



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

**HETEROGENEIDAD DE EFECTOS DE TRATAMIENTO EN COMPORTAMIENTO
FINANCIERO A TRAVÉS DE MÉTODOS DE APRENDIZAJE AUTOMÁTICO**

MEMORIA PARA OPTAR AL TÍTULO DE INGENIERO CIVIL INDUSTRIAL

ADOLFO ESTEBAN VILLANUEVA BONILLA

**PROFESOR GUÍA:
DANIEL SCHWARTZ PERLROTH**

**MIEMBROS DE LA COMISIÓN:
JUAN PABLO ROMERO GODOY
RICARDO MONTOYA MOREIRA**

**SANTIAGO DE CHILE
2020**

**RESUMEN DE LA MEMORIA PARA OPTAR AL
TÍTULO DE: INGENIERO CIVIL INDUSTRIAL
POR: ADOLFO ESTEBAN VILLANUEVA BONILLA
FECHA: 17 DE ABRIL DE 2020
PROFESOR GUÍA: DANIEL SCHWARTZ PERLROTH**

**HETEROGENEIDAD DE EFECTOS DE TRATAMIENTO EN
COMPORTAMIENTO FINANCIERO A TRAVÉS DE MÉTODOS DE
APRENDIZAJE AUTOMÁTICO**

Hoy en día científicos, investigadores, ingenieros y grandes empresas tienen como desafío comprender la heterogeneidad de los efectos de sus tratamientos e intervenciones (que pueden variar desde la medicina personalizada hasta recomendaciones de marketing de la misma naturaleza). En la presente memoria se implementa un experimento de campo y se utilizan métodos de aprendizaje automático para determinar la heterogeneidad asociada al efecto de enviar distintos recordatorios de pagos de tarjeta de crédito, y así evaluar qué tipo de recordatorio puede focalizarse a ciertos grupos.

Los tratamientos del experimento de campo destacan los montos referenciales de pago (monto mínimo y total facturado), así como la explicación de éstos. El efecto que se estudia es la consecuencia de caer en mora, esto es, que el sujeto no haya pagado o haya pagado menos que el monto mínimo. La inferencia del efecto de tratamiento y su heterogeneidad se realiza mediante Causal Forest, que es una herramienta proveniente del área de Aprendizaje Automático, no paramétrica y que extiende el algoritmo de Random Forest de Breiman. En el marco de resultados potenciales sin confusión, los Causal Forest han mostrado ser consistentes para el *verdadero* efecto del tratamiento y tienen una distribución de muestreo asintóticamente gaussiana y centrada.

El resultado de utilizar este método es que al agregar los montos referenciales de pago en los recordatorios, la mora promedio disminuye en cerca del 1% (comparando con sujetos que se les ha recordado el pago, pero no expuestos a estas cifras). Al analizar a los sujetos por cuartiles según sus respuesta al tratamiento, se encuentra que no hay efecto adverso detectable en ningún cuartil (aumentar la mora), y que el cuartil donde se presenta un mayor efecto de disminución de la mora se caracteriza por presentar sujetos con porcentajes de pago bajos, irregulares y que han caído previamente en mora reiteradas veces.

Finalmente, dividiendo la muestra en sets de entrenamiento y testeo, de manera de focalizar un posible tratamiento futuro en este último, se encuentra que de enviar distintos mensajes de pago a ciertos clientes (es decir, entregar el mensaje correcto a quién corresponde según el modelo), se puede reducir la mora promedio hasta en un 7%. Se concluye que la metodología utilizada en esta memoria es capaz de detectar patrones de heterogeneidad del efecto del tratamiento.

Dedicatoria

A todas y todos
Que han luchado por un mejor Chile
Y por un mejor mundo,
Donde se viva dignamente,
Donde exista justicia,
Donde exista solidaridad,
Donde prevalezca el amor.

Agradecimientos

Le agradezco a mi familia por el amor y la felicidad que me entregan desde que recuerdo, cursando conmigo y apoyándome en las etapas de mi vida. A mi madre Jacqueline por darme su confianza y enseñarme lo que es perseverancia y garra. A mi padre Emando por darme la frase precisa en el momento adecuado. A mi hermano mayor Sebastián por ser mi primer amigo. A mis hermanos menores Monserrat y Renato, por ayudarme a ser mejor persona, re-enseñándome lo esencial de ser niño.

A mi pareja Rose por llegar a mi vida y enseñarme lo bonito de caminar codo a codo en ésta. Por su dulce amor que me expresa día a día con cada palabra, con cada caricia y mostrarme lo bonito que es amar. Por sacar la mejor versión de mí a flote y por el constante empuje y fortaleza que me entrega en cada desafío que se presenta.

A mi tía Ale, por su crianza y principios que he recibido, junto con todo el amor que me ha entregado en la vida.

A mis amigos de infancia Mati, Malu, Paya, Peña y Osmar, por su leal amistad que me han dado durante todos estos años, en los que hemos compartido inolvidables procesos y experiencias juntos.

A mis amigos que he conocido desde la U y la convirtieron en un proceso único y significativo, Franco, Toledo, Gaba, Joaquín, Hodali, Basti, Zeta, Benja, Rorro, Javi, Gabi, Cris, Agus, Josefa, Marlene y Nico.

A mis amigos del Cheh-Rasmus Thomas, Lili, Jaki y Julien, con quienes descubrí hermosas perspectivas de la vida y me han enseñado que el cariño sobrepasa distancias.

A mi profesor Daniel Schwartz, por su constante apoyo y guía en este proceso.

Gracias a Dios, a la Vida, al Universo y a la Energía por todo este Amor y tan buenas vibras.

Tabla de Contenido

Capítulo 1. Introducción	- 1 -
i. Revisión de Literatura.....	- 2 -
ii. Objetivos	- 4 -
iii. Alcances	- 5 -
Capítulo 2. Metodología	- 6 -
i. Diseño experimental y tratamientos.....	- 6 -
ii. Efectos a estudiar e Hipótesis.....	- 11 -
iii. Universo a contactar e implementación del piloto.....	- 11 -
iv. Seguimiento del tratamiento	- 14 -
v. Causal Forest	- 14 -
vi. Implementación de la estimación de efectos con Causal Forest.....	- 17 -
vii. Best Linear Predictor	- 18 -
Capítulo 3. Resultados	- 19 -
i. Average Treatment Effect (ATE)	- 19 -
ii. Evaluación de heterogeneidad.....	- 20 -
iii. Heterogeneidad entre subgrupos.....	- 20 -
iv. Heterogeneidad entre covariables.....	- 21 -
v. ¿Qué mensaje enviar y a quién?	- 26 -
Capítulo 4. Conclusión	- 29 -
Capítulo 5. Trabajos futuros	- 31 -
Bibliografía	- 32 -
Anexo y apéndices	- 34 -

Índice de tablas

Tabla 1: Test de ANOVA para respaldar Blocked Randomization Assignment.	- 12 -
Tabla 2: Resumen de cantidad de sujetos por tratamiento.	- 13 -
Tabla 3: Resumen de las tasas de lectura del tratamiento agrupadas por grupo.	- 14 -
Tabla 4: Resumen del seguimiento de los tratamientos por grupo y por envío.	- 14 -
Tabla 5: ATE y su intervalo de confianza al 95%.	- 19 -
Tabla 6: Resultado del test de calibración Best Linear Predictor.	- 20 -
Tabla 7: Sample ATE y AIPW ATE por cuartil.	- 20 -
Tabla 8: Anova test y valores de variables X_1 , X_2 , X_3 y X_4	- 22 -
Tabla 9: Resumen test de Tukey HSD.	- 23 -
Tabla 10: ATE de los tratamientos 1, 2 y 3 versus el tratamiento 4.	- 27 -
Tabla 11: Ejemplo de selección de tratamiento.	- 27 -
Tabla 12: Resultado de optimizar el tratamiento en la mora promedio.	- 28 -
Tabla 13: Ajustes de parámetros de modelos Causal Forest (Anexo).	- 34 -

Índice de ilustraciones

Ilustración 1: Gráfica del tratamiento del Grupo 1.....	- 7 -
Ilustración 2: Gráfica del tratamiento del Grupo 2	- 8 -
Ilustración 3: Gráfica del tratamiento del Grupo 3	- 9 -
Ilustración 4: Gráfica del tratamiento del Grupo 4	- 10 -
Ilustración 5: Histogramas de Dificultad Transaccional de los cuatro grupos.....	- 13 -
Ilustración 6: Histogramas de Tasa de lectura de los cuatro grupos.	- 13 -
Ilustración 7 : Histograma de las estimaciones del CATE tipo OOB	- 19 -
Ilustración 8: ATE más intervalos de confianza al 90% por cuartil, mediante AIPW ATE y Sample ATE.	- 21 -
Ilustración 9: Variación de CATE según X_1	- 23 -
Ilustración 10: CATE variando según X_4	- 24 -
Ilustración 11: variación de CATE según mora en los últimos 12 meses.....	- 25 -
Ilustración 12: variación de CATE según mora en los últimos 4 meses	- 26 -
Ilustración 13: Histograma de Sexo entre los cuatro grupos (Anexo).....	- 34 -

Índice de ecuaciones

Ecuación 1: Conditional Average Treatment Effect.....	- 3 -
Ecuación 2: Supuesto de Unconfoundedness.....	- 15 -
Ecuación 3: Supuesto de Overlap.....	- 16 -
Ecuación 4: Propensión a ser tratado y función de respuesta, respectivamente.	- 16 -
Ecuación 5: Función de pérdida R-loss.....	- 16 -
Ecuación 6: Predicción CATE tipo OOB del Causal Forest	- 16 -
Ecuación 7: Augmented Inverse-Propensity Weighted Average Treatment Effect. .	- 17 -
Ecuación 8: Importancia de variable x_j del modelo Causal Forest.	- 17 -
Ecuación 9: Best Linear Predictor.....	- 18 -

Capítulo 1. Introducción

Cuando se desea establecer evidencia causal y conocer el efecto de, por ejemplo, tratamientos médicos, estrategias de marketing directo y programas sociales, los experimentos aleatorizados son protagonistas (Duflo, Glennerster & Kremer, 2007; Imbens & Rubin, 2015). Un ejemplo de estos son los experimentos de campo que se caracterizan por situar a los sujetos en un contexto real (Gerber & Green, 2012). Complementando la búsqueda del efecto promedio del tratamiento, los investigadores y gestores de políticas sociales están interesados en conocer la heterogeneidad de los efectos, por ejemplo, si varía o depende de factores o características individuales de los sujetos (socio-demográficas, conductuales, transaccionales, preferencias, entre otras). Esto anterior permite, entre otras cosas, optimizar las intervenciones, dado que se pueden conocer grupos específicos donde se debe focalizar los tratamientos (y donde no).

