basada en el emparejamiento en base al propensity score. Como resultados, obtuvo que los clientes que realizaron su primer canje de puntos en un determinado período de tiempo gastaron mensualmente \$3.785 [CLP] más que los clientes que no realizaron ningún canje de puntos, sin embargo, el efecto fue significativo sólo en 4 de esos segmentos. Al analizar los segmentos en los cuales el efecto fue significativo, el impacto sube a \$4.469 [CLP] pesos más que los clientes que no realizaron un canje. Ríos también midió el efecto que tenía el primer canje de puntos en el involucramiento de los clientes, medido a través de los productos que compraban mensualmente en promedio. El efecto fue significativo en 2 de los 6 segmentos y promedió 0,03 productos adicionales en comparación a los clientes que no realizaron un canje. Al analizar el involucramiento mediante la cantidad de productos distintos que compraba cada uno, el efecto fue significativo en sólo un segmento, promediando 0,27 productos distintos más que las personas que no realizaron su primer canje. Finalmente, Ríos no encontró efectos significativos en la permanencia de los clientes en la empresa.

Troncoso (2016) también realizó un estudio del impacto del primer canje de puntos en el comportamiento de los miembros de un programa de fidelización, pero esta vez de una empresa de retail, con clientes bajo régimen no contractual. En particular, Troncoso estudió el impacto en relación con sus decisiones de compra y canje, durante la permanencia de los clientes en el

programa, midiendo su incidencia y monto de compra. La investigación consistió en 3 partes: la primera fue estudiar la incidencia y el monto de compra de los clientes durante la trayectoria hacia el canje, mediante un modelo Tobit II. En segundo lugar, estudió la preferencia por productos hedónicos (asociados al lujo o placer), según la cantidad de puntos que requiere el canje. Finalmente, mediante métodos de matching y con una metodología muy similar a la de Ríos (2013), la investigadora estimó el efecto del primer canje de puntos en el comportamiento de compra del cliente, midiendo a nivel agregado variaciones en incidencia y en monto de compra. La investigadora pudo concluir que el primer canje de puntos genera un aumento de \$5.300 [CLP] en el gasto promedio mensual de los clientes, 3 meses posterior al canje.

Las investigaciones mencionadas se basan en distintas metodologías para hacer frente al sesgo por autoselección. En esta memoria se propone la metodología de los bosques causales (Athey & Wager, 2018), ya que se tiene evidencia de que logra resultados más robustos que otras metodologías. De hecho, Athey & Wager (2019) utilizaron los bosques causales para estimar el impacto de una intervención del National Study of Learning Mindsets en escuelas secundarias públicas de Estados Unidos, en la cual se busca inculcar una mentalidad de crecimiento positiva (growth mindset) en los estudiantes. Si bien el estudio educacional que interviene las escuelas selecciona a éstas aleatoriamente, los autores tuvieron evidencia suficiente como para tratar al estudio como el experimento como uno observacional. En este estudio los autores demuestran que a través de los bosques causales se logra una estimación del efecto de la intervención más robusta a las consecuencias que tiene el sesgo por autoselección que con otras metodologías basadas en el propensity score, en particular, en experimentos en donde el espacio de variables observables es relativamente grande y los métodos tradicionales pueden caer en el sobreajuste<sup>3</sup>.

---

<sup>3</sup> El sobre-ajuste ocurre cuando un algoritmo de aprendizaje queda ajustado a características muy específicas de los datos de entrenamiento, que no tienen relación causal con la función objetivo. En experimentos con un espacio de covariables relativamente grande, a diferencia de metodologías como la regresión lineal, el algoritmo de los bosques causales logra evitar el sobre-ajuste, ya que no toma en cuenta las variables que no tienen mucha relación con la variable de estudio.

## 3. OBJETIVOS

### 3.1. OBJETIVO GENERAL

Determinar el efecto que tiene el primer canje de puntos en el comportamiento de compra de un cliente.

### 3.2. OBJETIVOS ESPECÍFICOS

En línea con el objetivo general, se plantean los siguientes objetivos específicos.

1. Desarrollar una nueva metodología basada en Causal Forest para estimar el efecto de un club de fidelización.
2. Estimar el efecto del primer canje de puntos en el comportamiento de clientes bajo régimen contractual y no contractual.
3. Realizar una comparación de la estimación de este efecto con una obtenida con metodologías alternativas.
4. Caracterizar el efecto a nivel individual.

## 4. ALCANCES

Este trabajo de investigación tiene los siguientes alcances y limitaciones.

1. Sólo se estudiarán clientes que son personas naturales. Quedarán fuera empresas y otros tipos de clientes.
2. Cada uno de los estudios cuenta con pocas variables disponibles, por lo que, tanto la decisión de canjear puntos como la heterogeneidad del efecto podría estar determinada por una variable que no se controla en ninguna de las dos aplicaciones. Esto es una limitación muy importante para tener en cuenta en este estudio.
3. Sólo se estudiará el efecto del primer canje en el comportamiento de compra. No se estudiará la actitud que tiene el cliente frente a la empresa, por lo que el impacto estimado quedará restringido netamente a lo que es lealtad conductual.

# 5. MARCO TEÓRICO

## 5.1. CAUSALIDAD

Muchos trabajos empíricos en ciencias sociales buscan responder preguntas de causa y efecto. ¿X causa Y? Si la respuesta es sí, ¿qué tan grande es el efecto que tiene X sobre Y? Formalmente, el efecto causal que tiene una intervención (o tratamiento) en un individuo es la diferencia entre su respuesta potencial si fue sometido a esta intervención y la respuesta potencial si es que no fue sometido a ella. Naturalmente, sólo podemos observar uno de estos dos estados posibles, lo cual representa el mayor problema de la inferencia causal. Sin embargo, se ha desarrollado por más de 30 años un modelo para determinar causalidad, llamado modelo contrafactual (Morgan & Winship, 2007), con el fin de poder responder este tipo de preguntas.

El modelo contrafactual se compone de varias partes. En primer lugar, se contempla una variable binaria que indica si un individuo fue expuesto a la causa definida (tratamiento) o no. Esta variable es conocida como ‘variable de tratamiento’. Los individuos que son expuestos a la causa son clasificados como pertenecientes al ‘grupo de tratamiento’, mientras que los que no, son asignados al ‘grupo de control’. También, contempla una variable de respuesta que puede estar asociada a uno de los dos estados de la variable de tratamiento. Es decir, podemos observar la respuesta del individuo si es que fue tratado o si es que no fue tratado.

Entonces, una manera de estimar el efecto que tiene la causa, a través de este modelo, es estimar el efecto promedio del tratamiento – o Average Treatment Effect (ATE) – Este efecto se describe a continuación.

$$\tau = E[Y^{(1)} - Y^{(0)}]$$

Donde  $\tau$  es el efecto promedio del tratamiento,  $Y^{(1)}$  es la respuesta potencial cuando el cliente es tratado y  $Y^{(0)}$  es la respuesta potencial cuando el cliente no es tratado.

Como se mencionó anteriormente, el gran problema de este modelo es que, al estar tratando con observaciones de los individuos, sólo se tiene a disposición una de las dos respuestas de cada individuo. Como es imposible observar ambas respuestas para un mismo individuo, la atención es fijada a nivel agregado. Si los individuos que recibieron el tratamiento fueron asignados de manera aleatoria, un buen estimador del efecto queda representado de la siguiente manera.

$$\tau = E[Y^{(1)}] - E[Y^{(0)}]$$

Donde las esperanzas quedan determinadas por las medias de cada grupo. Los individuos del grupo de control son utilizados como contrafactual, es decir, son la representación de cómo se hubiesen comportado los individuos del grupo de tratamiento si no hubiesen sido sometidos a la intervención. Además, es posible estimar el efecto para distintas submuestras de los individuos, por ejemplo, el efecto promedio del tratamiento sobre el grupo que recibe el tratamiento – Average Treatment Effect on the treated (ATT) – o también en el grupo de control.

La estimación del efecto del tratamiento utilizando este modelo contrafactual tiene sentido cuando el tratamiento fue asignado aleatoriamente, dado que la asignación no depende sistemáticamente de las características de la unidad de análisis. Debido a eso, la aleatoriedad permite asegurar que la diferencia entre ambos grupos depende únicamente de la aplicación de la intervención. No obstante, en los estudios observacionales, los cuales se definen como “investigaciones empíricas de tratamientos, políticas o exposiciones y los efectos que causan, en los cuales el investigador no puede controlar la asignación del tratamiento a los individuos” (Rosenbaum, 2002), la estimación de este efecto no es tan directa. Esto se debe a que es posible que la asignación del tratamiento dependa de otras variables del individuo y se esté produciendo, como se mencionó anteriormente, un sesgo por autoselección. Cuando la asignación del tratamiento se puede ver afectada por otras variables, es posible hablar de *confounding* (Rosenbaum & Rubin, 1983). No obstante, tomando algunos supuestos, es posible concluir causalidad en situaciones en las que existe *confounding*.

Para determinar causalidad utilizando el modelo contrafactual bajo condiciones de *confounding*, primero que todo, es necesario definir una variante del efecto promedio del tratamiento, condicionado por las variables observables de un individuo: El Conditional Average Treatment Effect (CATE). Este efecto se expresa de la siguiente manera.

$$\tau(x) = E[Y^{(1)} - Y^{(0)} | X_i = x]$$

Donde  $\tau(x)$  es el efecto promedio del tratamiento para un individuo con un vector de características  $x$ .

Es posible estimar el CATE en un estudio observacional siempre y cuando se asuma un supuesto sumamente fuerte, llamado *unconfoundedness* (Rosenbaum & Rubin, 1983). Este supuesto se formaliza como supuesto de independencia condicional de la siguiente manera.

$$Y_i(1), Y_i(0) \perp W_i | X_i$$

El supuesto de *unconfoundedness* implica que la asignación del tratamiento  $W_i$  se considere aleatoria dentro de cada submuestra de individuos, indexados por sus características  $X_i = x$ . En otras palabras, se asume que cuando se tiene conocimiento de todas las características de un individuo que determinan su tratamiento, la asignación de éste no nos entrega información adicional acerca de él.

Dado esto, si se considera la linealidad de las esperanzas, el estimador del CATE expuesto más arriba puede quedar expresado de la siguiente manera.

$$\tau(x) = E[Y^{(1)} | X_i = x] - E[Y^{(0)} | X_i = x]$$

Finalmente, tomando el supuesto de *unconfoundedness*, el CATE puede ser expresado mediante la siguiente ecuación.

$$\tau(x) = E[Y^{(1)} | W_i = 1, X_i = x] - E[Y^{(0)} | W_i = 0, X_i = x]$$

Por ende, tomando en cuenta el CATE y el supuesto de *unconfoundedness*, es posible estimar el efecto promedio del tratamiento mediante metodologías que sean capaces de encontrar individuos similares en base a sus características. Algunas de estas metodologías se basan en la estimación de

una variable que agrupe a todas las demás, llamada *propensity score*. Otras logran estimar el efecto sin tener que pasar por la estimación del propensity score, como el algoritmo Causal Forest.

## 5.2. PROPENSITY SCORE

El propensity score se define como la probabilidad de que un individuo reciba el tratamiento, condicional a sus características observables:  $e_i = \Pr(W_i = 1 | X_i)$  (Rosenbaum & Rubin, 1983). Este puntaje también es un puntaje de balance, ya que, condicional al propensity score, la distribución de las características observables es similar entre los individuos tratados y no tratados.

El propensity score existe tanto en estudios aleatorizados como observacionales. En los primeros, es definido por el estudio y, por lo tanto, conocido. En cambio, en los estudios observacionales no se conoce, sin embargo, puede ser estimado usando los datos disponibles, generalmente mediante regresión logística.

Los métodos basados en el propensity score reducen todas las variables que describen a un individuo a solamente una (el propensity score) y la utilizan para comparar sujetos que tengan valores similares de esta variable. Tomando el supuesto de unconfoundedness, es posible asumir que para individuos con propensity score similar, la asignación del tratamiento es aleatoria entre ellos. De esta manera se puede estimar el efecto promedio del tratamiento en los tratados – Average Treatment Effect on the treated (ATT) – usando como contrafactual los individuos que son similares a ellos en base al propensity score.

Existen distintos métodos utilizados para estimar causalidad usando el propensity score. Este estudio utiliza como metodologías alternativas dos de ellos: *Propensity Score Matching* e *Inverse Propensity Score Weighting*.

El método de Propensity Score Matching consiste en asociar sujetos que recibieron el tratamiento con sujetos que no, y que comparten un propensity score similar (Rosenbaum & Rubin, 1983). Un sujeto se puede emparejar con un sujeto que no recibió el tratamiento (matching 1:1), o bien se puede asociar a más de uno (matching 1:N). Mediante esta metodología es posible estimar el Average Treatment Effect on the Treated (ATT). La implementación más común de esta metodología realiza la estimación del propensity score mediante una regresión logística, mientras que el emparejamiento tiene un carácter de uno-a-uno, generalmente emparejándose un individuo tratado con el no tratado que tenga el propensity score más parecido (nearest neighbor matching (Rosenbaum & Rubin, 1985). Una vez que el emparejamiento ya fue realizado, el efecto se puede calcular simplemente realizando la diferencia entre el promedio de la respuesta de los individuos tratados con el promedio de la respuesta de sus pares del grupo de control.