En este último propósito (inferencia de efectos heterogéneos de tratamientos) es donde ha aparecido la disciplina de Machine Learning y sus herramientas, que a diferencia de métodos tradicionales (como métodos paramétricos, regresión y otros) no sufren dificultades como el riesgo y limitantes del Overfitting, trabajar con submuestras, asumir supuestos estrictos como distribución de datos, entre otras (Chernozhukov, Demirer, Duflo & Fernández-Val, 2018). Siguiendo lo anterior, han surgido diversos trabajos e investigaciones de variadas áreas (y aplicados en diversos rubros) que aplican métodos y técnicas de este mundo, es más, es una de las disciplinas que ha impactado directamente a la hora de realizar estudios en la Economía y Econometría, entre otras áreas (Athey, 2018).

En este trabajo se abordará el diseño e implementación de un experimento aleatorio estratificado, que consta de 4 tipos de recordatorios (tratamientos) a clientes que cuentan con deuda crediticia. El foco del trabajo consiste en determinar la heterogeneidad del tratamiento. Esta inferencia del efecto y su heterogeneidad se realizará mediante herramientas provenientes del área de Machine Learning, particularmente, a través de Causal Forest.

Los tratamientos que se utilizan en esta memoria están basados en montos referenciales de mínimo facturado que han mostrado influir en las decisiones de las personas (Navarro-Martinez, Salisbury, Lemon, Stewart, Matthews & Harris, 2011; Keys & Wang, 2018). Uno de los mecanismos que explican este efecto es basado en el sesgo de anclaje y ajuste (Kahneman & Tversky, 1982) en el cual las personas varían sus decisiones influenciadas por valores numéricos.

La relevancia de este trabajo va relacionada con que la oferta y operación de tarjetas de créditos es una de las actividades claves de las instituciones financieras, como bancos y de otros emisores no bancarios (como el caso de algunas casas comerciales). En Chile, estos emisores de tarjetas de crédito son regulados y supervisados por la Comisión para el Mercado Financiero (CMF), servicio público descentralizado que se relaciona con el Presidente de la República de Chile a través del Ministerio de Hacienda, desde el 01 de junio del año 2019 y previamente por la Superintendencia de Bancos e Instituciones Financieras (SBIF), la cual se integró a la CMF en la misma fecha.

Según el Informe de Tarjetas de Crédito (CMF, 2019), hoy en día existen más de 17 millones de tarjetas de crédito vigentes, por lo que una gran parte de personas tiene acceso a este tipo de servicio financiero. Según indicadores del Informe de Endeudamiento (SBIF, 2018), el endeudamiento de los clientes bancarios ha aumentado los últimos años. Por ejemplo, el nivel de deuda promedio, de 2017 a 2018, ha aumentado en 7,3% (siendo en 2018, de 3,2 MM\$CLP, definido a partir de la mediana de la distribución). Análogamente, el indicador de carga financiera de 2018 (25,7%) y de apalancamiento de 2018 (5,7 veces el ingreso), superaron a los observados doce meses antes (23,7% y 5,2 veces, respectivamente).

Cuando el cliente no cumple con el pago de su deuda en las fechas estipuladas, entra en un estado de morosidad. En el Informe de Tarjetas de Crédito (SBIF, 2018), en junio de 2018 se registró un monto de créditos en mora promedio entre emisores bancarios de 3.6% (con respecto al total), mientras que, en el Informe de Tarjetas de Crédito del año siguiente (CMF, 2019) este indicador en junio de 2019 ha aumentado, alcanzando un nivel de 5.0%.

Uno de los factores claves para que este tipo de actividad (ofrecer dinero y permitir un pago de diversas formas, como lo son las cuotas) siga siendo rentable para la institución, es que efectivamente el cliente cumpla con su compromiso de pago. Por parte del cliente, el cumplir con su deber le evita que le cobren intereses y otros cobros por no pago (aumentando su deuda). Finalmente, a nivel país, el no cumplimiento de estos compromisos financieros tiene perjuicios económicos, dado que aumentan los precios de los créditos, disminuye el consumo, entre otras consecuencias.

i. Revisión de Literatura

El efecto de tratamiento heterogéneo, o Heterogeneous Treatment Effect (HTE) en inglés, nace cuando se busca entender como el efecto del tratamiento varía de acuerdo a las características individuales de los sujetos. En estadísticas se puede entender como la búsqueda de variables que predicen un mayor efecto del tratamiento, en psicología experimental como “moderadores” del efecto, entre otras descripciones que se le asocian. Según lo anterior, el efecto de tratamiento promedio, o Average Treatment Effect (ATE), es una herramienta que para este propósito se encuentra limitada (pues es un resultado *para todos*) y no serviría, por ejemplo, para poder focalizar algún tratamiento en alguna población específica que tenga una característica en común.

En base a lo anterior es que surge el efecto de tratamiento promedio condicional, o Conditional Average Treatment Effect (CATE) en inglés, que se define como el valor del efecto dando un nivel de observables. En un contexto de regresión lineal se puede entender como una interacción entre una característica (covariable) y el tratamiento, si la interacción resulta significativa (relevante en magnitud y estadísticamente hablando), se puede inferir que el efecto del tratamiento depende de esta covariable. Matemáticamente y bajo el marco de resultados potenciales (Rubin, 1974), denotando $Y_i(1)$ la variable aleatoria que representa la potencial respuesta del sujeto i que haya recibido el tratamiento y por el otro lado $Y_i(0)$ la potencial respuesta del sujeto i que no haya recibido el tratamiento (u otro tratamiento con el que se desee comparar), se escribe como la Ecuación 1:

$$CATE: \tau(x) := E[Y_i(1) - Y_i(0) | X_i = x]$$

Ecuación 1: Conditional Average Treatment Effect

Si, como se dijo anteriormente, con una regresión lineal se pueden obtener estimaciones del CATE, rápidamente entonces emerge la pregunta de por qué han surgido tantas herramientas, específicamente del mundo del Machine Learning, que buscan contestar estas preguntas y estimaciones.

La razón se debe a que las herramientas y enfoques tradicionales se encuentran limitados (Athey, 2018). Por ejemplo, una regresión lineal se puede sobre-ajustar si como entrada recibe múltiples covariables con sus correspondientes interacciones al tratamiento. También, la tradición de investigadores en economía y econometría es de especificar un modelo, estimando en el conjunto de datos completo (no óptimo pues no es “honesto” como se profundizará en los siguientes párrafos de esta sección) y basándose en teoría estadística para estimar los intervalos de confianza para los parámetros estimados, reportando dos o tres alternativas para verificar la robustez, sin embargo, usualmente se verificaron muchos más de los reportados y en general no se comunica pues invalidaría los intervalos de confianza reportados (o sea, intervalos de confianza no válidos) dada las múltiples pruebas y búsqueda de resultados deseados. Entre otras razones, como falta de capacidad (como computacional) para trabajar en ambientes de alta dimensión, el tener que hacer supuestos fuertes sobre los datos (como su distribución) cuando no se cumplen y otras.

Ejemplos de inferencia causal a través de Machine Learning (y que buscan evitar las dificultades recién mencionadas) son trabajos como Belloni, Chernozhukov y Hansen (2014), donde utilizan regresiones regularizadas en un ambiente de muchas covariables (de alta dimensión) y pocas de éstas son las que inciden en el resultado (previamente se desconocían cuáles observables eran ni las importancias de éstos), Athey, Imbens y Wager (2016) donde proponen métodos de balance residual con un foco “doblemente robusto”, donde combinan pesos de equilibrio con ajuste de regresión regularizado. También Chernozhukov, Chetverikov, Demirer, Duflo, Hansen y Newey (2017) donde se propone un método “aprendizaje automático doblemente insesgado” (“double debiased machine learning”), que consta de una regresión semi-paramétrica residual sobre residuos con fin de estimar el efecto promedio del tratamiento y el efecto promedio del tratamiento sobre tratados bajo el supuesto de “falta de confusión” (“unconfoundedness”).

Por otro lado, ejemplos por el lado de la estimación de la heterogeneidad en los efectos de tratamientos mediante Machine Learning son trabajos como Athey e Imbens (2016) donde proponen métodos “honestos” (nueva forma de dividir bases de datos de entrenamiento y prueba que difieren de las tradicionales) para detectar heterogeneidad, los cuales mostraron una cobertura nominal para intervalos de confianza del 90%, a diferencia de los enfoques tradicionales (“no honestos”) donde la cobertura fue entre el 74% y 84%; Athey y Wager (2017) junto con Athey, Tibshirani y Wager (2019) que desarrollan el método de Bosques Causales o Causal Forest en inglés (entendiéndose como el promedio de un conjunto de árboles causales), donde a diferencia de un método tipo vecino más cercano, tiene un enfoque basado en datos, por lo que da distinta importancia a las covariables (en contraste del método vecino más cercano), además, tiene la capacidad de trabajar en espacios de alta dimensión y esparcidos (al igual que en los ejemplos

anteriores), por lo que no tiene el riesgo de caer en un sobreajuste (a diferencia de una regresión lineal). Añadiendo a lo anterior, los árboles causales (que componen a un Causal Forest) son honestos, es decir (en conjunto a lo que se dijo anteriormente), utilizan datos diferentes para construir una partición del espacio de covariables a partir de los datos utilizados para estimar los efectos del tratamiento dentro de los nodos terminales (hojas del árbol), a diferencia de un Bosque Aleatorio (o Random Forest en inglés), donde todos los datos se utilizan tanto para la construcción del bosque como para la estimación de efecto (no honesto, pues un dato de la construcción de un árbol puede encontrarse en la muestra de estimación del efecto de otro árbol), entre otros ejemplos.

En presente trabajo de memoria, se hará inferencia del efecto y heterogeneidad de los tratamientos mediante el modelo Causal Forest. Añadiendo a lo dicho en el párrafo anterior, este método tiene como propósito buscar buenos predictores de los efectos heterogéneos de tratamiento. Estos efectos heterogéneos son descritos por estimaciones del *CATE* (Ecuación 1) siendo capaz de entregar una estimación por cada sujeto o un Individual Treatment Effect (ITE) para cada individuo. Este método se ha popularizado y ha sido aplicado en diferentes trabajos y rubros, mostrando un buen desempeño (en base a variados test), como por ejemplo O’Neill y Weeks (2019) en donde se examina los efectos específicos del hogar al introducir esquemas de precios de electricidad por tiempo de uso, observando heterogeneidad en covariables de cuartiles creados a partir de las estimaciones de CATE, argumentando que la inferencia es razonable y testeando el desempeño del modelo. Otro ejemplo es Davis y Heller (2017), que, en un contexto de pruebas aleatorias controladas de programa de empleos, estiman efectos heterogéneos y muestran la utilidad de predecir en base a los observables (al examinar qué sujetos responden y cuáles no), además de enseñar cómo esta herramienta muestra un mejor desempeño que otros métodos tradicionales. Similares conclusiones se muestran en Lechner (2018), Suk, Kim y Kang (2019) y en Athey y Wager (2019), aunque en distintos rubros de aplicación, y además mostrando la versatilidad de este modelo al poder complementarlo con clústeres predefinidos y complementarlo con otros modelos para realizar la predicción. Se explicará en profundidad la mecánica de esta herramienta en el Capítulo 2, sección v.

ii. Objetivos

Objetivo General

“Determinar la existencia de heterogeneidad de una intervención basada en un experimento de campo aleatorio de recordatorios de pago para disminuir la mora financiera”

Objetivos Específicos

- Implementar piloto en base a los recordatorios de pago determinados.
- Determinar el efecto promedio de los recordatorios de pago.
- Determinar la heterogeneidad del efecto, identificando la importancia relativa de las variables y los segmentos relevantes de consumidores.
- Evaluar la metodología de focalizar el tratamiento en grupos específicos.

iii. Alcances

- El trabajo de se limitará a operaciones de la organización, donde se realiza el trabajo.
- Para medir el efecto y contestar la pregunta de investigación, se observará como variable dependiente el caer en estado de mora. Para luego predecir el efecto de los recordatorios.
- Se trabajará con los productos tipo tarjeta de crédito que permitan pago mínimo, dado el diseño experimental.
- La memoria tiene carácter investigativo. El diseño experimental y metodológico corresponden a una agenda más amplia de la investigación realizada y originada por el profesor guía de esta memoria. Por ello esta memoria se concentra en el desarrollo de la implementación de estos experimentos y su respectivo análisis principal, sin pretender hacer un análisis exhaustivo de los mecanismos que originan los efectos.

Capítulo 2. Metodología

La metodología contiene lo siguiente:

i. Diseño experimental y tratamientos

Se definen los tratamientos a los que los sujetos serán expuestos junto con su diseño experimental:

i.i Sesgo cognitivo econométrico asociado.

Dadas las hipótesis que se tienen, se opta por el sesgo cognitivo econométrico de Anclaje (“Anchoring” en inglés). Éste se puede definir como un fenómeno en el que un individuo basa sus ideas y decisiones en base a una información inicial (ancla).

En este caso y de acuerdo a las hipótesis anteriores, se busca anclar el Monto Mínimo (MM) y el Monto Total Facturado (MTF). Además, existirán dos tratamientos extras, en que uno estarán ambas anclas y en el otro ninguna.

i.ii Cantidad de toques.

A la persona se le recordará dos veces el pago de su tarjeta de crédito (mediante alguno de los tratamientos) previo a su fecha de vencimiento. Uno será 5 días antes de la fecha de vencimiento y el otro toque serán dos días previo a la fecha de vencimiento.

i.iii Vía.

Los tratamientos serán enviados vía correo electrónico.

i.iv Fecha y horario.

Dadas las seis fechas de vencimiento que existen al mes (días 5, 10, 15, 20, 25 y 30 de cada mes) y que se realizará una vez el piloto, es que se recorrerán una vez cada fecha de vencimiento de manera seguida, comenzando desde la fecha de vencimiento 10 de septiembre hasta la fecha 5 de octubre.

Todos los toques se realizarán a la misma hora, siendo definida a las 15:30h.

i.v Encabezado, sub-encabezado y gráficas.

El correo electrónico tiene como encabezado:

“Adolfo, Nombre Organización tiene información importante para ti”

Notar que Adolfo varía por cliente, o sea su nombre. Además, Nombre Organización corresponde al nombre del banco donde se realiza el proyecto.

El sub-encabezado se encontrará vacío.

Considerando los puntos anteriores de esta sección, se definen las siguientes gráficas que recibirán los clientes (por motivos de confidencialidad, se mostrarán incompletas y con escala de grises como filtro).



Ilustración 1: Gráfica del tratamiento del Grupo 1



Ilustración 2: Gráfica del tratamiento del Grupo 2



Ilustración 3: Gráfica del tratamiento del Grupo 3

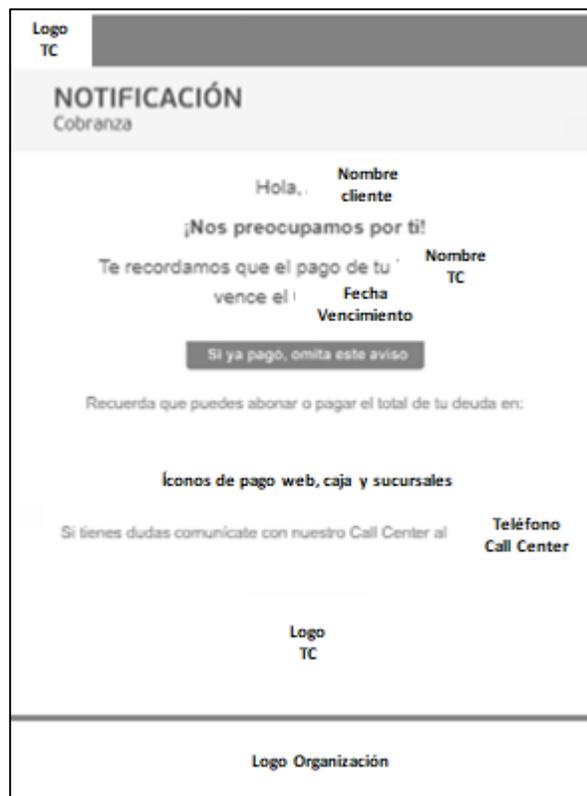


Ilustración 4: Gráfica del tratamiento del Grupo 4

i.vi Asignación aleatoria en bloques.

Gerber y Green (2012) describen la asignación aleatoria en bloques o Block Random Assignment (BRA) en inglés como un procedimiento donde los sujetos están divididos en subgrupos (llamados bloques) y la asignación aleatoria completa ocurre dentro de cada bloque, asegurando, que se encuentre la cantidad deseada de individuos por tratamiento dada ciertas características por la que se ha dividido. En consecuencia, la aleatorización en bloques se hace cargo de dos preocupaciones estadísticas relevantes. Primero, reduce la variabilidad de la muestra eliminando la posibilidad de desbalance (donde por casualidad se ha dado que en el tratamiento se asignaron sujetos con respuestas a favor del tratamiento, por ejemplo). Segundo, esta técnica asegura que ciertos subgrupos estén disponibles para análisis por separado, evitando estimaciones imprecisas (tratamientos compuestos por subgrupos desbalanceados a nivel de observables, por ejemplo).