Al realizar un emparejamiento según similitudes en los propensity scores de los individuos, es posible que queden individuos del grupo de control que no sean emparejados con alguno del grupo de tratamiento, teniendo como consecuencia el desperdicio de una cantidad de datos no menor. El método Inverse Propensity Score Weighting (Rosenbaum, 1987) deja de lado el efecto secundario mencionado anteriormente, al considerar a todas las observaciones, pero otorgándoles un peso en función de su propensity score.



Al igual que con el método propensity score matching, con el método inverse propensity score weighting se estima un propensity score para cada individuo en función de sus variables descriptivas. La diferencia es que, en vez de emparejar, ahora se utilizan todas las observaciones, cada una con un peso determinado. Las observaciones que recibieron el tratamiento tienen un peso de  $\frac{1}{p}$  y las del grupo de control tienen un peso de  $\frac{1}{1-p}$ , donde  $p$  es el propensity score de cada individuo. En palabras, cada observación es ponderada según el inverso de la probabilidad de ser asignados al grupo que fueron asignados, por lo que, intuitivamente, los individuos tratados que tenían alta probabilidad de no ser tratados y que sí lo fueron tienen mayor relevancia, así como los individuos que no fueron tratados y que tenían un propensity score alto.

## 5.3. CAUSAL FOREST

Muchos métodos de inferencia causal se basan en estimar el propensity score de cada individuo y, a través de él, poder estimar el efecto promedio del tratamiento. Sin embargo, se ha desarrollado un algoritmo consistente, basado en árboles de decisión (Breiman, 2001), llamado Causal Forest (Athey & Wager, 2018), el cual permite construir un estimador no paramétrico<sup>4</sup> del efecto del tratamiento, sin la necesidad de estimar la propensión a recibir éste.

En primer lugar, se define un modelo basado en árboles como un modelo de decisiones secuenciales y jerárquicas que tiene por objetivo predecir el valor de una variable de destino en función de diversas variables de entrada. Si la variable de destino puede tomar valores finitos, se trata de un árbol de clasificación, mientras que, si la variable de destino toma valores continuos, se trata de un árbol de regresión.

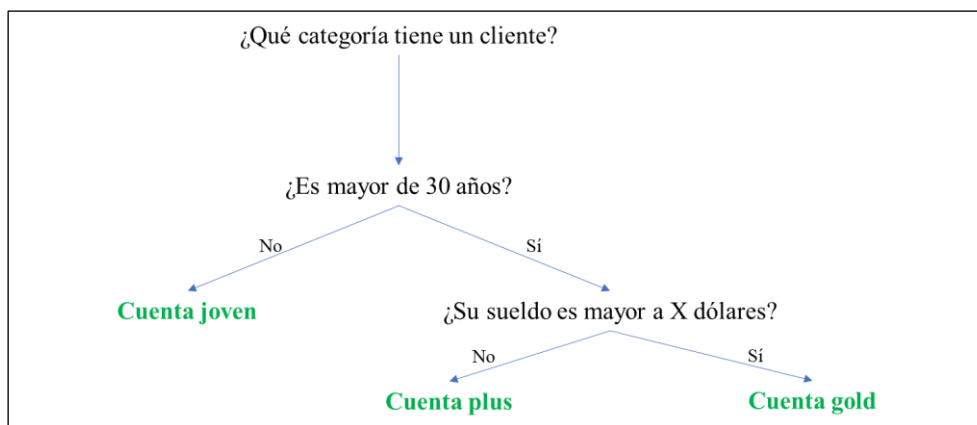
Los modelos de árboles están compuestos por nodos interiores, en los cuales se debe tomar una decisión. Dependiendo de esa decisión, se va avanzando a otros nodos interiores en distintos niveles para, finalmente, llegar a una hoja, la cual clasifica o entrega el valor de la variable de salida. Cada vez que se realiza una decisión y siguen dos caminos posibles, se dice que se realiza un *split* en el árbol. Estos árboles de decisión son unas de las técnicas de aprendizaje supervisado<sup>5</sup> más usadas actualmente en Machine Learning. A continuación, se presenta una figura con un ejemplo de un árbol de clasificación.

---

<sup>4</sup> La estadística no paramétrica es una rama de la estadística que no se basa solamente en familias parametrizadas de distribuciones probabilísticas (media y varianza son ejemplos de parámetros). Estos modelos pueden no tener una distribución o tener una con parámetros no especificados.

<sup>5</sup> En Machine Learning, el aprendizaje supervisado es una técnica para deducir una función a partir de datos de entrenamiento, dentro de los cuales se encuentra la variable de destino de cada individuo. A diferencia de esta técnica, en el aprendizaje no supervisado, no se cuenta con la variable de destino de los individuos.

Ilustración 1: Ejemplo de árbol de clasificación



En el ejemplo de la ilustración 1, se busca clasificar a un cliente en función de dos datos: su edad y su sueldo. Las preguntas corresponden a los nodos en los cuales se realizan los splits, mientras que las clasificaciones (en verde) Cuenta joven, Cuenta plus y Cuenta gold son los nodos terminales, o también llamados hojas, los cuales entregan la variable de salida, en este caso la clasificación del cliente.

El método de los árboles causales (Athey & Wager, 2018) se vale de estos algoritmos basados en árboles para poder determinar causalidad en un experimento bajo condiciones de confounding. Para ello, en primer lugar, se construye un árbol, con el fin de clasificar a los individuos dadas sus características (covariables observables). Estos árboles quedan con un conjunto de  $L$  hojas, dentro de las cuales existe un mínimo de individuos que recibieron y que no recibieron el tratamiento. En este caso, los individuos similares son los que caen dentro de la misma hoja, a diferencia de los métodos basados en el propensity score, en los cuales los individuos similares son aquellos con el propensity score más cercano. El efecto del tratamiento promedio condicional (CATE) dentro de cada hoja queda estimado por la diferencia que existen entre las medias de la variable a estudiar de los individuos que recibieron el tratamiento y los que no, dentro de esa hoja. Así, el efecto del tratamiento condicional para un individuo dentro de una hoja en un árbol causal queda expresado por la siguiente fórmula.

$$\hat{\tau}(x) = \frac{1}{|\{i: W_i = 1, X_i \in L\}|} \sum_{\{i: W_i=1, X_i \in L\}} Y_i - \frac{1}{|\{i: W_i = 0, X_i \in L\}|} \sum_{\{i: W_i=0, X_i \in L\}} Y_i$$

Donde  $\hat{\tau}(x)$  es el estimador del efecto causal buscado condicional a las características  $X = x$  de un individuo y  $L$  es la hoja en la que cae dicho individuo.

Es importante destacar que estos árboles deben cumplir una condición sumamente importante llamada *Honestidad*, con el fin de reducir el sesgo producido cuando se utiliza la variable dependiente para construir un árbol en un submuestreo tradicional. Un árbol es honesto si utiliza las respuestas de los individuos  $Y_i$  para realizar la construcción del árbol o para estimar el efecto promedio del tratamiento, pero no para ambas (Athey & Wager, 2018).

Existen dos algoritmos de árboles causales que satisfacen la condición de honestidad: double-sample trees y propensity trees (Athey & Wager, 2018). En este estudio se utilizan bosques basados

en double-sample trees para la estimación del efecto. En la ilustración 2, se presenta, a grandes rasgos, cómo logran satisfacer la condición de honestidad los double-sample trees.

## Ilustración 2: Algoritmo Double-Sample Trees

### Procedimiento: Double-Sample Trees

Los Double-Sample trees dividen la muestra de entrenamiento en dos submuestras: una para construir el árbol (fijar los splits) y otra para estimar el efecto dentro de cada hoja.

Entrada del modelo:  $n$  ejemplos de entrenamiento de la forma  $(X_i, Y_i, W_i)$ , donde  $X_i$  son las variables características de cada individuo,  $Y_i$  es la variable en la que se quiere medir el efecto y  $W_i$  la asignación del tratamiento. Además, se debe definir un tamaño mínimo de las hojas  $k$ .

1. Obtener una submuestra aleatoria de tamaño  $s$  de  $\{1, \dots, n\}$  sin reposición y luego dividirla en dos subconjuntos disjuntos de tamaño  $|I| = \lfloor s/2 \rfloor$  y  $|J| = \lceil s/2 \rceil$ .
2. Construir un árbol vía partición recursiva. Estos árboles son construidos tomando en cuenta cualquier variable de la submuestra  $J$  y solamente las observaciones  $X$  o  $W$  de la muestra  $I$ , pero sin utilizar las variables  $Y$  de la muestra  $I$ .
3. Asignar a los clientes de la submuestra  $I$  en las hojas correspondientes y estimar el efecto del tratamiento dentro de cada hoja, mediante la metodología antes mencionada.

Siguiendo lo expuesto por Athey & Imbens (2016), los splits de los árboles son escogidos maximizando la varianza de  $\hat{\tau}(x)$  para todo  $i \in J$ . En el estudio de Athey & Wager (2018) es posible encontrar más información acerca de la construcción de los árboles y el proceso de splitting.

Bajo la condición de honestidad y unconfoundedness, es posible construir muchos árboles para estimar el efecto del tratamiento, pero no se puede saber a ciencia cierta cuál es el mejor árbol. Dado esto, como menciona Breiman (2001), es más conveniente construir muchos árboles y promediar sus resultados que construir un solo árbol altamente optimizado. Por lo tanto, el efecto condicional de un bosque causal quedaría expresado de la siguiente manera.

$$\hat{\tau}(x) = B^{-1} \sum_{b=1}^B \hat{\tau}_b(x)$$

Donde  $\hat{\tau}_b$  es el efecto del tratamiento calculado con un árbol  $b$  y  $B$  es la cantidad de árboles del bosque.

Así como es posible estimar un efecto para cada cliente, condicionando por sus variables observables, también es posible estimar, a través de bosques causales, un efecto promedio del tratamiento para toda la muestra (Chernozhukov et. al., 2018a). Las diferencias entre el efecto promedio y el efecto condicional del tratamiento pueden dar indicios sobre la heterogeneidad de dicho efecto.

Para llevar a cabo la implementación de este modelo, se utilizó el paquete de R llamado *grf* desarrollado por Tibshirani et. al. (2018). Este paquete incluye una función para estimar el ATE

basada en una variación de la metodología Augmented Inverse-Propensity Weighting de Robins, Rotnitzky & Zhao (1994), la cual se caracteriza por ser una estimación doblemente robusta. Esta estimación se puede realizar para todos los individuos (ATE), para los individuos tratados solamente (ATT) o para los individuos pertenecientes al grupo de control. En este estudio se decide estimar el efecto promedio del tratamiento en los tratados, ya que lo que se quiere estudiar es precisamente qué impacto tiene el primer canje de puntos en los individuos que efectivamente canjean.

Por otra parte, dentro de los hiperparámetros<sup>6</sup> que se deben definir en un bosque causal, los más importantes son: el número de árboles a construir en el bosque; la fracción del total de datos que serán utilizados para construir cada árbol; la proporción de honestidad, es decir, la proporción de variables de salida que son utilizadas para decidir los splits y las que son utilizadas para estimar el efecto; y la cantidad de variables independientes que son utilizadas para construir cada árbol (si este número es menor que el total de variables independientes, éstas se seleccionan aleatoriamente para cada árbol). El algoritmo permite optimizar estos parámetros mediante validación cruzada o seleccionarlos manualmente.

Para evaluar el desempeño del modelo, existe un método, motivado por la estimación del mejor predictor lineal (best linear predictor o BLP) de Chernozhukov et. al. (2018b), que busca ajustar el CATE como una función lineal de las predicciones del bosque causal. Este ajuste permite evaluar las estimaciones realizadas por el modelo y también realizar una prueba estadística sobre la capacidad de detección de un efecto heterogéneo que tiene el modelo. El mejor predictor lineal queda determinado por la siguiente expresión.

$$Y_i - \hat{m}^{(-i)}(X_i) = \beta_1 \tau(W_i - \hat{e}^{(-i)}(X_i)) + \beta_2 (\hat{\tau}^{(-i)}(X_i) - \tau)(W_i - \hat{e}^{(-i)}(X_i))$$

Los parámetros de este modelo son estimados mediante el método de los mínimos cuadrados ordinarios (OLS) para los parámetros  $\beta_1$  y  $\beta_2$ . A partir de sus resultados, es posible realizar las siguientes interpretaciones: El parámetro  $\beta_1$  habla sobre el desempeño que tiene el modelo para detectar un ATE para la muestra (mientras más cercano a 1 sea, las estimaciones del efecto promedio del tratamiento serán mejores). Por otro lado, el parámetro  $\beta_2$  habla de la capacidad que tiene el modelo para detectar heterogeneidad en el efecto del tratamiento. Si el coeficiente  $\beta_2$  se acerca a 1, las estimaciones heterogéneas están bien calibradas, mientras que, si el coeficiente es positivo y significativo, hay evidencia de que vale la pena estudiar la heterogeneidad dentro del modelo. Si el parámetro  $\beta_2$  no resulta significativo, no implica necesariamente que no exista un efecto heterogéneo en la población, sino que el modelo, con los datos disponibles que trabaja, no es capaz de detectar dicha heterogeneidad.