En la presente investigación, la asignación de los clientes a los tratamientos se hace mediante la técnica de asignación aleatoria en bloques. Las covariables de Monto Total Facturado (del Estado de Cuenta a recordar), Monto Mínimo (del Estado de Cuenta a recordar), Porcentaje de pago del mes anterior (con respecto a su Monto Total Facturado) y Cantidad de meses que no ha pagado el Monto Total Facturado en los últimos 12 meses (de ahora en adelante, se referenciará a esto como Dificultad Transaccional) son las que se utilizarán para dividir en bloques. Dado que son cuatro tratamientos, se crearán cuartiles de éstas. La utilización de estas se debe obtener grupos balanceados y comparables en este nivel de covariables, asegurando estimaciones precisas y reduciendo

la variabilidad de la muestra además de que se igualen las medias poblacionales de los observables y también de los no observables.

ii. Efectos a estudiar e Hipótesis.

Como se ha dicho anteriormente, se busca saber si existe algún efecto en la mora de los clientes recordados el hecho de exponerlos a montos de pago de su deuda. Por lo que, al comparar los tratamientos 1, 2 y 3 (Ilustraciones 1, 2 y 3 respectivamente) contra el tratamiento 4, se tiene la hipótesis que la mora será menor en los grupos que mantienen cifras numéricas por el sesgo cognitivo econométrico de Anclaje (Anchoring en inglés) y Ajuste (Kahneman & Tversky, 1982), sesgo donde el juicio y por ende las decisiones de las personas varían, siendo influenciadas por valores numéricos. En otras palabras, se cree que las personas expuestas a valores numéricos ajustarán su pago a estos, en contraste con las que no fueron expuestas al Anclaje y Ajuste. Así, para personas que iban a pagar ciertos montos, o no iban a pagar, los valores y consecuencias de cada tratamiento deberían ayudar a que disminuya la tasa de morosidad.

Siguiendo la idea anterior y añadiendo el hecho que las personas varían sus conductas de pago cuando son expuestas al Monto Mínimo y Facturado (Navarro-Martinez, Salisbury, Lemon, Stewart, Matthews & Harris, 2011; Keys & Wang, 2018), cuando se comparan los grupos 1, 2 y 3 por separados contra el grupo 4, se tienen las hipótesis que el efecto será mayor en el grupo que muestra ambos montos (grupo 1), luego el grupo que ancla el monto mínimo (grupo 2) y finalmente el grupo que ancla el monto total facturado (grupo 3), dada lo restrictivo que pueden interpretarse estos y su información asociada.

iii. Universo a contactar e implementación del piloto

Se seleccionan clientes a partir de los siguientes criterios:

- ❖ Clientes que no hayan pagado el Monto Total Facturado al menos dos veces en los últimos doce meses.
- ❖ Que presenten una tasa de lectura, de correos enviados por parte de la organización en los últimos seis meses, de al menos un 20%.
- ❖ Que hayan recibido al menos seis correos en los últimos seis meses.
- ❖ Que su correo electrónico se pueda contactar (de cara a la organización).
- ❖ Que el correo electrónico este asignado exclusivamente a un cliente.
- ❖ Que el Monto Mínimo (del Estado de Cuenta que será recordado previo a su fecha de vencimiento) sea de al menos 1,000 CLP (por lo que serán productos que admitan el pago de un monto distinto al Monto Total Facturado).
- ❖ Que el Monto Total Facturado (del Estado de Cuenta que será recordado previo a su fecha de vencimiento) sea de al menos 5,000 CLP.
- ❖ Si el cliente tiene más de un producto (tarjeta de crédito), será recordado el Estado de Cuenta de su producto principal.
- ❖ Productos que no sean pagados automáticamente.
- ❖ Clientes que no sean trabajadores del holding al que pertenece la organización.
- ❖ Clientes que no sean directores ni familiares de éstos del holding al que pertenece la organización.

Estos criterios fueron seleccionados para aumentar el poder estadístico de la intervención. Aplicando estos filtros a la cartera total de clientes de la organización, queda un universo de 179,911 clientes a contactar.

A modo de explorar los datos y verificar si la asignación a los cuatro tratamientos cumplió su propósito, se observa en primer lugar un test de ANOVA, y así comprobar estadísticamente que los grupos tratados no presentan diferencias estadísticamente significativas (en otras palabras, que las medias poblacionales no difieran estadísticamente). En la Tabla 1 se puede ver el resultado del test indicando que, tanto los observables utilizados para el BRA como el resto, no presentan diferencias estadísticamente significativas.

Variable	P-Value	Uso para BRA
Monto Mínimo	0.81	Sí
Monto Total Facturado	0.91	
Porcentaje de pago anterior	0.22	
Dificultad Transaccional	0.83	
Monto Total Facturado anterior	0.79	No
Correos recibidos	0.93	
Correos abiertos	0.44	
Tasa de lectura	0.65	
Monto de renta	0.66	
Edad	0.41	
Sexo	0.83	
Cantidad de hijos	0.76	
Cantidad de vehículos	0.92	

Tabla 1: Test de ANOVA para respaldar Blocked Randomization Assignment.

En último lugar, se observaron las distribuciones de las variables de manera gráfica, las cuales se pueden ver que distribuyen homogéneamente entre grupos. A modo de ejemplo, se muestran las distribuciones de las variables Dificultad Transaccional (recordar, como se dijo más arriba, que equivale a la cantidad de meses que no pagó su MTF en los últimos 12 meses) y Tasa de lectura. El histograma de Sexo se encuentra en el Anexo A.

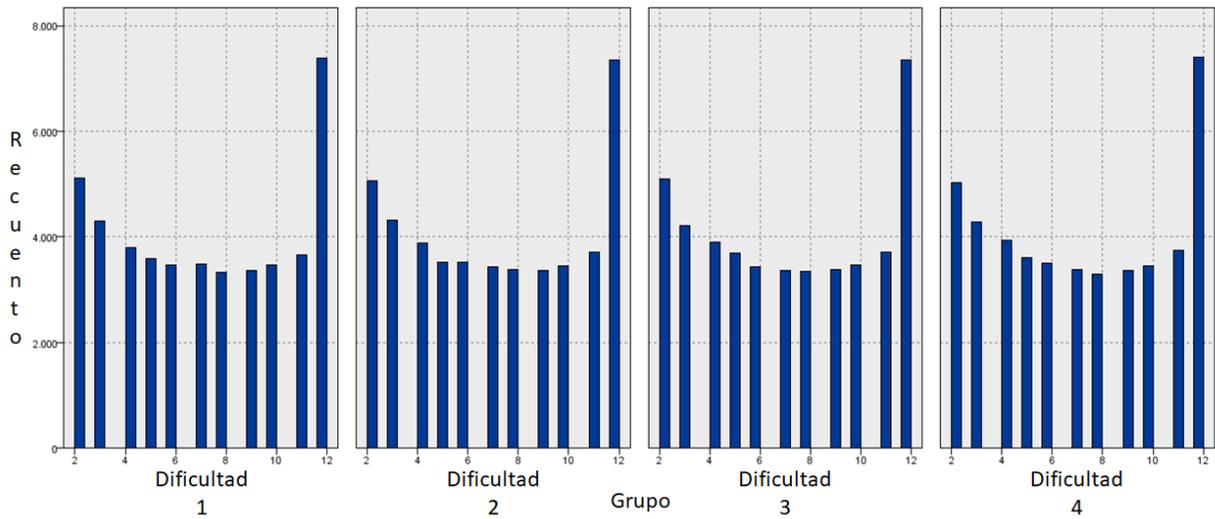


Ilustración 5: Histogramas de Dificultad Transaccional de los cuatro grupos.

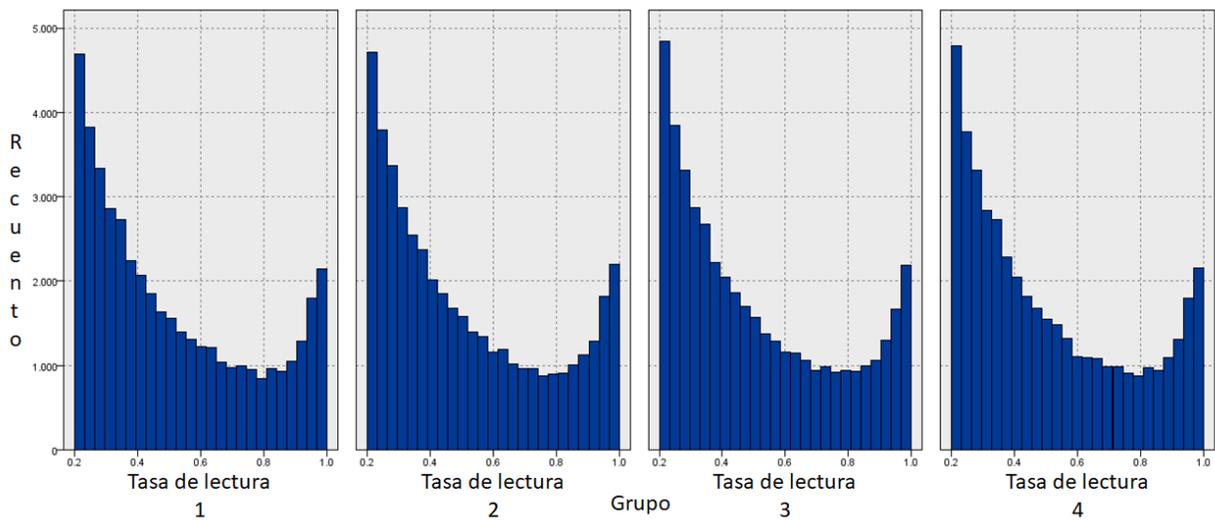


Ilustración 6: Histogramas de Tasa de lectura de los cuatro grupos.

Finalmente, la asignación de clientes a los cuatro tratamientos se puede ver resumida en la Tabla 2.

Tratamiento	Clientes
Ambos montos	44,978
Monto Mínimo	44,977
Monto Total Facturado	44,978
Sin montos	44,978
Total clientes	179,911

Tabla 2: Resumen de cantidad de sujetos por tratamiento.

iv. Seguimiento del tratamiento

De los 179,911 sujetos seleccionados, 178,779 recibieron el tratamiento (99.4%) y presentaron una tasa de lectura promedio del 72.0%. A continuación, la Tabla 3 resume la tasa de lectura promedio por grupo, donde se puede observar que presentan tasas de lecturas similares.

Tratamiento	Tasa de lectura
Ambos montos	71.7%
Monto Mínimo	71.8%
Monto Total Facturado	72.2%
Sin montos	72.3%
General	72.0%

Tabla 3: Resumen de las tasas de lectura del tratamiento agrupadas por grupo.

Complementando lo anterior y recordando que el tratamiento consta de dos envíos (cinco y dos días previo a la fecha de vencimiento del pago), es que se muestra la Tabla 4 que resume la tasa de lectura promedio, las entregas, aperturas únicas y tasas de lectura, por grupo y envío.

Grupo	Envío	Entregados	Aperturas	Tasa de lectura
Ambos montos	Primer envío	44,706	32,717	73.2%
	Segundo envío	44,707	31,357	70.1%
Monto Mínimo	Primer envío	44,682	32,654	73.1%
	Segundo envío	44,697	31,556	70.6%
Monto Total Facturado	Primer envío	44,688	32,839	73.5%
	Segundo envío	44,686	31,674	70.9%
Sin Montos	Primer envío	44,689	32,973	73.8%
	Segundo envío	44,689	31,623	70.8%
Total		357,544	257,393	72.0%

Tabla 4: Resumen del seguimiento de los tratamientos por grupo y por envío.

Constantemente se ve que el primer envío presenta una mayor tasa de lectura que el segundo envío. Esto se puede relacionar a que el segundo envío del tratamiento pierde el “factor sorpresa” que presenta el primer envío.

v. Causal Forest

Se puede decir que el inicio de esta herramienta se encuentra en el trabajo de Athey e Imbens (2016) donde los autores critican las divisiones de muestras convencionales (que, como reportan, se preocupan de controlar el sesgo y la varianza) y muestran como los árboles de decisión que utilizan la validación cruzada convencional entregan estimaciones inconsistentes. Es aquí donde nace el concepto de honestidad, que consiste en utilizar una sub-muestra de los datos para construir el modelo (construir el árbol dividiendo el espacio de covariables según heterogeneidad) y otra sub-muestra para estimar los efectos del tratamiento. Muestran que este método conlleva a estimaciones imparciales y asintóticamente normales. Junto a lo anterior, para garantizar intervalos de confianza de

calidad, se penalizan nodos terminales pequeños (hojas) de alta varianza que implican intervalos de confianza de alto ruido. Los árboles que utilizan este método pasaron a llamarse árboles causales u honestos.

Es en una segunda publicación (Athey & Wager, 2017) donde se pasa a la idea de árboles causales a bosques causales o Causal Forest en inglés. Como breve descripción, esta herramienta se construye en base a árboles honestos y cada árbol honesto se construye en base a sub-muestras aleatorias del conjunto de datos original (no desperdiciándose ningún dato ya que cada árbol utilizará una parte de los datos). Estos bosques demostraron tener estimaciones consistentes y asintóticamente normales.

Con respecto a cómo se construye el modelo, Athey, Tibshirani y Wager (2019) comunican puntos importantes sobre esto. Primero, muestra que los Causal Forest se proyectan como un estimador ponderado localmente, en otras palabras, esto significa que cuando se está estimando el efecto de tratamiento en un sujeto con $X_i = x$, se da más peso a las observaciones más similares. La diferencia de este modelo con respecto a uno tipo K-NN es que la función de pesos Kernel weighting es reemplazada por una Forest-based weights, dando importancia a las variables cercanas a x mientras que uno tipo K-NN da igual importancia a todas las variables.

Lo anterior, supone una gran ventaja cuando existe un ambiente de alta dimensión, y se traduce en que el desempeño será mayor en un Causal Forest al ser capaz de identificar que variables son más importantes para un sujeto determinado, con características x . Mejor aún será el desempeño de un Causal Forest cuando se está en ambiente de alta dimensión y además gran parte de la variación en los efectos se encuentra en unas pocas covariables.

Lo anterior también es una ventaja en comparación a modelos de regresión, pues evita caer en un problema de sobreajuste al ser capaz de trabajar con una gran cantidad de variables y discriminar según importancia (mientras que una regresión no).

Otro punto importante que se menciona es que cada árbol que compone al bosque es honesto (como se mencionó más arriba). Esto junto con los siguientes tres supuestos (los cuales se cumplen dado el diseño experimental), implica que las estimaciones del Causal Forest serán insesgadas y tendrán intervalos de confianza válidos.

1. Data iid.
2. Unconfoundedness: Este supuesto indica que la respuesta del sujeto no tiene relación con su asignación a alguno de los tratamientos, dado sus observables. Esto se cumple pues la asignación de tratamientos fue aleatoria. Matemáticamente se ve como:

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

Ecuación 2: Supuesto de Unconfoundedness.

3. Overlap: Este supuesto indica que todos los individuos (en este caso) están asignados a alguno de los tratamientos. Matemáticamente:

$$\forall x \in \text{supp}(X), \quad 0 < P(W = 1|X = x) < 1$$

Ecuación 3: Supuesto de Overlap.

Por otro lado, un modelo Causal Forest puede ser entendido como un método basado en bosques e inspirado por el R-Learner. Para seguir entendiendo, se describe el R-Learner en su forma general, definiendo previamente:

$$e(x) = P[W = 1|X = x] \quad \text{y} \quad m(x) = E[Y|X = x]$$

Ecuación 4: Propensión a ser tratado y función de respuesta, respectivamente.

Definido lo anterior, el R-Learner consiste en:

- i. Mediante algún método (en este caso, Regression Forest), estimar las funciones de respuesta $\hat{e}(x)$ y $\hat{m}(x)$.
- ii. Minimizar el “squared loss”, motivado por (Robinson, 1988) [14], usando cross-fitting para los componentes tipo “nuisance”. Por lo tanto, se define el R-loss:

$$\hat{\tau}(\cdot) = \underset{\tau}{\text{argmin}} \sum_{i=1}^n \left((Y_i - \hat{m}^{(-i)}(X_i)) - \tau(X_i) (W_i - \hat{e}^{(-i)}(X_i)) \right)^2 + \Lambda_n(\tau(\cdot))$$

Ecuación 5: Función de pérdida R-loss.

Donde el exponente $(-i)$ indica que son predicciones tipo out-of-bag (en simples palabras, las predicciones son resultado de árboles cuyos datos de entrenamiento no incluyen a la observación i) y Λ_n es algún regularizador.

Robinson (1988) mostró que cuando $\tau(x) = \tau$ es constante, ejecutar Ordinary Least Squares (OLS) para minimizar el R-loss (Ecuación 5) logra una eficiencia semiparamétrica. Nie y Wager (2017) amplían este resultado a $\tau(\cdot)$ no paramétricos con fines de estimar los efectos heterogéneos de los tratamientos.

Basado y motivado en lo anterior, un modelo tipo Causal Forest primero ejecuta dos Regression Forest para estimar $\hat{e}(\cdot)$ y $\hat{m}(\cdot)$. Luego, realiza predicciones tipo OOB (usando los dos primeros bosques) para calcular:

$$\hat{\tau}(x) = \frac{\sum_{i=1}^n \alpha_i(x) (Y_i - \hat{m}^{(-i)}(X_i)) (W_i - \hat{e}^{(-i)}(X_i))}{\sum_{i=1}^n \alpha_i(x) (W_i - \hat{e}^{(-i)}(X_i))^2}$$

Ecuación 6: Predicción CATE tipo OOB del Causal Forest

Con $\alpha_i(x) = \frac{1}{B} \sum_{b=1}^B \frac{1_{\{X_i \in L_b(x), i \in B\}}}{|\{i: X_i \in L_b, i \in B\}|}$ que corresponde a los pesos adaptativos usando random forests (Forest-based weights), B la cantidad total de árboles del bosque, $L_b(x)$ la hoja donde x ha caído del árbol b . Por último, y en este trabajo de memoria, los parámetros propios del bosque se escogieron mediante validación cruzada sobre el R-loss (Ecuación

5), con fin de minimizarlo lo máximo posible. Esto último se profundizará en la sección vi del Capítulo 2.

Ahora, para calcular el efecto promedio con las estimaciones del tratamiento o Average Treatment Effect en inglés (ATE) de algún grupo de observaciones I_q determinado, se utiliza el método Augmented Inverse-Propensity Weighted Average Treatment Effect (AIPW ATE), que se puede entender como promediar los puntajes de manera doblemente robusto, como se ve a continuación (notar que $\hat{\tau}^{(-i)}(X_i)$ y $\hat{e}^{(-i)}(X_i)$ son estimaciones tipo OOB):

$$\frac{1}{|I_q|} \sum_{i \in I_q} \hat{\tau}^{(-i)}(X_i) + \frac{W_i - \hat{e}^{(-i)}(X_i)}{\hat{e}^{(-i)}(X_i)(1 - \hat{e}^{(-i)}(X_i))} (Y_i - \hat{\mu}_{W_i}(X_i))$$

Ecuación 7: Augmented Inverse-Propensity Weighted Average Treatment Effect.