Finalmente, es posible estimar la importancia relativa de cada covariable en la construcción de un bosque causal. Esta importancia relativa se estima tomando el porcentaje de splits en los que participa dicha variable del total de splits realizados en una profundidad dada del árbol. Esta importancia para cada profundidad del bosque es ponderada por el nivel de profundidad, siendo más importantes los splits de los primeros niveles del árbol que los que se encuentran más cercanos a las hojas.

---

<sup>6</sup> En Machine Learning, los hiperparámetros son parámetros que se fijan antes de que el proceso de aprendizaje comience, a diferencia de los parámetros, que se obtienen a través del entrenamiento de un modelo.

A continuación, se presenta la expresión para el cálculo de la importancia relativa estimada, en donde  $K$  representa el máximo de niveles de profundidad que se quieren considerar. En este trabajo se consideró un máximo de 4 niveles.

$$imp(x_j) = \frac{\sum_{k=1}^K \left[ \frac{\sum_{\text{todos los árboles}} \text{número de splits en nivel } k \text{ por } x_j}{\sum_{\text{todos los árboles}} \text{número total de splits en nivel } k} \right] k^{-2}}{\sum_{k=1}^K k^{-2}}$$

# 6. DESARROLLO METODOLÓGICO

## 6.1. APLICACIÓN N° 1: RÉGIMEN CONTRACTUAL

La primera aplicación de este estudio se lleva a cabo con datos de una empresa de telecomunicaciones que opera en Chile y que presta servicios de telefonía, internet y televisión, tanto móviles como de hogar. Esta empresa cuenta con un programa de lealtad en el cual los clientes que cumplen una cantidad de requisitos pueden acumular puntos, los cuales tienen un tiempo determinado de vigencia. Los premios que pueden ser canjeados por estos puntos pueden ser entradas al cine, productos dentro de la empresa o donaciones a organizaciones de caridad, entre otros. También, al pertenecer al club de lealtad, los clientes pueden acceder a descuentos en eventos como recitales, o en determinados cines.

Esta primera aplicación se basa en el estudio de Ríos (2013). Sin embargo, en este caso se realiza la estimación del efecto del primer canje de puntos mediante la metodología de los bosques causales para dos variables de comportamiento: el monto mensual gastado en la empresa y la cantidad de productos comprados mensualmente en la empresa. Además, se realiza una caracterización individual de los clientes con efecto significativo, con el fin de analizar qué características de los clientes pueden incidir en la magnitud del efecto que tiene el primer canje de puntos en el comportamiento de compra.

### 6.1.1. DATOS DISPONIBLES

La base de datos cuenta con información demográfica de los clientes, como su identificación, región, comuna, ciudad, edad y género. También, contiene información transaccional de los mismos, agregada a nivel mensual como, por ejemplo, su ARPU<sup>7</sup> mensual y los productos que consumió mensualmente.

Por otra parte, la base cuenta con información de los canjes de puntos que ha realizado cada cliente, como el mes en que realizó el canje, los puntos que utilizó en él, la descripción del canje y su categoría. Para el estudio se consideraron sólo los clientes que ingresaron en el año 2011, estudiando su comportamiento entre los meses de septiembre del 2011 y enero del 2013.

Para este estudio no se cuenta con la información sobre la cantidad de puntos acumulados de los clientes ni la antigüedad de ellos. El hecho de no tener estas variables para controlar, junto con muchas otras más que podrían tenerse, se considera una limitación fuerte, debido a que el tratamiento podría estar determinado por alguna de esas variables. Esto afecta fuertemente al supuesto de unconfoundedness.

---

<sup>7</sup> El ARPU (Average Revenue per User) es una medida usada principalmente por las empresas de telecomunicaciones y que mide el ingreso promedio por usuario.

## 6.1.2. ASIGNACIÓN DE GRUPOS EXPERIMENTALES Y VARIABLES A ESTUDIAR

En primer lugar, el tratamiento se determinó por una ventana de tiempo. Todos los clientes que hubiesen realizado al menos un canje en el período de tratamiento quedan considerados como tratados. La ventana de tiempo correspondiente al período de tratamiento fueron los meses que acumularon la mayor cantidad de canjes (44,74% del total de canjes), es decir, entre marzo del 2012 y julio del 2012. Los clientes que realizaron algún canje fuera de esa ventana de tiempo no fueron considerados en el estudio, mientras que los clientes que no realizaron ningún canje fueron considerados como el grupo de control. El comportamiento de los clientes previo al primer canje quedó determinado como el promedio mensual de los 6 meses anteriores al período de tratamiento, mientras que el comportamiento posterior al canje quedó determinado por los 6 meses siguientes a tal período<sup>8</sup>.

Ilustración 3: Ventana temporal del tratamiento



La mayor parte de los canjes dentro del período de tratamiento se realizó en el mes de mayo, seguido por el mes de julio. Por otro lado, el mes con menos canjes fue abril. En la tabla 1 se puede ver en detalle cómo distribuyen los canjes dentro del período de tratamiento.

Tabla 1: Distribución de los canjes por mes dentro del período de tratamiento para la primera aplicación

Mes	Número de canjes	Porcentaje del total de canjes
Marzo	332	18,73%
Abril	207	11,67%
Mayo	494	27,86%
Junio	306	17,26%
Julio	434	24,48%
<b>Total</b>	<b>1.773</b>	<b>100%</b>

Por lo tanto, en esta aplicación fueron considerados 7.778 clientes, de los cuales 1.457 pertenecen al grupo de tratamiento (algunos canjearon más de una vez dentro del período de tratamiento).

<sup>8</sup> Se probó paralelamente otra división del horizonte de tiempo (la usada en la segunda aplicación) y se llegó a un resultado distinto al obtenido con este criterio. Por lo tanto, se puede decir que la metodología aplicada es sensible al período de tiempo estudiado.

Las variables de comportamiento, es decir, en las cuales se espera ver algún cambio posterior al canje, son dos: el monto gastado mensualmente en la empresa, medido a través del ARPU mensual promedio, y la cantidad de productos o servicios contratados, medida a través de la cantidad promedio de productos adquiridos por mes<sup>9</sup>.

Las variables utilizadas para describir a estos clientes son la zona geográfica del país en donde viven (norte, centro o sur), su edad, el género y las variables de comportamiento de compra antes mencionadas.

Las variables descriptivas, o covariables, son divididas en dos grupos: continuas y categóricas. Las variables categóricas son el género y la zona geográfica (norte, centro y sur) y están codificadas dummy. Es importante mencionar que, para el estudio de cada una de las dos variables de comportamiento posterior al tratamiento, se utilizó también el promedio de éstas durante el período previo al tratamiento como covariables explicativas. Por lo tanto, el grupo de variables continuas está compuesto por la edad, el ARPU promedio mensual durante el período previo al tratamiento y el promedio de productos comprados mensualmente durante el período previo al tratamiento. Estas variables continuas fueron escaladas de manera que sigan una distribución normal estándar.

Tabla 2: Tipos de variables de la primera aplicación

Tipo de variable	Variables
Variables categóricas	Género
	Zona geográfica
Variables continuas	Edad
	ARPU mensual
	Productos comprados por mes

El tratamiento quedó definido por una variable binaria  $W$  que indica si los clientes fueron tratados o no.

$$W_i \begin{cases} 1 & \text{si el cliente } i \text{ fue tratado} \\ 0 & \text{si el cliente } i \text{ no fue tratado} \end{cases}$$

Antes de estimar el efecto, se realizó un análisis de las variables disponibles en el estudio. Con respecto al género, los clientes están bien balanceados, existiendo un 51,8% de clientes hombres y un 48,2% de clientas mujeres. En cuanto a la edad, la media es 44 años, mientras que la mediana 43. El cliente más joven tiene 19 años y el más longevo tiene 101 años. Sin embargo, es importante mencionar que el 75% de los clientes tiene menos de 53 años. Con respecto a la región en la cual viven los clientes, 4.507 clientes pertenecen a la zona central del país, es decir, el 57,9% de ellos. El 25,1% pertenecen a la zona sur, mientras que el 17% de ellos pertenece a la zona norte. Los individuos dentro del grupo de tratamiento promediaron un ARPU mensual previo al período de tratamiento de \$48,92 [US\$]<sup>10</sup> y compraron 1,36 productos mensuales en promedio. Mientras que

<sup>9</sup> Se probó agregar las desviaciones estándar del ARPU mensual y de los productos comprados por los clientes mensualmente como variables adicionales, sin embargo, los resultados no variaron, por lo que se decidió dejar solamente los promedios como variables de comportamiento, para las dos aplicaciones.

<sup>10</sup> Los datos originales se encontraban en pesos chilenos. Para convertirlos en montos en dólares, se consideró el promedio del precio del dólar observado mes a mes durante el período del estudio. Este promedio fue de \$488,8 [CLP]. Por simplicidad, se consideró un precio del dólar de \$500 [CLP].



los individuos en el grupo de control promediaron un ARPU mensual de \$52,72 [US\$] y 1,58 productos comprados mensualmente en promedio.

Adicionalmente, en la siguiente tabla se pueden comparar los rasgos generales de cada grupo experimental.

Tabla 3: Caracterización de grupos experimentales según variables disponibles para la primera aplicación

<b>Variable</b>	<b>Promedio grupo de tratamiento</b>	<b>Promedio grupo de control</b>	<b>p-valor test de diferencia medias</b>
Edad promedio [años]	42,16	44,94	<0,01
Promedio de productos comprados mensualmente antes del canje	1,37	1,58	<0,01
ARPU mensual promedio antes del período de canje [US\$]	48,92	52,72	<0,01
Porcentaje de clientes hombres	52,78%	51,59%	0,41
Porcentaje de clientes de la zona norte	17,64%	17,72%	0,41
Porcentaje de clientes de la zona centro	54,15%	58,82%	<0,01
Porcentaje de clientes de la zona sur	28,21%	24,46%	<0,01

De la tabla 3 se desprende que existen diferencias estadísticamente significativas entre los grupos de control y tratamiento para todas las variables, excepto para la proporción de hombres y la proporción de clientes de la zona norte. Sin embargo, en magnitud, estas diferencias no son muy relevantes en la práctica.

### 6.1.3. ESTIMACIÓN DEL EFECTO MEDIANTE CAUSAL FOREST

Se realizó por separado la estimación del efecto del primer canje en el comportamiento del cliente para las dos variables antes mencionadas: ARPU mensual y productos comprados mensualmente en promedio. Los árboles fueron construidos utilizando todas las variables, excepto la variable de comportamiento a estudiar. Se construyeron 2.000 árboles para cada bosque. Para cada árbol se tomó una muestra aleatoria del 50% de los datos para construirlo. De esa muestra, el 50% se utilizó para generar los splits, mientras que el otro 50% fue utilizado para calcular el efecto dentro de cada hoja. Los hiperparámetros de los bosques fueron optimizados mediante validación cruzada. Para el modelo del efecto sobre el ARPU mensual, los árboles se construyeron con 3 covariables cada uno, mientras que el mínimo de individuos pertenecientes a cada grupo en cada hoja fue de 77 clientes.

Por otro lado, para la construcción de los árboles para medir el efecto sobre los productos comprados mensualmente, los árboles fueron construidos utilizando 3 de las covariables, mientras que el mínimo de personas de cada grupo por hoja fue de 39 clientes.

A través de los dos bosques construidos, se estimó que los clientes que realizan su primer canje de puntos gastan en promedio \$6,97 [US\$] más por mes que los que no realizan un canje, mientras que consumen en promedio 0,04 productos más mensualmente. En la tabla 4 se resumen los resultados para ambos casos.

Tabla 4: Resultados de la estimación del efecto mediante Causal Forest para la primera aplicación

Caso	ATT	p-valor	Porcentaje de individuos con CATE significativo
Efecto sobre ARPU mensual [US\$]	6,97	<0,001	100,00%
Efecto sobre productos comprados [productos]	0,04	0,04	19,28%

Con respecto a la importancia relativa de las covariables en la construcción del bosque, la variable más importante en la construcción de ambos bosques fue el ARPU mensual promedio previo al período de canje. En la tabla 5 se presenta un resumen con las tres covariables más importantes de cada caso.