Finalmente, para medir la importancia de las variables se utilizará la siguiente medida, la cual es propia del modelo Causal Forest:

$$imp(x_j) = \frac{\sum_{k=1}^4 \left[\frac{\sum_{all\ trees} \text{number depth } k \text{ splits on } x_j}{\sum_{all\ trees} \text{total number depth } k \text{ splits}} \right] k^{-2}}{\sum_{k=1}^4 k^{-2}}$$

Ecuación 8: Importancia de variable x_j del modelo Causal Forest.

La cual se entiende como un recuento de la proporción de divisiones en la variable x_j , con una profundidad máxima a considerar de 4 y una ponderación específica de profundidad de nivel 2 (exponente de descomposición).

vi. Implementación de la estimación de efectos con Causal Forest

Siguiendo el espíritu de Athey y Wager (2019), primero se entrena un bosque causal con los valores de ajuste predeterminados (explicitados en el Anexo B y que muestran un desempeño razonable, según los autores), con fin de encontrar variables relevantes y descartar variables irrelevantes. Luego, se entrena un segundo bosque que tendrá como objetivo realizar las predicciones finales. Este segundo bosque causal se caracteriza por utilizar las variables filtradas gracias al bosque anterior y además ajustar sus parámetros vía Cross-Validation sobre el R-Loss (Ecuación 5), minimizándolo lo máximo posible.

Los parámetros del bosque que son ajustables son:

- ❖ Sample Fraction. Fracción de la data usada para construir cada árbol honesto.
- ❖ Mtry. Número de variables que se intentan para cada división (splits).
- ❖ Minimum Node Size. Objetivo del número mínimo de observaciones in cada hoja de los árboles.
- ❖ Honesty Fraction. Fracción de la data que será usada para determinar las divisiones (splits) de los árboles honestos.
- ❖ Honesty Prune Leaves. Poda de hojas del árbol de la muestra que no tiene hojas vacías.
- ❖ Alpha: Parámetro que controla el máximo del desbalance de una división (split).

- ❖ Imbalance Penalty: Parámetros que controla que tan duras son penalizadas las divisiones (splits) desbalanceadas.

vii. Best Linear Predictor

Para testear la heterogeneidad y evaluar la calidad de las estimaciones del Causal Forest se utiliza el Best Linear Predictor (BLP) propuesto por Chernozhukov, Demirer, Duflo y Fernández-Val (2018). Este sigue la siguiente fórmula (con $\bar{\tau} = \frac{1}{n} \sum_{i=1}^n \hat{\tau}^{(-i)}(X_i)$ donde, de nuevo, el exponente $(-i)$ corresponden a estimaciones tipo OOB):

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

Ecuación 9: Best Linear Predictor.

El modelo tipo OLS ponderado anterior entrega las siguientes interpretaciones. Primero, β_1 habla del desempeño del modelo al detectar el ATE (mientras más cercano a 1, las estimaciones son más correctas). Segundo, β_2 mide la calidad de las estimaciones para detectar los HTE (al igual que el indicador anterior, mientras más cercano a 1, más certero es), por lo que si β_2 es positivo y significativo, el modelo Causal Forest (modelo de esta tesis) encuentra asociaciones productivas entre $\hat{\tau}^{(-i)}(X_i)$ y $\tau(X_i)$, por lo que el p-value del coeficiente β_2 se puede utilizar como un test para verificar si el modelo ha capturado la heterogeneidad subyacente (así que si β_2 es significativo y mayor que 0, se rechaza la hipótesis nula de no heterogeneidad).

Capítulo 3. Resultados

Los resultados son en base a comparar las condiciones con montos y consecuencias versus la condición 4, que sólo recordaba el pago, estudiándose entonces el efecto en la mora de los sujetos al agregar algún monto en el recordatorio de pago (para ser explícitos, el tratamiento es que se agregue un monto en el recordatorio, visitar Capítulo 2, sección i para recordar las condiciones experimentales).

i. Average Treatment Effect (ATE)

A modo de entender el efecto de agregar algún monto a un recordatorio de manera general, es que se estudia el ATE. Éste mostró un efecto positivo en los sujetos (pues en promedio se ha disminuido la mora) y significativo al 95% de confianza, como se ve en la Tabla 5. Además, en la Ilustración 7 se puede ver el histograma de este efecto.¹

ATE	Intervalo de confianza 95%
-0.83%	[-1.16,%, -0.50%]

Tabla 5: ATE y su intervalo de confianza al 95%.

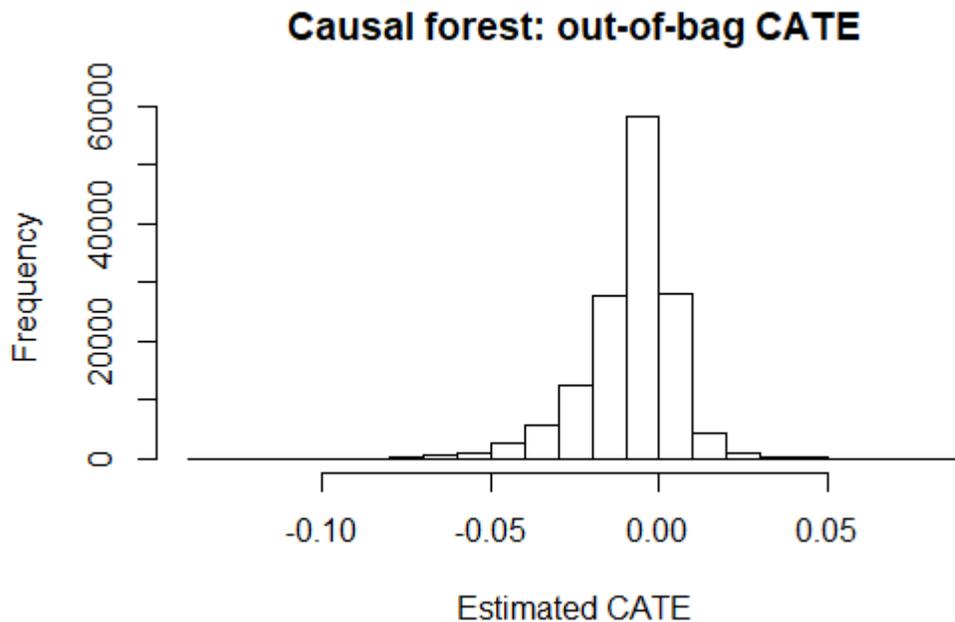


Ilustración 7 : Histograma de las estimaciones del CATE tipo OOB

A modo de recordar, a cada sujeto i se le estima su $OOB\ CATE_i$ (Ecuación 1). Dado que, como se dijo anteriormente, se observa una variable binaria como variable dependiente Y (1: sujeto i no paga o paga menos que el mínimo, 0: si no), el valor $CATE_i$ del sujeto i , se interpreta como el cambio estimado en la probabilidad de Y asociado con el tratamiento W . Por ende, el valor $E[Y_i(1) - Y_i(0)|X_i = x]$ para el sujeto i se encontrará entre -1 y 1.

¹ Tal como se adelantó en el Capítulo 2, sección vi, los parámetros ajustados vía Cross-Validation se encuentran en el Anexo B.

ii. Evaluación de heterogeneidad

Ya habiendo visto que existe un ATE práctico y estadísticamente significativo, y también que la Ilustración 7 muestra un efecto esparcido, ahora el foco es abordar su heterogeneidad. Sin embargo, previamente es bueno verificar si las estimaciones son consistentes, o en otras palabras, si existe una relación correcta entre $\hat{\tau}^{(-i)}(X_i)$ y $\tau(X_i)$ y por ende, la búsqueda de la heterogeneidad en base a las predicciones sea acertada.

Para el propósito anterior (y como se adelantó en el Capítulo 2, sección vii), se calculará el BLP. El resultado de este test de calibración son coeficientes positivos y significativos, lo que nos diría que las estimaciones del Causal Forest $\hat{\tau}^{(-i)}(X_i)$ son consistentes con $\tau(X_i)$. Detallando ahora con cada coeficiente, se ve que el desempeño del bosque es bueno en relación con las predicciones promedio (β_1) y algo menor, sin embargo correcto, para detectar la heterogeneidad adecuadamente (β_2). Además, el hecho de que β_2 sea significativo, indica que vale la pena evaluar la heterogeneidad detectada y presente en los datos (en otras palabras, se puede rechazar la hipótesis nula de no heterogeneidad). Los resultados del test se encuentran en la Tabla 6. Notar que (***) indica p-value<0.001 mientras que (*) indica p-value<0.05

Coficiente	Estimación	P-value	Significancia
β_1	0.99	<0.001	(***)
β_2	0.32	0.028	(*)

Tabla 6: Resultado del test de calibración Best Linear Predictor.

iii. Heterogeneidad entre subgrupos

Con el fin de detectar patrones de heterogeneidad, se crean cuartiles en base a las predicciones realizadas por el Causal Forest. Luego, se calcula el ATE para cada cuartil, esto mediante el método AIPW ATE (descrito en Capítulo 2, sección v) y mediante el Sample ATE el cual que se entiende como, para cada cuartil, una regresión lineal controlando por el tratamiento (en otras palabras, una diferencia entre el resultado promedio para los sujetos tratados y los no tratados, dentro del cuartil). Los resultados de ambos métodos se pueden ver juntos en la Tabla 7 (entre paréntesis las desviaciones estándar).