Tabla 5: Importancia relativa de las covariables en la construcción de los bosques de la primera aplicación

Posición	Aplicación: ARPU mensual		Aplicación: Productos mensuales	
	Covariable	Imp. Relativa	Covariable	Imp. Relativa
1	ARPU mensual promedio previo al canje	36,69%	ARPU mensual promedio previo al canje	45,34%
2	Promedio Productos comprados mensualmente antes del canje	22,15%	Edad	23,53%
3	Edad	22,01%	Promedio Productos comprados mensualmente antes del canje	18,71%

Adicionalmente, se realizó la estimación del mejor predictor lineal para ambos modelos, con el fin de probar que éstos realicen buenas estimaciones del efecto y que éste sea heterogéneo. En la tabla 6 se presentan los resultados para el BLP estimado.

Tabla 6: Resultados de la estimación del mejor estimador lineal para la primera aplicación

Parámetro	Estimación para modelo ARPU (p-valor)	Estimación para modelo productos comprados (p-valor)
$\beta_1$	1,012 ( $<0,01$ )	1,11 (0,01)
$\beta_2$	1,029 (0,04)	1,23 (0,03)

De la tabla 6 se puede deducir que, mediante la metodología de bosques causales, es posible encontrar un efecto promedio del primer canje de puntos en el ARPU mensual de los clientes pertenecientes a un club de fidelización, y que este efecto es heterogéneo, por lo que vale la pena hacer una caracterización individual de este efecto. Por otra parte, si bien se obtienen parámetros significativos para el modelo del efecto sobre la cantidad de productos mensuales adicionales comprados posterior al canje, la magnitud del efecto es tan pequeña, que no vale la pena realizar una caracterización de este efecto, puesto que no tiene significancia práctica (una diferencia de 0,04 productos en promedio por mes no parece ser una diferencia relevante).

## 6.1.4. ESTIMACIÓN DEL EFECTO MEDIANTE METODOLOGÍAS ALTERNATIVAS

Con el fin de comparar el resultado obtenido mediante bosques causales, se estimó el mismo efecto en el ARPU, mediante las metodologías de propensity score matching y propensity score weighting.

En primer lugar, se llevó a cabo la estimación del efecto, mediante propensity score matching. El propensity score fue construido mediante una regresión logística, tomando la asignación del tratamiento como variable dependiente, mientras que las variables independientes fueron las covariables antes mencionadas. A continuación, se muestra la fórmula que fue utilizada para explicar la asignación del tratamiento.

$$W \sim \text{Edad} + \text{ARPU}_{\text{antes}} + \text{productos mensuales}_{\text{antes}} + \text{género} + \text{zona\_norte} + \text{zona\_sur}$$

Donde  $W$  corresponde a la asignación del tratamiento. El método utilizado para emparejar a los individuos fue el de los  $K$  vecinos más cercanos, es decir, cada individuo del grupo de tratamiento fue emparejado con los  $K$  individuos del grupo de control que tuvieran el propensity score más parecido. Se hicieron pruebas emparejando con los  $K = 1$  y los  $K = 5$  vecinos más cercanos. El emparejamiento fue con reposición, es decir, los individuos no tratados pudieron ser emparejados con más de un individuo del grupo de tratamiento.

Para el caso de la estimación mediante propensity score weighting, la fórmula para estimar el propensity score fue la misma. El peso utilizado fue el correspondiente para la estimación del

ATT<sup>11</sup>. A continuación, se presenta una tabla resumen comparativa de los efectos estimados mediante las tres metodologías.

Tabla 7: Comparación de las estimaciones del efecto mediante las tres metodologías para la primera aplicación

<b>Modelo</b>	<b>ATT en el ARPU mensual [US\$] (p-valor)</b>	<b>ATT en productos comprados mensuales [productos] (p-valor)</b>
Causal Forest	6,97 (<0,001)	0,04 (0,04)
Propensity Score Weighting	7,28 (<0,001)	0,04 (0,04)
Propensity Score Matching 1:1	5,22 (<0,001)	0,004 (0,90)
Propensity Score Matching 1:5	5,99 (<0,001)	-0,04 (0,11)

En la tabla 7 se puede observar que las estimaciones del efecto obtenidas mediante las 4 metodologías son consistentes. Un efecto significativo y positivo en el monto gastado mensualmente en la empresa y un efecto muy pequeño para la cantidad de productos comprados.

## 6.1.5. CARACTERIZACIÓN INDIVIDUAL DEL EFECTO ESTIMADO MEDIANTE CAUSAL FOREST

El modelo de bosques causales también permite realizar una estimación individual del CATE para cada cliente. Valiéndose de esta ventaja, es posible analizar a los clientes tratados y el efecto que tuvo en cada uno el primer canje, con el fin de estudiar qué características pueden influir en la magnitud del efecto estimado. Se realizó una caracterización individual para el efecto que tiene el primer canje de puntos sobre el ARPU mensual promedio, mientras que para el efecto que tiene sobre los productos mensuales comprados, solamente se caracterizó el grupo de individuos con un CATE significativo.

### 6.1.5.1. Caracterización del efecto en el ARPU mensual

Para realizar la caracterización individual del efecto del primer canje, primero se analizó cómo distribuye el efecto según cada una de las covariables, para analizar si es que existe alguna diferencia significativa por género, zona geográfica o edad.

En primer lugar, se analizó la diferencia entre hombres y mujeres. Los hombres, que son un 52,78% del total de clientes tratados, tuvieron un efecto promedio de \$7,07 [US\$], mientras que las mujeres

---

<sup>11</sup> Dependiendo de la submuestra en la cual se quiere estimar el efecto (tratados, no tratados, total, etc.) se utiliza un peso distinto para cada cliente en la estimación mediante propensity score weighting. En este caso se utilizó la que corresponde a la estimación del efecto en los tratados. Para más información ver Hirano, Imbens & Ridder (2003).

tuvieron un efecto promedio de \$6,62 [US\$], diferencia estadísticamente significativa con un nivel de significancia del 1%. En la tabla 8 se muestra un resumen de las diferencias.

Tabla 8: Diferencias por género del efecto estimado mediante Causal Forest de la primera aplicación

<b>Género</b>	<b>CATE promedio [US\$]</b>	<b>Cantidad de personas</b>	<b>Porcentaje del total</b>
Hombre	7,07	769	52,78%
Mujer	6,62	688	47,22%

Luego, se analizaron las diferencias por zona geográfica de los clientes. Los clientes de la zona norte (18% del total) tuvieron un efecto promedio de \$7,02 [US\$] y los de la zona centro (54% del total) un efecto de \$7,06 [US\$], no habiendo diferencias estadísticamente significativas entre las medias de estos dos grupos. En cambio, los clientes de la zona sur (28% del total) tuvieron un efecto de \$6,38 [US\$], presentando una diferencia de medias estadísticamente significativa con los primeros dos grupos, a un nivel de significancia del 1%. A continuación, se presenta una tabla con las diferencias resumidas.

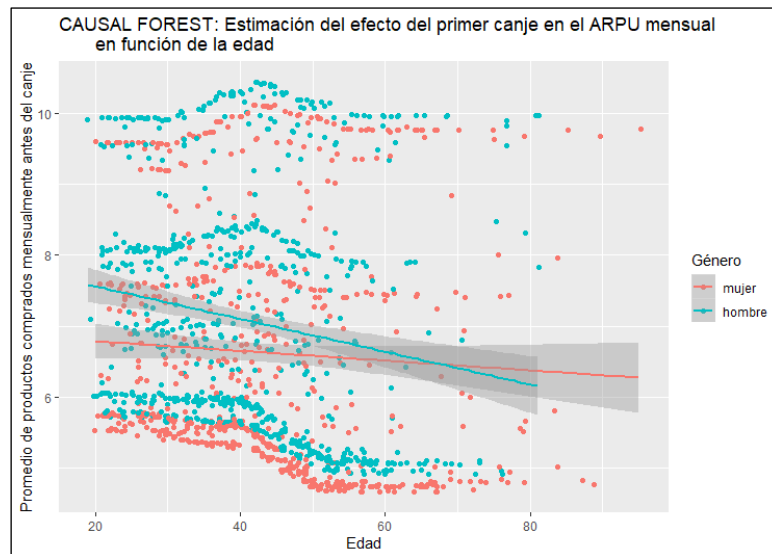
Tabla 9: Diferencias por zona geográfica del efecto estimado mediante Causal Forest de la primera aplicación

<b>Zona geográfica</b>	<b>CATE promedio [US\$]</b>	<b>Cantidad de personas</b>	<b>Porcentaje del total</b>
Norte	7,02	257	17,63%
Centro	7,06	789	54,16%
Sur	6,38	411	28,21%

Posteriormente, se analizó cómo impactaba el comportamiento previo, medido a través de los productos comprados mensualmente antes del canje, al efecto del primer canje en el ARPU mensual. Se realizó una regresión log-log, usando el logaritmo del efecto del primer canje como variable dependiente y el logaritmo de la cantidad de productos comprados mensualmente antes del primer canje como variable independiente. La regresión se realizó para hombres y mujeres por separado. La pendiente de la curva para hombres fue de 0,42, mientras que para mujeres fue de 0,43, ambos parámetros significativos a un nivel del 1%. También se realizó una regresión lineal del efecto del primer canje como variable dependiente en función del género, la cantidad de productos comprados mensualmente antes del canje y el cruce entre ambas variables, resultando significativas las dos primeras, pero no el cruce, por lo que no existen diferencias estadísticamente significativas en el impacto del comportamiento previo en el efecto del primer canje en el ARPU mensual, entre hombres y mujeres.

El análisis del impacto de la edad en el efecto del primer canje en el ARPU mensual se realizó gráficamente. En la siguiente figura se presenta un gráfico de dispersión que muestra en el eje X la edad de los clientes y en el eje Y el efecto del primer canje de puntos en el ARPU mensual.

Ilustración 3: Estimación mediante Causal Forest del efecto del primer canje en el ARPU mensual en función de la edad



En el gráfico es posible ver que para hombres y mujeres a medida que aumenta la edad, el impacto del primer canje en el ARPU mensual es menor, sin embargo, existe un peak para edades cercanas a los 40 años.

Posteriormente, se realizó una regresión lineal del CATE en función de las covariables, con el fin de corroborar si estas variables independientes tenían un efecto significativo en la estimación del impacto del primer canje. Para esta regresión, se utilizó como variable dependiente el logaritmo del efecto del primer canje en el ARPU mensual, mientras que las variables independientes fueron el logaritmo de la edad, el logaritmo del ARPU mensual promedio antes del canje, el logaritmo de la cantidad promedio de productos comprados mensualmente antes del primer canje, una variable dummy de género, una variable dummy que indica si el cliente pertenece a la zona norte y una que indica si el cliente pertenece a la zona sur. Además, se incluyó una variable dummy para identificar a los clientes que tienen una edad entre los 35 y los 45 años, que son los que presentaron el peak gráfico del efecto. En la tabla 10 se presentan los resultados de dicha regresión lineal.

Tabla 10: Resultados de la regresión lineal del CATE estimado mediante Causal Forest en función de las covariables para la primera aplicación

Coefficiente	Estimación	Error estándar	P-valor
Intercepto	1,375	0,037	<0,001
Log(Edad)	-0,115	0,008	<0,001
Rango edad (35-45)	0,040	0,005	<0,001
Log(ARPU promedio mensual antes del primer canje)	0,238	0,005	<0,001
Log(Productos comprados mensuales)	0,197	0,007	<0,001

promedio antes del primer canje)			
Género (hombre)	0,052	0,004	<0,001
Zona norte	0,012	0,006	0,060
Zona sur	-0,033	0,005	<0,001
Número de observaciones	1.457		

El coeficiente de determinación  $R^2$  de la regresión fue 0,848, al igual que el coeficiente de determinación ajustado, es decir, el modelo explica gran parte de la varianza de la estimación del efecto del primer canje en el ARPU mensual.

De los resultados de la regresión lineal, es posible desprender que la magnitud del efecto del primer canje disminuye conforme aumenta la edad del cliente, específicamente, un aumento de un 1% en la edad del cliente implica una disminución de un 0,11% en la magnitud del efecto. Con respecto a las personas que tienen entre 35 y 45 años, éstas presentan una diferencia positiva de un 4% en la magnitud del efecto con respecto a personas de características similares que no se encuentran dentro de ese rango etario. Asimismo, los clientes hombres presentan una diferencia positiva de un 5,2% en la magnitud del efecto en relación a las clientas mujeres de similares características. Al comparar a los clientes de la zona centro y norte con los clientes de la zona sur, presentan una disminución del 3,3% en la magnitud del efecto con respecto a los clientes del resto de las zonas geográficas del país. Finalmente, una variación de un 1% en el ARPU mensual promedio antes del canje significará una variación de un 0,24% en la magnitud del efecto sobre el ARPU mensual después del canje, siendo la variable que más afecta positivamente en el impacto del primer canje, y siendo ese impacto mayor en mujeres que en hombres (ver anexo 1.5). Por otra parte, una variación de un 1% en la cantidad promedio de productos que consumía un cliente antes del canje, significará un aumento de un 0,19% en la magnitud del efecto sobre el ARPU mensual posterior al canje.

Posteriormente, para completar la caracterización, se segmentó a los clientes en 3 grupos: de bajo, mediano y alto impacto del primer canje de puntos en el ARPU mensual. El método utilizado para segmentar fue el de método de clasificación de los cortes naturales de Jenks (Jenks, 1967). Este método, fue diseñado originalmente para ser utilizado en cartografía, fue escogido porque permite realizar una segmentación basada en datos unidimensionales, muy parecida a lo que realiza el método K-means, pero cuyos resultados pueden ser interpretados de mejor manera más al estar optimizado para datos en una sola dimensión.

El método de Jenks recibe como entrada la cantidad de cortes que se desean posicionar en los datos y mediante un algoritmo que maximiza la varianza entre segmentos y minimiza la varianza dentro de los segmentos, entrega los cortes óptimos a realizar. Con el fin de comparar los resultados de esta segmentación, se realizó una segmentación utilizando los mismos datos, pero con el método K-Means, obteniendo la misma segmentación. Dado esto, se confió en el método de Jenks y se procedió a realizar el análisis basado en la segmentación obtenida con él.

Basado en la segmentación realizada, los clientes que tuvieron un efecto del primer canje en el ARPU mensual por debajo de los \$6,53 [US\$] fueron asignados al segmento de bajo impacto (53,53% del total), los clientes que tuvieron un efecto entre \$6,53 [US\$] y \$8,66 [US\$] fueron asignados al segmento de mediano impacto (28,87% del total), mientras que los que presentaron

un efecto por sobre los \$8,66 [US\$] se asignaron al segmento de alto impacto (18,19% del total). A continuación, se presenta una tabla resumen de las características de cada segmento.

Tabla 11: Caracterización de los segmentos de clientes de la primera aplicación

<b>Segmento</b>	<b>Bajo impacto</b>	<b>Mediano impacto</b>	<b>Alto impacto</b>
CATE promedio [US\$]	5,48	7,58	9,79
Tamaño [Personas]	780	412	265
Edad promedio [años]	43,39	38,93	43,53
ARPU mensual promedio antes del primer canje	25,59	59,85	100,60
Promedio de productos comprados mensualmente antes del primer canje	1,00	1,11	2,81
Porcentaje de personas de la zona norte	16,03%	21,36%	16,60%
Porcentaje de personas de la zona centro	49,62%	56,55%	63,78%
Porcentaje de personas de la zona sur	34,35%	22,09%	19,62%
Porcentaje de hombres	47,95%	61,89%	52,83%

De la segmentación, es posible concluir que el segmento de bajo impacto tiene una concentración mayor de personas de la zona sur. Los tres segmentos se encuentran balanceados en cuanto a edad y género, excepto el segmento de mediano impacto, que está compuesto en más de un 60% por hombres. Con respecto a la zona geográfica, el segmento de alto impacto está compuesto en su mayoría por clientes de la zona centro. Las variables de consumo son las que se llevan la mayor correlación con la segmentación, ya que los clientes consumen más del triple que los clientes del segmento de bajo impacto, tanto en monto mensual como en cantidad de productos mensuales. Esto sugiere que el impacto del primer canje de puntos en el consumo de los clientes es mayor en los clientes que consumen más.

Todas las diferencias de medias entre los segmentos de bajo y mediano impacto son significativas a un 5%, a excepción de la proporción de clientes de la zona norte. Para los segmentos de mediano y alto impacto, todas las diferencias de medias son significativas a un 5%, exceptuando el porcentaje de clientes de la zona sur, por lo que se puede decir que efectivamente existe heterogeneidad en el efecto del primer canje de puntos en el ARPU mensual.



### 6.1.5.2. Caracterización del efecto en los productos comprados mensualmente

Dado que el efecto del primer canje en la cantidad de productos comprados mensualmente por los clientes es estadísticamente significativo en menos del 20% de ellos, la caracterización se basó en analizar las variables que definen a este grupo de personas con efecto significativo. Además, como el efecto oscila entre los 0,04 y los 0,11 productos por mes para los clientes con efecto estadísticamente significativo, este efecto no tiene significancia en la práctica, ya que es una fracción muy menor de productos, que no hace sentido y, por consiguiente, no tiene sentido tampoco analizar a qué se puede estar debiendo este pequeño efecto. Por lo tanto, a continuación, se presenta una tabla resumen de las características del segmento de clientes con efecto estadísticamente significativo.

Tabla 12: Caracterización de los clientes con efecto estadísticamente significativo para el segundo caso de la primera aplicación

Variables descriptivas	
Edad promedio [Años]	46,84 años
ARPU promedio mensual antes del primer canje [US\$]	\$74,59 [US\$]
Promedio de productos comprados mensualmente antes del primer canje	1,74
Porcentaje de hombres	57,30%
Porcentaje de clientes de la zona norte	21,35%
Porcentaje de clientes de la zona centro	58,72%
Porcentaje de clientes de la zona sur	19,93%

## 6.1.6. CONCLUSIONES DE LA PRIMERA APLICACIÓN

De esta primera aplicación usando datos de clientes bajo régimen contractual en una empresa de telecomunicaciones, es posible concluir que el primer canje de puntos sí tiene un efecto significativo en el comportamiento del cliente. En particular, este efecto se ve reflejado en el monto que gastan los clientes en la empresa, medido a través del ARPU mensual. Este efecto presenta heterogeneidad: es diferente para hombres y mujeres, siendo significativamente mayor en hombres. Si bien, a medida que aumenta la edad de los clientes, el primer canje suele tener un efecto menor en el ARPU mensual, para edades cercanas a los 40 años existe un aumento del efecto transversal a todos los clientes, sin importar sus demás características. Además, las personas de la zona sur tienen un efecto del primer canje en el monto gastado mensualmente significativamente menor a las personas de la zona norte y centro. Sin dudas la conclusión más importante respecto a la heterogeneidad del efecto del primer canje es que mientras mayor sea el consumo promedio en el período previo al canje de los clientes, mayor será el efecto que tiene éste en su comportamiento de compra.

Por otro lado, el primer canje no parece tener un efecto en la cantidad de productos comprados mensualmente por los clientes, ya que menos del 20% de éstos presenta un efecto estadísticamente significativo en el promedio de productos comprados mensualmente después del canje. Dentro del grupo de clientes con efecto del primer canje significativo en la cantidad de productos, éstos compran en promedio entre 0,04 y 0,11 productos más por mes, lo cual no tiene significancia práctica, ya que las cantidades son fracciones muy pequeñas.

Sin embargo, resulta interesante analizar por qué sucede esto. ¿De qué manera los clientes pueden estar gastando más dinero si no están comprando más productos? Una posible respuesta a esta interrogante es que los clientes que realizan su primer canje de puntos se cambian a un mejor plan de telefonía que el que tienen actualmente, el cual aumenta el monto que abonan mes a mes en la empresa, sin cambiar el número de productos que compran. De hecho, al tener un plan mejor, puede que necesiten comprar menos productos como bolsas de mensajes o datos de navegación, haciendo que la cantidad de productos que consumen mensualmente pueda disminuir.

Por último, es importante mencionar que, debido la metodología utilizada, se pierde el análisis de los meses inmediatamente posterior al canje real de un cliente, dado que se está utilizando una ventana de meses como tratamiento. Esto podría estar generando una subestimación o sobrestimación del efecto. Sin embargo, pese a esta limitación, fue posible encontrar resultados consistentes y significativos estadísticamente para el efecto del primer canje de puntos y la heterogeneidad subyacente en él.

## 6.2. APLICACIÓN N° 2: RÉGIMEN NO CONTRACTUAL

La segunda aplicación de esta investigación estudia el impacto del primer canje de puntos en el comportamiento de los clientes de una tienda por departamentos en Chile. El programa de lealtad de esta empresa permite a los clientes acumular puntos cada vez que utilizan su tarjeta como medio de pago. Los clientes obtienen aproximadamente 3,3 puntos por cada dólar gastado en una compra. Los puntos tienen una vigencia de 12 meses desde su fecha de emisión, luego expiran. Estos puntos pueden ser canjeados por premios, los cuales son productos definidos previamente en un catálogo renovado todos los años. Estos premios están clasificados en 8 niveles, según la cantidad necesaria de puntos para canjearlos. En la tabla 13 se describen los 8 niveles.

Tabla 13: Niveles de canje

<b>Nivel</b>	<b>Puntos</b>
1	5.000
2	9.000
3	12.000
4	24.000
5	48.000
6	60.000
7	90.000
8	120.000

Esta segunda aplicación se basa en el estudio de Troncoso (2016), pero realizando la estimación del primer canje de puntos en el comportamiento del cliente – al igual que en la aplicación anterior – mediante bosques causales. Las variables en las que se estima el efecto que tiene el primer canje de puntos son: monto de compra, frecuencia promedio de compra y meses que pasan desde el primer canje hasta la siguiente compra. Además, al igual que en la primera aplicación, se realiza una caracterización individual del efecto, para analizar qué características de los individuos inciden en la magnitud del efecto en su comportamiento de compra.

### 6.2.1. DATOS DISPONIBLES

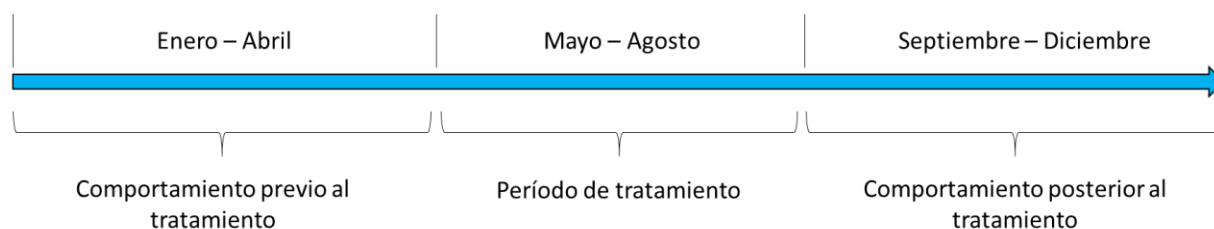
La base de datos construida cuenta con información transaccional de clientes pertenecientes al club de lealtad. Los datos están agregados a nivel mensual, y fueron considerados todos los clientes que tuvieran su información completa para cada año por separado, desde el 2004 al 2006, es decir, clientes que para cada año tuvieran información desde enero a diciembre, sin excepción. Según este criterio, la base de datos quedó conformada por 26.539 clientes pertenecientes al club de lealtad.

La información contiene datos sobre las compras mensuales de los clientes<sup>12</sup>, junto con los puntos que acumularon mes a mes. También se tiene información sobre los puntos que tienen acumulados, los que tienen por expirar y los que expiraron. Por último, se tienen los datos sobre los canjes de los clientes.

## 6.2.2. DETERMINACIÓN DE GRUPOS EXPERIMENTALES Y VARIABLES A ESTUDIAR

Al igual que en la primera aplicación, el tratamiento fue definido por una ventana de tiempo. En este caso, la ventana de tratamiento se definió entre los meses de mayo y agosto, para cada año. Por ejemplo, para el año 2005, todos los clientes que canjearon por primera vez entre los meses de mayo y agosto de ese año fueron asignados al grupo de tratamiento. Los clientes asignados al grupo de control fueron los que no canjearon ese año ni los anteriores, pero que sí pudieron haberlo hecho, es decir, los que en algún momento del período de tratamiento tuvieron más de 5.000 puntos, pero que terminaron el año sin canjes acumulados. Se utilizó el mismo criterio para los 3 años disponibles, para luego combinar las 3 instancias en una gran base de datos. A diferencia del experimento anterior, no se definió la ventana de tratamiento como la que tenía la mayor cantidad de canjes, debido a que la mayor cantidad de canjes se encontraba en diciembre, lo cual podría generar que el efecto del primer canje fuese confundido con el efecto que tiene la navidad en los clientes, período en el que el consumo aumenta notablemente. Se decidió utilizar la ventana de mitad de año debido a que se podría tener información de los clientes antes del canje (los 4 primeros meses del año) y después del canje (los últimos 4 meses del año). Además, utilizar datos de cada año por separado permitía observar mejor a aquellos clientes que dejaban de comprar de un año a otro. En la ilustración 4 se presenta los períodos utilizados para cada ventana temporal. Es importante mencionar que la estacionalidad de la industria del retail no afecta este estudio, ya que se está comparando el comportamiento de compra de los clientes tratados y no tratados para el mismo período de tiempo (septiembre a diciembre) y no una diferencia con respecto a meses pasados. Cabe también destacar que hay algunos clientes que pueden aparecer más de una vez en la base de datos, por ejemplo, un cliente que perteneció al grupo de control en el año 2005 y también en el 2006, aparecerá dos veces, sin embargo, son considerados como clientes distintos ya que sus comportamientos pueden haber sido distintos en cada año.

Ilustración 4: Ventana temporal del tratamiento para la segunda aplicación



<sup>12</sup> Al igual que en la primera aplicación, los datos monetarios se encontraban en pesos chilenos. Análogamente, en esta aplicación se realizó una transformación de la moneda, tomando un valor promedio del dólar de \$600 [CLP].