Cuartil	Valor cuartil q	Sample ATE	AIPW ATE
1	\overline{ATE}	-0.0184	-0.0185
	SD	(0.0035)	(0.0044)
2	\overline{ATE}	-0.0069	-0.0060
	SD	(0.0035)	(0.0031)
3	\overline{ATE}	-0.0036	-0.0051
	SD	(0.0035)	(0.0025)
4	\overline{ATE}	-0.0019	-0.0036
	SD	(0.0035)	(0.0030)
P-value		0.003 (**)	<0.001 (***)

Tabla 7: Sample ATE y AIPW ATE por cuartil.

La Tabla 7 además de adjuntar el OOB ATE por cuartil con ambos métodos, contiene el p-value de dos test. Para la columna Sample ATE es un F-test y para la columna AIPW ATE es un Wald test. Ambos test contienen como hipótesis nula H_0 : ATE es constante entre los subgrupos. Por lo que se puede decir, con evidencia estadística, que se rechaza la hipótesis nula que las estimaciones OOB CATE son constantes entre los cuartiles, por lo que existen grupos que presentan un mayor efecto que otros. Notar que (***) indica $p\text{-value} < 0.001$ mientras que (**) indica $p\text{-value} < 0.01$. Gracias a esto, se procede a investigar si existen otras diferencias entre estos subgrupos. La información de la Tabla 7 puede verse resumida en el siguiente gráfico.

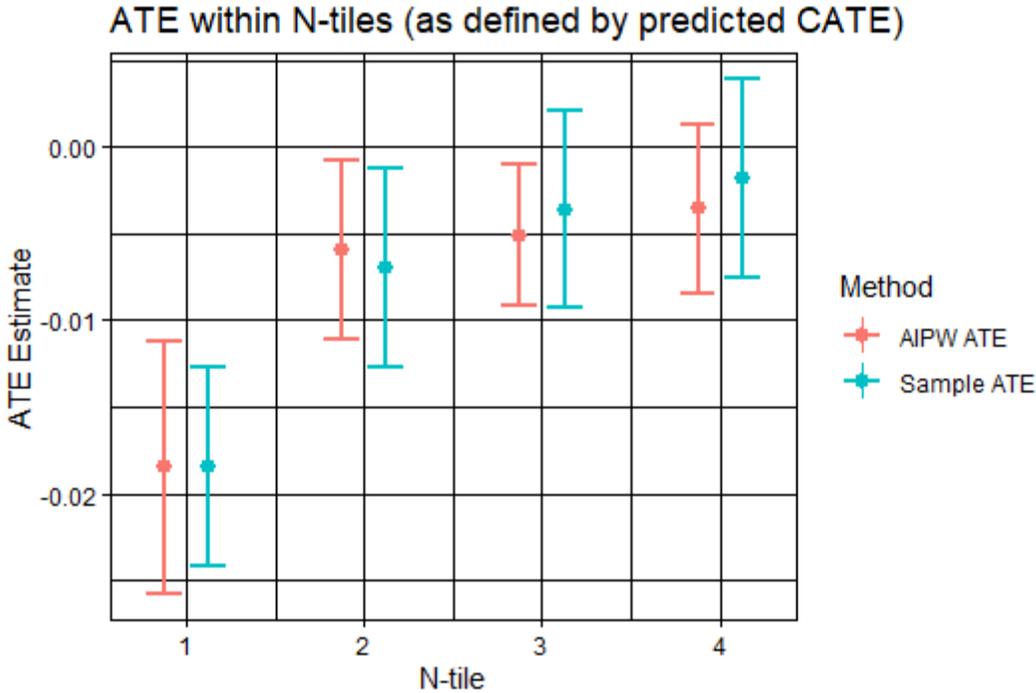


Ilustración 8: ATE más intervalos de confianza al 90% por cuartil, mediante AIPW ATE y Sample ATE.

Es importante mencionar que, por cuartil, se encuentra el ATE de cada subgrupo junto con su intervalo de confianza correspondiente al 90% (empleando ambos métodos). Notando donde está el valor 0 en el eje, se pueden dar dos observaciones: i) Los dos grupos que presentan un mayor ATE son estadísticamente significativos al 90% de nivel de confianza, considerando ambos métodos (y serían 3 de los 4 cuartiles que se logró un efecto positivo para disminuir la mora según el Causal Forest); ii) y más importante aún, se puede decir que no se ha realizado ningún efecto contraproducente (aumentar la mora) en ninguno de los cuartiles, lo que es muy positivo en la práctica.

iv. Heterogeneidad entre covariables

El propósito de encontrar heterogeneidad entre covariables es poder caracterizar los grupos definidos en la sección anterior y también encontrar qué variables guardan mayor relación con el efecto. Para esto, se escogerán las variables que para el bosque fueron más importantes y las que más variaron más entre los cuartiles. Luego, las variables

seleccionadas se analizan por separadas con un test ANOVA para avalar con evidencia estadística la diferencia (o no) entre los cuartiles en base a los observables seleccionados (recordando que la hipótesis nula de un test ANOVA es que las medias de los cuartiles son las mismas, mientras que la hipótesis alternativa es que alguna de las medias es diferente).

Según lo anterior, las variables seleccionadas (en orden de relevancia para el modelo) son la desviación estándar del porcentaje de pago del último cuatrimestre ($X1$), la cantidad de veces que no pagaron el monto mínimo en los últimos 12 meses ($X2$), la cantidad de veces que no pagaron el monto mínimo específicamente en el último cuatrimestre ($X3$) y el promedio del porcentaje de pago del último cuatrimestre ($X4$). Todas previas al pago que fue recordado, claramente. Esto concuerda con las hipótesis que se tenían previamente, donde el comportamiento del pasado es el que explica el comportamiento del futuro.

Un valor alto de $X1$ indica que el sujeto presenta un rango de porcentajes de pago amplios, por ejemplo, a veces paga e 100% de su deuda, otras veces no paga y a veces una fracción (mientras que un valor bajo de $X1$ indica que el sujeto paga casi siempre la misma cantidad de su deuda, pudiendo ser la totalidad, nada, monto mínimo u otra fracción). Por su lado, $X2$ y $X3$ nos indican la conducta de no-pago de la persona con respecto al monto mínimo, mientras más cercano a 12 y 4 respectivamente, significa que reiteradas veces cae en mora (en cambio, mientras más cercano a 0, significa que la persona se caracteriza por presentar una conducta constante de no caer en mora). Finalmente, $X4$ se debe interpretar como el porcentaje de pago característico del sujeto, por lo que si este es más cercano a 1, significa que el individuo se caracteriza por pagar constantemente la totalidad de su deuda (mientras que si es más cercano a 0 significa que la persona constantemente no paga su deuda).

En la Tabla 8 se puede ver el valor de $X1$, $X2$, $X3$ y $X4$ en cada cuartil (junto con su desviación estándar entre paréntesis) y el valor del p-value del test ANOVA para cada variable. Rápidamente se observa que se rechaza la hipótesis nula de igualdad de medias y, viendo los valores por cuartil junto con seguir las interpretaciones del párrafo anterior, los sujetos que presentaron un mayor efecto de no caer en mora (en otras palabras, un efecto más negativo hacia la mora) son quienes presentan un comportamiento de pago disperso y bajo (según su porcentaje de pago) y han caído en mora más veces en el último año y en el último cuatrimestre.

Variable	Valor	Cuartil 1	Cuartil 2	Cuartil 3	Cuartil 4	P-value
$X1$	Media	0.285	0.127	0.073	0.098	<0.001(***)
	SD	(0.001)	(0.001)	(0.001)	(0.001)	
$X2$	Media	2.067	0.770	0.388	0.691	<0.001(***)
	SD	(0.008)	(0.008)	(0.008)	(0.008)	
$X3$	Media	0.747	0.207	0.101	0.200	<0.001(***)
	SD	(0.003)	(0.003)	(0.003)	(0.003)	
$X4$	Media	0.617	0.777	0.867	0.774	<0.001(***)
	SD	(0.002)	(0.002)	(0.002)	(0.002)	

Tabla 8: Anova test y valores de variables $X1$, $X2$, $X3$ y $X4$.

Ahora se realiza un test de Tukey HSD (Tukey Honest Significant Differences) y apoyándose en los valores de la Tabla 8 para comprender qué cuartil o qué cuartiles

difieren de otros y en qué diferencia. Como resultado, se puede ver que para $X1$ y $X2$, todos los cuartiles difieren entre sí con significancia práctica y estadística (realizando diferencias entre pares de cuartiles), mientras que para $X3$ y $X4$ los cuartiles 4 y 2 presentan similitudes en las medias. Tras análisis, se puede concluir que el cuartil 1, que presenta mayor efecto de tratamiento, difiere de los demás presentando características como un porcentaje de pago variante ($X1$) y bajo ($X4$), además de caer en mora reiteradas veces en el pasado (viéndose reflejado en la cantidad de veces que no paga el monto mínimo en $X2$ y $X3$).

Pares cuartiles	X1 adj p-value	X2 adj p-value	X3 adj p-value	X4 adj p-value
2 - 1	<0.001(***)	<0.001(***)	<0.001(***)	<0.001(***)
3 - 1	<0.001(***)	<0.001(***)	<0.001(***)	<0.001(***)
4 - 1	<0.001(***)	<0.001(***)	<0.001(***)	<0.001(***)
3 - 2	<0.001(***)	<0.001(***)	<0.001(***)	<0.001(***)
4 - 2	<0.001(***)	<0.001(***)	0.35	0.38
3 - 4	<0.001(***)	<0.001(***)	<0.001(***)	<0.001(***)

Tabla 9: Resumen test de Tukey HSD.