Además, como el 96,32% de los canjes pertenecen al nivel 1 (5.000 puntos), solamente se consideraron los canjes de ese nivel. De esta manera, la base de datos queda conformada por 10.313 individuos, de los cuales 1.773 pertenecen al grupo de tratamiento, es decir, entre el 17% y el 18% del total.

Con respecto a la distribución de los canjes dentro de la ventana de tratamiento, en esta aplicación los canjes se encuentran mucho más balanceados dentro del período de canje que en la primera aplicación. En la tabla 14 se presenta la distribución de los canjes totales dentro del período de canje, los cuales consideran a individuos dentro y fuera del estudio.

Tabla 14: Distribución de los canjes por mes dentro del período de canje para la segunda aplicación

Mes	Número de canjes	Porcentaje del total de canjes
Mayo	691	24,87%
Junio	698	25,13%
Julio	708	25,49%
Agosto	681	24,51%
<b>Total</b>	<b>2778</b>	<b>100%</b>

Por otra parte, las variables utilizadas para describir a los clientes fueron: su frecuencia promedio de compra, el monto de compra condicional (cuánto gastan mensualmente en promedio cada vez que compran), el gasto promedio mensual (cuánto gastan mensualmente en promedio, independientemente de si compran o no), la recencia de la última compra (es decir, los meses que pasaron entre su última compra del período previo al tratamiento y el período de tratamiento), la recencia de la siguiente compra<sup>13</sup> (es decir, los meses que pasaron entre el período de tratamiento y su primera compra dentro del período post tratamiento), los puntos que tenía acumulados al comenzar el período de canje, los puntos que le expiraron durante el período previo al canje, los puntos que le iban a expirar durante el período de canje y una variable dummy que indicaba si el cliente cumplía el requisito mínimo para canjear (tener acumulados 5.000 puntos) en el mes en el que comenzaba el período de canje. Finalmente, también fue considerado el año en el cual fue observado cada cliente. En la tabla 15 se presenta un resumen de las variables estudiadas.

Tabla 15: Tipos de variables para la segunda aplicación

Tipo de variable	Variables
Variables continuas	Monto de compra condicional
	Gasto promedio mensual
	Frecuencia de compra
	Recencia de última compra
	Recencia de siguiente compra
	Puntos acumulados antes del período de canje

<sup>13</sup> Cuando se mide el efecto en esta variable se puede estar sobrestimando o subestimando el efecto, ya que no considera el mes dentro del período de tratamiento en el que canjeó el cliente tratado, pero como no se tiene un mes exacto de canje para el grupo de control, se optó por comparar el tiempo que pasa entre el tratamiento y la primera compra, considerando el último mes del período de canje.

	Puntos por expirar durante período de canje
	Puntos expirados en el período previo al canje
Variables categóricas	Requisito de canje previo al período
	Año

Al igual que en la primera aplicación, el tratamiento quedó definido por una variable binaria  $W$  que indica si los clientes fueron tratados o no.

$$W_i \begin{cases} 1 & \text{si el cliente } i \text{ fue tratado} \\ 0 & \text{si el cliente } i \text{ no fue tratado} \end{cases}$$

Antes de estimar el efecto, se realizó un análisis de las variables seleccionadas. En relación con los puntos, los clientes promedian 7.255 puntos acumulados al inicio del período de canje, sin embargo, solamente el 66,1% de los clientes llegaba a los 5.000 puntos al comenzar el período de tratamiento. En promedio, los clientes tenían 443 puntos por expirar durante el período de tratamiento, sin embargo, al menos el 50% de los casos no tenía puntos por expirar.

Con respecto a los puntos expirados entre enero y abril, los clientes promediaron 467 puntos expirados, pero, al igual que en el caso de los puntos por expirar, más del 50% de los clientes no tuvo puntos expirados en dicha fecha. En relación con el comportamiento de compra, los clientes presentaron una frecuencia de compra promedio del 56% en los 4 meses previos al período de canje. En cada compra gastaron aproximadamente \$171,7 [US\$] en promedio, por lo que los clientes tuvieron en promedio un gasto mensual aproximado de \$110,7 [US\$]. Con respecto a la recencia de la última compra, al menos el 50% de los clientes había comprado en el mes anterior al período de tratamiento. De igual manera, para la recencia de la primera compra posterior al tratamiento, al menos el 50% de los clientes compró al mes siguiente. En la tabla 16 se pueden apreciar un resumen de estas variables por grupo experimental.

Tabla 16: Caracterización de grupos experimentales según variables disponibles para la segunda aplicación

Variable	Promedio grupo de tratamiento	Promedio grupo de control	p-valor test de diferencia medias
Puntos acumulados al iniciar el período de canje	6.060	7.503	<0,01
Puntos por expirar durante período de canje	508	430	<0,01
Puntos expirados antes del período de canje	765	406	<0,01
Porcentaje de clientes que cumple requisito de canje	40,95%	71,27%	

Frecuencia de compra promedio <sup>14</sup>	72,93%	52,83%	<0,01
Monto promedio de compra [US\$]	152,07	175,78	<0,01
Gasto promedio mensual [US\$]	118,94	109,02	<0,01
Recencia de la última compra antes del período de canje [meses]	1,50	2,36	<0,01

Finalmente, con respecto al año de las observaciones, la base contó con 3.821 clientes del año 2004, 3.935 clientes del año 2005 y 2.557 clientes del año 2006.

### 6.2.3. ESTIMACIÓN DEL EFECTO MEDIANTE CAUSAL FOREST

La estimación del efecto del primer canje de puntos en el comportamiento del cliente se realizó en 4 variables: monto condicional de compra, gasto promedio mensual, frecuencia promedio de compra y recencia de la primera compra posterior al canje. En la tabla 17 se puede ver qué variables fueron utilizadas como covariables independientes para cada caso.

Tabla 17: Covariables utilizadas para cada caso de la segunda aplicación

<b>Caso</b>	<b>Covariables utilizadas</b>
Monto mensual de compra condicional	Puntos acumulados, Puntos expirando, Puntos expirados, Frecuencia promedio de compra, Recencia de última compra antes del canje, Gasto promedio mensual, Condición de requisito mínimo, Año.
Monto mensual de compra incondicional	Puntos acumulados, Puntos expirando, Puntos expirados, Frecuencia promedio de compra, Recencia de última compra antes del canje, Monto de compra mensual condicional, Condición de requisito mínimo, Año.
Frecuencia promedio de compra	Puntos acumulados, Puntos expirando,

<sup>14</sup> La frecuencia promedio de compra se calculó como el porcentaje de los meses del período previo al tratamiento en los que el cliente realizó una compra del total de meses en el período previo al tratamiento.

	Puntos expirados, Frecuencia promedio de compra, Recencia de última compra antes del canje, Monto de compra mensual condicional, Condición de requisito mínimo, Año.
Recencia de primera compra posterior al primer canje	Puntos acumulados, Puntos expirando, Puntos expirados, Frecuencia promedio de compra, Recencia de última compra antes del canje, Monto de compra mensual condicional, Condición de requisito mínimo, Año.

Al igual que en la primera aplicación, cada bosque se construyó con 2.000 árboles, utilizando una muestra aleatoria del 50% del total para construir cada árbol. De estas submuestras, se usó un 50% de la muestra para seleccionar los splits y un 50% para estimar el efecto dentro de cada hoja. Para los casos de estudio, los hiperparámetros de los bosques fueron optimizados mediante validación cruzada. En la tabla 18 se pueden observar los hiperparámetros para cada caso.

Tabla 18: Hiperparámetros para cada caso de la segunda aplicación

Caso	Número de variables utilizadas en cada árbol	Mínimo de personas de cada grupo en cada hoja
Monto mensual de compra condicional	3	105
Gasto promedio mensual	3	48
Frecuencia de compra	6	20
Recencia de primera compra posterior al primer canje	7	8

A través de los dos bosques construidos, se estimó que los clientes que realizan su primer canje de puntos gastan en promedio \$38,6 [US\$] más por compra posterior al canje que los que no realizan un canje, mientras que tienen un gasto promedio mensual 30,57 [US\$]. Además, tienen una frecuencia de compra 8,84% veces mayor que la frecuencia de los que no realizan un canje, y se demoran en promedio una semana menos en realizar su próxima compra posterior al canje. En la tabla 19 se resumen los resultados.



Tabla 19: Resultados de la estimación mediante Causal Forest para la segunda aplicación

Caso	ATT	p-valor	Porcentaje de individuos con CATE significativo
Monto mensual de compra condicional	41,03 [US\$]	<0,001	100,00%
Gasto promedio mensual	31,99 [US\$]	<0,001	100,00%
Frecuencia de compra	8,85%	<0,001	56,91%
Recencia de primera compra posterior al primer canje	-0,22 [meses]	<0,001	35,98%

De la misma manera que en la primera aplicación, se realizó la estimación de la importancia relativa de cada covariable en la construcción de los bosques. En la tabla 20 y 21 se presentan las importancias relativas de las 3 covariables más importantes para todos los casos, en los cuales los puntos acumulados, expirados y por vencer muestran ser las variables más importantes.

Tabla 20: Importancia relativa de las covariables en la construcción de los bosques para el efecto en monto de compra de la segunda aplicación

Posición	Aplicación: <b>Monto condicional</b>		Aplicación: <b>Gasto promedio</b>	
	Covariable	Imp. Relativa	Covariable	Imp. Relativa
1	Puntos acumulados	24,08%	Puntos acumulados	20,77%
2	Puntos expirando	18,05%	Puntos expirados	21,61%
3	Puntos expirados	17,14%	Puntos expirando	18,37%

Tabla 21: Importancia relativa de las covariables en la construcción de los bosques para el efecto en frecuencia y recencia de compra de la segunda aplicación

Posición	Aplicación: <b>Frecuencia de compra</b>		Aplicación: <b>Recencia de compra</b>	
	Covariable	Imp. Relativa	Covariable	Imp. Relativa
1	Puntos expirando	40,39%	Puntos expirando	39,53%
2	Puntos acumulados	28,02%	Puntos acumulados	37,68%
3	Monto condicional de compra	8,87%	Puntos expirados	10,23%

Adicionalmente, se realizó la estimación del mejor predictor lineal para los modelos, con el fin de probar que éstos realicen buenas estimaciones del efecto y que éste sea heterogéneo. En la tabla 22 se presentan los resultados para los BLPs estimados.

Tabla 22: Resultados de la estimación del mejor estimador lineal para la segunda aplicación

Parámetro	Aplicación: Monto condicional de compra (p-valor)	Aplicación: Gasto promedio mensual (p-valor)	Aplicación: Frecuencia promedio de compra	Aplicación: Recencia primera compra post canje
$\beta_1$	0,85 (<0,001)	0,99 (<0,001)	0,99 (<0,001)	0,98 (<0,001)
$\beta_2$	-2,58 (0,95)	0,56 (0,23)	1,33 (<0,001)	1,24 (<0,001)

De la tabla 22 es posible concluir que existe un efecto significativo en las 4 variables estudiadas, sin embargo, no se cuenta con suficiente evidencia estadística para demostrar que este efecto sea heterogéneo en el monto de compra ni en el gasto promedio mensual. Por otro lado, en las variables frecuencia promedio de compra y recencia de primera compra posterior al canje, si bien se cuenta con evidencia estadística para hablar de un efecto heterogéneo, en el primer caso los individuos con CATE significativo son poco más del 50% del total de clientes tratados, mientras que para el segundo caso no superan el 50%. Esto genera dudas sobre el input sobre el cual fue construido el modelo, ya que aproximadamente de la mitad representan un efecto no significativo.

Si bien los efectos anteriormente estimados miden el impacto del primer canje a corto-mediano plazo, ya que toman en cuenta el comportamiento del cliente en los 4 meses posteriores al canje, se creyó interesante tomar como punto de comparación cómo cambia el comportamiento del cliente inmediatamente después del éste. Para aquello, utilizando la misma metodología que para estimar el efecto en el monto condicional de compra y la frecuencia promedio de compra, se estimó el efecto sobre el monto gastado en el mes siguiente al período de canje<sup>15</sup> y la propensión del cliente a comprar en el mismo mes. Los resultados de estos dos efectos se presentan en la siguiente tabla.

Tabla 23: Resultados de la estimación del efecto al mes siguiente mediante Causal Forest para la segunda aplicación

Caso	ATT	p-valor	Porcentaje de individuos con CATE significativo
Monto gastado al mes siguiente	33,38 [US\$]	<0,001	98,69%
Propensión de compra al mes siguiente	9,26%	<0,001	40,19%

En la tabla es posible apreciar que el efecto medido a mediano plazo se repite a corto plazo, ya que se encuentra un efecto estadísticamente significativo en el monto de compra gastado al mes siguiente, cercano a los 30 [US\$], mientras que la diferencia en la propensión a comprar el mes

<sup>15</sup> El siguiente mes al período de canje no es necesariamente el primer mes posterior al canje de los individuos, pero es la mejor manera de estandarizar para realizar posteriormente una comparación entre los individuos que canjearon y los únicos. Es importante destacar que se puede estar subestimando o sobrestimando el efecto en el mes siguiente al canje.

siguiente sigue la misma tendencia de la frecuencia promedio de compra, aumentando cerca de un 10% para los clientes que realizan su primer canje de puntos.

## 6.2.4. ESTIMACIÓN DEL EFECTO MEDIANTE METODOLOGÍAS ALTERNATIVAS

Al igual que en la primera aplicación, se comparó el resultado obtenido mediante bosques causales, con las dos metodologías alternativas basadas en el propensity score. Se estimó el mismo efecto en el monto condicional de compra, el gasto promedio mensual, la frecuencia promedio de compra y la recencia de la primera compra posterior al canje.

En primer lugar, se llevó a cabo la estimación mediante propensity score matching. Al igual que en la primera aplicación, el propensity score fue construido mediante una regresión logística, tomando la asignación del tratamiento como variable dependiente, mientras que las variables independientes fueron asignadas según corresponda el caso, acordes con la tabla 17. Se probó el método realizando el matching con los  $K = 1$  y los  $K = 5$  vecinos más cercanos. En este caso el emparejamiento también fue con reposición.

Para el caso de la estimación mediante propensity score weighting, la fórmula para estimar el propensity score fue la misma. El peso utilizado fue el correspondiente para la estimación del ATT<sup>16</sup>. A continuación, se presenta una tabla resumen comparativa de los efectos estimados mediante las tres metodologías.

Tabla 24: Comparación de las estimaciones del efecto mediante las tres metodologías para la primera aplicación

<b>Modelo</b>	<b>ATT en el monto condicional de compra [US\$] (p-valor)</b>	<b>ATT en el gasto promedio mensual [US\$] (p-valor)</b>	<b>ATT en la frecuencia promedio de compra (p-valor)</b>	<b>ATT en la recencia de la primera compra post canje [meses] (p-valor)</b>
Causal Forest	41,03 [US\$] ( $<0,001$ )	31,99 [US\$] ( $<0,001$ )	8,85% ( $<0,001$ )	-0,22 ( $<0,001$ )
Propensity Score Weighting	43,33 [US\$] ( $<0,001$ )	31,70 [US\$] ( $<0,001$ )	7,64% ( $<0,001$ )	-0,15 ( $<0,001$ )
Propensity Score Matching 1:1	39,95 [US\$] ( $<0,001$ )	26,98 [US\$] ( $<0,001$ )	5,85% ( $<0,001$ )	-0,11 (0,053)
Propensity Score Matching 1:5	39,43 [US\$] ( $<0,001$ )	29,45 [US\$] ( $<0,001$ )	7,51% ( $<0,001$ )	-0,14 ( $<0,002$ )

<sup>16</sup> Dependiendo de la submuestra en la cual se quiere estimar el efecto (tratados, no tratados, total, etc.) se utiliza un peso distinto para cada cliente en la estimación mediante propensity score weighting. En este caso se utilizó la que corresponde a la estimación del efecto en los tratados. Para más información ver Hirano, Imbens & Ridder (2003).

De la tabla 24 se desprende que, al igual que en la primera aplicación, la estimación del efecto es consistente entre las distintas metodologías. Considerando sólo esto, la metodología de Causal Forest no agrega mucho valor con respecto a las metodologías clásicas como las basadas en propensity score.

## 6.2.5. CARACTERIZACIÓN INDIVIDUAL DEL EFECTO ESTIMADO MEDIANTE CAUSAL FOREST

Los resultados de la sección anterior indicaron que para esta aplicación no hay suficiente evidencia estadística como para determinar que existe heterogeneidad en el efecto del primer canje de puntos en el comportamiento del cliente para los modelos que tienen una mayor cantidad de individuos con efecto significativo, por lo tanto, en esta sección se procede a comprobar que no se justifica realizar una caracterización individual del efecto, dado que el modelo no fue capaz de detectar heterogeneidad en éste.

Para comprobar que no existe heterogeneidad en el efecto del primer canje de puntos, se realizó – al igual que en la primera aplicación – una regresión lineal del tipo log-log para ver la incidencia que tiene cada una de las covariables en la estimación del CATE. Los resultados de dicha regresión lineal se presentan a continuación.

Tabla 25: Resultados de la regresión lineal del CATE estimado mediante Causal Forest en función de las covariables para la segunda aplicación

<b>Coefficiente</b>	<b>Estimación</b>	<b>Error estándar</b>	<b>P-valor</b>
Intercepto	3,370	0,034	<0,001
Log (Puntos acumulados)	0,030	0,004	<0,001
Log (Puntos expirando)	-0,014	0,001	<0,001
Log (Puntos expirados)	-0,024	0,001	<0,001
Log (Frecuencia promedio de compra)	0,025	0,005	<0,001
Log (Monto compra condicional)	-0,001	0,002	0,470
Log (Recencia de última compra antes del canje)	-0,020	0,009	0,042
Requisito de canje	0,044	0,008	<0,001
Año (2005)	0,024	0,007	<0,001
Año (2006)	-0,006	0,007	0,393

El coeficiente de determinación  $R^2$  de la regresión es de 0,36, al igual que el coeficiente de determinación ajustado. Esto significa que gran parte de la varianza del CATE no logra ser explicada con el modelo, a diferencia de la aplicación con datos contractuales, en la cual el

coeficiente de determinación indicaba que casi el 90% de la varianza del CATE lograba ser explicada por el modelo.

Adicionalmente, para detectar una relación entre las variables más importantes en la construcción del árbol y el CATE estimado por el mismo, se calculó el coeficiente de correlación entre este último y las covariables, obteniendo correlaciones muy bajas. En la siguiente tabla se presentan los coeficientes de correlación estimados.

Tabla 26: Correlación entre el CATE estimado mediante Causal Forest y las covariables más importantes para la segunda aplicación

Covariable	Coefficiente de correlación con CATE
Puntos acumulados	0,31
Puntos expirando	-0,19
Puntos expirados	-0,42
Monto condicional de compra	0,24

De ambas tablas es posible desprender que, si bien los CATE estimados para la muestra de clientes que realizaron su primer canje presentan una varianza, el modelo no la logra explicar de buena manera. Por lo tanto, no existe suficiente evidencia estadística para afirmar que el efecto del primer canje de puntos es heterogéneo en los modelos de gasto promedio.

## 6.2.6. CONCLUSIONES DE LA SEGUNDA APLICACIÓN

De este estudio sobre el impacto del primer canje de puntos en el comportamiento de los clientes pertenecientes a un programa de lealtad de una empresa que no tiene un contrato asociado con ellos, se pueden obtener distintas conclusiones.

En primer lugar, hay suficiente evidencia estadística para determinar que el primer canje de puntos sí afecta el comportamiento del cliente posterior a la realización del canje. En particular, los clientes aumentan su monto de compra promedio en 41,03 [US\$] y su frecuencia promedio de compra en un 8,85% con respecto a los clientes que no realizan su primer canje, pero que sí podrían haberlo realizado. Esto se refleja también en una diferencia de 32 [US\$] en el gasto promedio mensual entre los clientes que realizan su primer canje y los que no. Además, los clientes que realizan su primer canje, pasado el período definido como de tratamiento, vuelven a comprar, en promedio, casi una semana antes que los clientes que no canjearon.

Por ende, existe un efecto significativo en los clientes posterior a su primer canje, sin embargo, no se tuvo suficiente evidencia estadística para afirmar que este efecto es heterogéneo. Esto no significa que en la realidad no exista heterogeneidad en el efecto del primer canje de puntos en el comportamiento de los clientes bajo régimen no contractual, sino que, con los datos utilizados, no se logra reunir suficiente evidencia estadística para poder rechazar con un alto grado de certeza que el efecto del primer canje es homogéneo en la población.

Son dos las limitaciones principales de este estudio que pueden estar afectando en la capacidad del modelo de detectar un efecto heterogéneo. La primera es la cantidad de clientes observados, ya que, si bien la cantidad es similar a la de la primera aplicación, aun así, las bases de datos de este tipo de empresa cuentan con mucha más información que podría ser utilizada para robustecer el resultado. En la misma línea, la cantidad de variables observables de cada individuo podría mejorarse también, ya que en esta aplicación se contó solamente con información transaccional, no obstante, información demográfica de los clientes, como su edad, residencia o género, podrían ser útiles para caracterizarlos de mejor manera y realizar un matching más preciso.

Si bien, este modelo falló en detectar un efecto heterogéneo del primer canje de puntos en el comportamiento del cliente, todavía quedan dudas sobre la verdadera naturaleza de este efecto. Por lo tanto, se deja propuesto para trabajos futuros una investigación más robusta sobre el efecto del primer canje de puntos para clientes bajo régimen no contractual.

## 7. CONCLUSIONES

Mediante este trabajo de investigación ha sido posible estimar un efecto significativo del primer canje de puntos en el comportamiento del cliente, tanto para clientes bajo régimen contractual como para clientes bajo régimen no contractual.

Es importante recalcar nuevamente que este estudio se basa en un cuasiexperimento, es decir, el tratamiento no fue asignado aleatoriamente. A través de distintas decisiones y metodologías se buscó reducir la endogeneidad inherente a este tipo de estudios, causada principalmente por el sesgo por autoselección, sin embargo, se debe tener en cuenta que muy difícilmente puede ser eliminada. De hecho, dada la falta de variables en este experimento, se toma un supuesto sumamente fuerte al asumir que la decisión de canjear está dada solamente por las variables disponibles, lo cual muy probablemente no sea así. Si fuese posible tener más variables observables, este sesgo podría ser reducido.

A través de este estudio, fue posible poner a prueba la metodología de los bosques causales, la cual, bajo el supuesto de unconfoundedness, fue capaz de estimar un efecto significativo y heterogéneo, muy similar al estimado mediante otras metodologías basadas en el propensity score.

El hecho de haber obtenido estimaciones muy similares entre bosques causales y las metodologías alternativas podría llevar a pensar que no suponen ninguna ventaja en relación con ellos. Sin embargo, es importante mencionar que las variables utilizadas en este estudio fueron muy reducidas, por lo que no se pudo comprobar el verdadero potencial del algoritmo, ya que éste presenta evidentes ventajas en relación con otros métodos en espacios de alta dimensionalidad. El hecho de que, al agregar variables al modelo como la varianza de las variables de comportamiento, los resultados no variaran mucho, no es un indicio de que el modelo no sirva. Por el contrario, esto podría indicar que tal vez esas variables no están relacionadas con la decisión de canjear puntos.

Por lo tanto, en lo que respecta a este estudio en particular, mediante bosques causales no se gana mucho más, con respecto a las metodologías alternativas, que detectar heterogeneidad en el efecto del primer club de lealtad estudiado.

Con respecto al efecto estimado, para los clientes bajo régimen contractual no fue posible encontrar un efecto significativo en la cantidad de productos que compran mensualmente. Sin embargo, se encontró una diferencia significativa de 6,97 [US\$] en el monto promedio gastado mensualmente en la empresa entre los clientes que realizaron su primer canje de puntos y los que no, por lo que se podría decir que el primer canje aumenta el ARPU mensual de los clientes que lo realizan en un 12,7%. Además, fue posible comprobar que este efecto del primer canje de puntos no es igual para todos los clientes. Los clientes más activos, es decir, los que gastan más y compran más cosas en la empresa, tienden a modificar en mayor magnitud su consumo posterior al primer canje que los clientes menos activos. Además, el efecto también suele ser mayor en hombres que en mujeres. Por alguna razón, el efecto estimado tiende a ser mayor en personas cercanas a los 40 años.

Por otro lado, para el estudio realizado con clientes bajo régimen no contractual, también se encontró un efecto del primer canje de puntos significativo estadísticamente. En particular, los clientes que realizan su primer canje de puntos gastan en promedio 41 [US\$] más por compra en los 4 meses siguientes al primer canje, es decir, aumentan su monto de compra en un 25,4% con respecto a los que no realizan su primer canje de puntos, siendo este efecto mayor que en el caso

de los clientes bajo régimen contractual. Además, aumentan su frecuencia promedio de compra en un 8,85% y vuelven a comprar después del período designado para el primer canje de puntos una semana antes que los clientes que no realizaron su primer canje en el período determinado. Este patrón se repite cuando se estima en el mes siguiente al período de canje definido, dado que los clientes que realizan su primer canje de puntos gastan en promedio 33,38 [US\$] en el mes posterior al período designado de canje, mientras que tienen una propensión a comprar 9,26% mayor que los que no realizan el canje en el período estimado, en el mismo mes.

Para el caso de estudio con clientes bajo régimen no contractual, no fue posible confirmar que este efecto no es igual para todos los clientes, pero esto no significa que sea así en la realidad, sino que el modelo no tuvo la capacidad de detectar heterogeneidad en el efecto. Con una base de datos más sólida, con más observaciones y una mejor caracterización de cada individuo, podría ser posible detectar heterogeneidad en el efecto, pero no se puede asegurar que realmente exista.

Los resultados de este estudio indican que, para estas dos aplicaciones en particular, se encontró un efecto significativo del primer canje para clientes bajo régimen contractual y no contractual, por lo que se puede decir que la participación de un cliente en un programa de lealtad modifica su comportamiento. Sin embargo, se debe ser cuidadoso con la interpretación de estos resultados, ya que un aumento de alrededor del 20% de compra mensual solamente por haber canjeado genera dudas acerca de la veracidad del efecto estimado. El efecto podría estar sobrestimado debido a que el primer canje podría estar motivado por otras características de los clientes no observables en este efecto.

Por lo tanto, los efectos estimados en este trabajo y sus caracterizaciones no necesariamente son extrapolables a otros programas de lealtad. No obstante, este estudio sienta las bases para desarrollar metodologías que permitan estudiar efectos causales heterogéneos en el comportamiento del cliente, ya sea causados por promociones, programas de lealtad o cualquier intervención que pueda ser realizada en los clientes.



## 8. TRABAJOS FUTUROS

Debido a los datos disponibles, este estudio se enfocó en estimar el impacto del primer canje de puntos en la incidencia y monto de compra de los clientes, para ambas aplicaciones. Sin embargo, no se pudo estimar el efecto en otras variables relacionadas a la lealtad como, por ejemplo, la duración del individuo como cliente de la empresa. Por lo tanto, se propone para trabajos futuros el estudio del efecto del primer canje de puntos en otras variables relacionadas a la lealtad de los clientes.

En la misma línea del estudio, éste tiene una limitación muy importante en la cantidad de variables por las cuales se controló el tratamiento. Dada la carencia de estas variables de control, se tuvo que tomar un supuesto sumamente fuerte, al asumir que los clientes controlados por esas variables eran comparables entre sí. Debido a esta misma limitante, existe también la chance de que no haya sido posible detectar heterogeneidad en el efecto del primer canje de puntos en el comportamiento del cliente bajo régimen no contractual. Sin embargo, esto no quiere decir que de verdad el efecto no sea heterogéneo, sino que los datos utilizados no permiten que el modelo detecte dicha heterogeneidad. Los bosques causales tienen una gran ventaja, en comparación a otros métodos, ya que funcionan particularmente bien cuando el vector que caracteriza a cada observación tiene muchas dimensiones. Por lo tanto, es posible utilizar esta misma metodología, pero con más variables observables para estimar un efecto más robusto y, quizás, detectar una heterogeneidad significativa para la misma aplicación.

Por otro lado, este estudio se enfoca en detectar un impacto del primer canje de puntos en el comportamiento de los individuos asociadas a su lealtad conductual, sin embargo, no se está midiendo el impacto que tiene en la lealtad actitudinal de los clientes. Volviendo a la definición de lealtad, ésta se conforma por dos componentes (Dick & Basu, 1994): lealtad conductual (asociada a los patrones de consumo) y actitudinal (asociada con la actitud del cliente frente a la empresa y el nivel de apego). Por lo tanto, este estudio está detectando un efecto en la lealtad conductual, pero no necesariamente significa que el cliente sea leal realmente. Queda propuesto realizar estudios que midan el impacto del primer canje de puntos en la lealtad actitudinal de los clientes.

Finalmente, es importante mencionar que este estudio mide el efecto en el comportamiento del cliente a corto y mediano plazo, sin embargo, no se ocupa de los efectos a largo plazo. Como se mencionó anteriormente, los clientes leales aumentan significativamente el monto de sus compras a largo plazo, no obstante, no existen muchos estudios que se encarguen de estudiar el efecto de los programas de lealtad a largo plazo (Breugelmans et. al., 2015). Es por esta razón que se propone también estudiar el efecto del primer canje de puntos en la evolución del monto de compra para períodos más largos de tiempo.

## 9. BIBLIOGRAFÍA

Athey, S., Imbens, G. W., & Wager, S. (2018). Approximate residual balancing: debiased inference of average treatment effects in high dimensions. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 80(4), 597-623.

Athey, S., & Wager, S. (2018). Estimation and inference of heterogeneous treatment effects using random forests. *Journal of the American Statistical Association*, 113(523), 1228-1242.

Athey, S., & Wager, S. (2019). Estimating treatment effects with causal forests: An application. *arXiv preprint arXiv:1902.07409*.

Bijmolt, T. H., Dorotic, M., & Verhoef, P. C. (2011). Loyalty programs: generalizations on their adoption, effectiveness and design. *Foundations and Trends® in Marketing*, 5(4), 197-258.

Breiman, L. (2001). Random forests. *Machine learning*, 45(1), 5-32.

Breugelmans, E., Bijmolt, T. H., Zhang, J., Basso, L. J., Dorotic, M., Kopalle, P., ... & Wunderlich, N. V. (2015). Advancing research on loyalty programs: a future research agenda. *Marketing Letters*, 26(2), 127-139.

Carboni, A. (2012). Estudio de las respuestas de los consumidores ante un programa de lealtad. Tesis para optar al grado de magíster en gestión de operaciones. Memoria para optar al título de ingeniero civil industrial. Universidad de Chile.

Chernozhukov, V., Chetverikov, D., Demirer, M., Duflo, E., Hansen, C., Newey, W., & Robins, J. (2018a). Double/debiased machine learning for treatment and structural parameters.

Chernozhukov, V., Demirer, M., Duflo, E., & Fernandez-Val, I. (2018b). Generic machine learning inference on heterogeneous treatment effects in randomized experiments (No. w24678). National Bureau of Economic Research.

Dick, A. S., & Basu, K. (1994). Customer loyalty: toward an integrated conceptual framework. *Journal of the academy of marketing science*, 22(2), 99-113.

Dorotic, M., Bijmolt, T. H., & Verhoef, P. C. (2012). Loyalty programmes: Current knowledge and research directions. *International Journal of Management Reviews*, 14(3), 217-237.

Heckman, J. J. (1979). Sample selection bias as a specification error. *Econometrica: Journal of the econometric society*, 153-161.

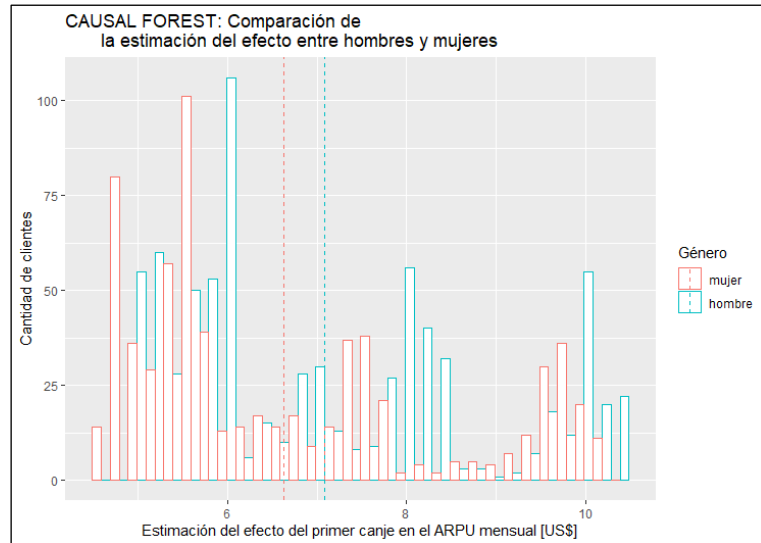
Hirano, K., Imbens, G. W., & Ridder, G. (2003). Efficient estimation of average treatment effects using the estimated propensity score. *Econometrica*, 71(4), 1161-1189.

Jenks, G. F. (1967). The data model concept in statistical mapping. *International yearbook of cartography*, 7, 186-190.

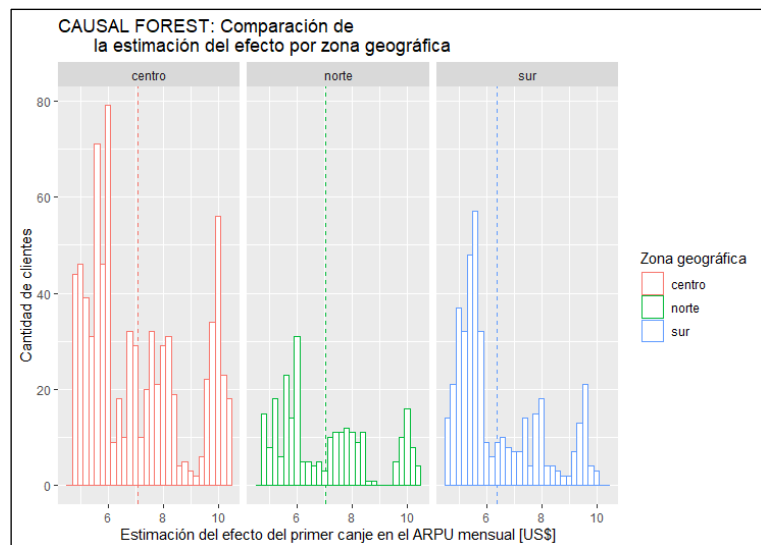
- Leenheer, J., Van Heerde, H. J., Bijmolt, T. H., & Smidts, A. (2007). Do loyalty programs really enhance behavioral loyalty? An empirical analysis accounting for self-selecting members. *International Journal of Research in Marketing*, 24(1), 31-47.
- Liu, Y. (2007). The long-term impact of loyalty programs on consumer purchase behavior and loyalty. *Journal of marketing*, 71(4), 19-35.
- Montoya, R., & Flores, C. (2019). Buying free rewards: the impact of a points-plus-cash promotion on purchase and reward redemption. *Marketing Letters*, 30(1), 107-118.
- Morgan, S., & Winship, C. (2007). *Counterfactuals and Causal Models*.
- Nie, X., & Wager, S. (2017). Quasi-oracle estimation of heterogeneous treatment effects. arXiv preprint arXiv:1712.04912.
- Reichheld, F. F., & Sasser, W. E. (1990). Zero defections: Quality comes to services. *Harvard business review*, 68(5), 105-111.
- Ríos, C. (2013). Análisis del efecto de un club de fidelización en el comportamiento de clientes bajo régimen contractual. Memoria para optar al título de ingeniero civil industrial. Universidad de Chile.
- Robins, J. M., Rotnitzky, A., & Zhao, L. P. (1994). Estimation of regression coefficients when some regressors are not always observed. *Journal of the American statistical Association*, 89(427), 846-866.
- Rosenbaum, P. R., & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, 70(1), 41-55.
- Rosenbaum, P. R., & Rubin, D. B. (1985). Constructing a control group using multivariate matched sampling methods that incorporate the propensity score. *The American Statistician*, 39(1), 33-38.
- Rosenbaum, P. R. (1987). Model-based direct adjustment. *Journal of the American Statistical Association*, 82(398), 387-394.
- Rosenbaum, P. R. (2002). Observational studies. In *Observational studies* (pp. 1-17). Springer, New York, NY.
- Tibshirani, J., Athey, S., Friedberg, V., Hadad, V., Miner, L., Wager, S., & Wright, M. (2018). grf: Generalized Random Forests (Beta). URL <https://github.com/grf-labs/grf>. R package version 0.10.2.
- Wang, Y., Lewis, M., Cryder, C., & Sprigg, J. (2016). Enduring effects of goal achievement and failure within customer loyalty programs: A large-scale field experiment. *Marketing Science*, 35(4), 565-575.

# 10. ANEXOS

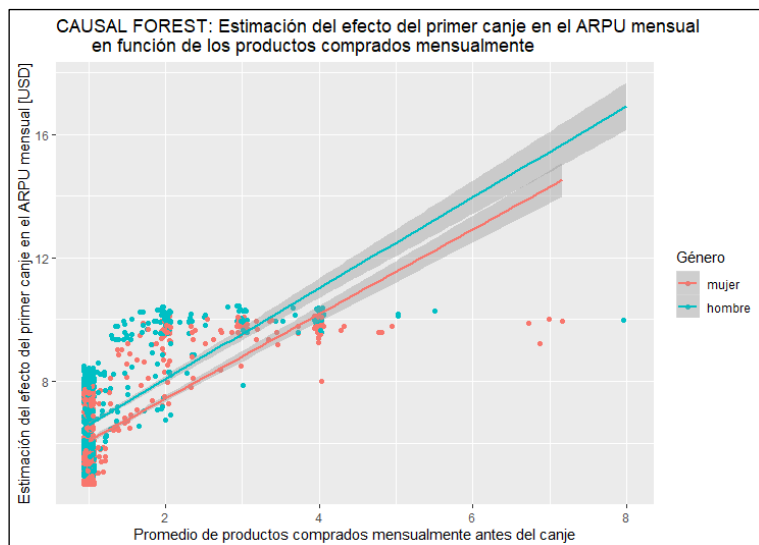
## Anexo 1. Histograma del efecto estimado mediante Causal Forest por género



## Anexo 2. Histograma del efecto estimado mediante Causal Forest por zona geográfica



### Anexo 3. Efecto estimado mediante Causal Forest en función de los productos comprados mensualmente, por género



### Anexo 4. Efecto estimado mediante Causal Forest en función del ARPU mensual promedio antes del primer canje, por género