Dados los resultados anteriores, resulta interesante conocer como varía el CATE predicho (con el cual se construyeron los cuartiles) en las variables identificadas ($X1$, $X2$, $X3$ y $X4$) de manera gráfica.

Se puede ver en la Ilustración 9 que para la desviación estándar del porcentaje de pago ($X1$) la intuición y conclusiones se mantienen, mientras más dispersos son sus porcentajes de pago del sujeto, mayor es el efecto del tratamiento.

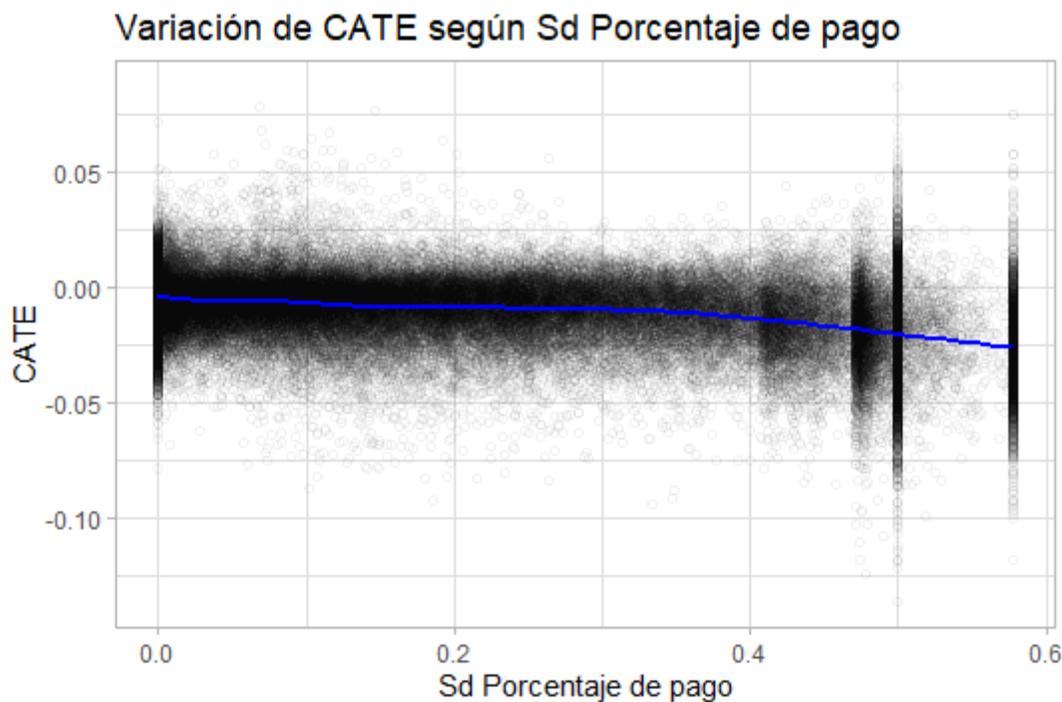


Ilustración 9: Variación de CATE según $X1$.

Ahora, para el promedio de los porcentajes de pago ($X4$) se ve que el efecto es muy cercano a 0 (casi nulo) cuando el sujeto tiende constantemente a no pagar o constantemente paga toda su deuda, mientras varía lejos de 0 (de manera positiva en la práctica) para sujetos con porcentaje de pago promedio distinto a estos valores extremos.

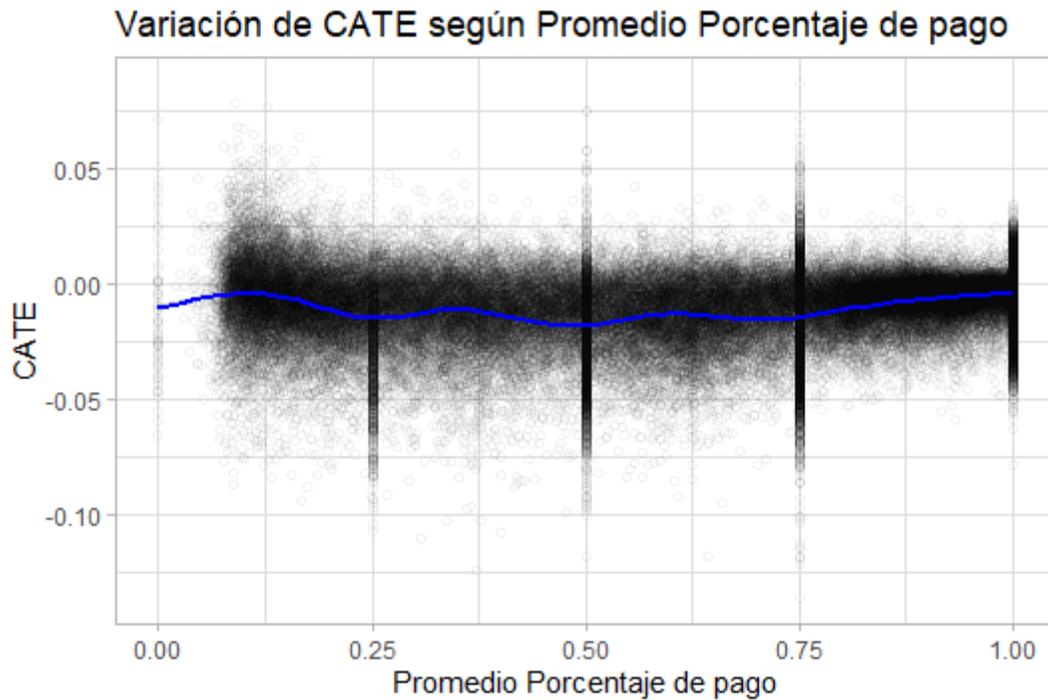


Ilustración 10: CATE variando según $X4$

Ahora, para la cantidad de veces que el sujeto no pagó el monto mínimo (tanto para los últimos 12 meses $X2$, como para los últimos 4 meses $X3$) presentan resultados similares entre sí: mientras más veces ha caído en mora, mayor efecto presenta el sujeto, sin embargo el efecto va disminuyendo mientras presenta una conducta de nunca pagar el monto mínimo en el plazo.

Variación de CATE según Cantidad de veces en mora

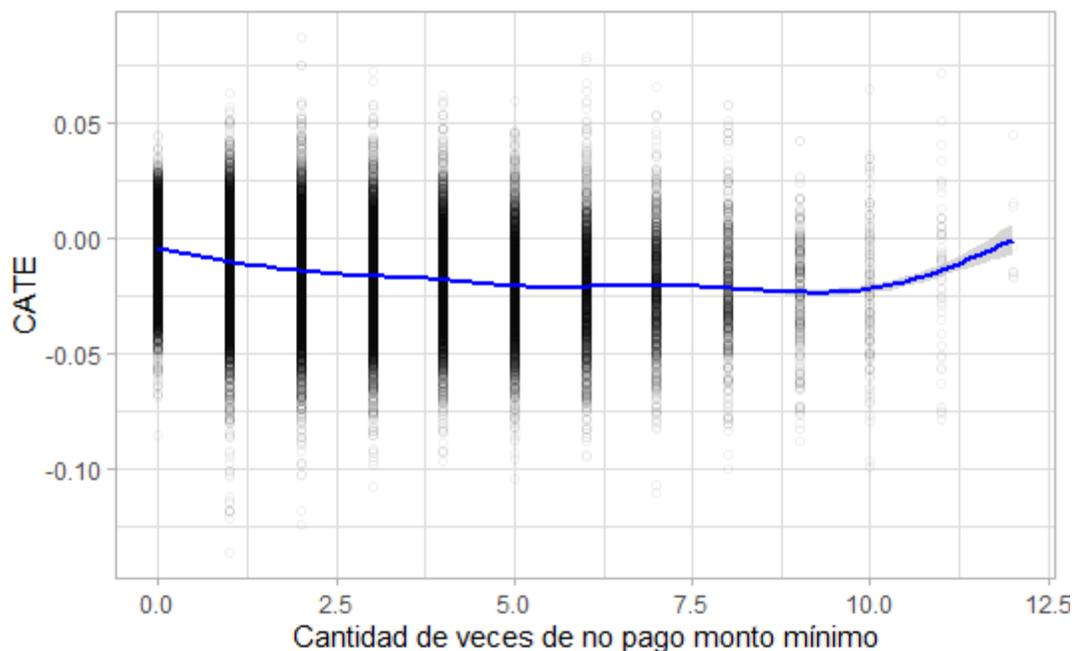


Ilustración 11: variación de CATE según mora en los últimos 12 meses

Dado que X_2 y X_3 miden lo mismo, pero en distintos plazos, se puede ver en qué punto la recurrencia del comportamiento hace que el efecto empiece a disminuir en un plazo largo y otro más corto. En la Ilustración 11 se ve que para X_2 es cuando ha caído 10 veces en mora en 12 meses, en cambio, en la Ilustración 12 se ve que es 3 veces en un plazo de 4 meses (X_3). Esto se puede entender que es difícil “mover” una conducta cuando ha mostrado ser consistente en su comportamiento de caer en mora. La conclusión anterior aplica también para cuando la persona recurrentemente no cae en mora, por lo que cuando X_2 y X_3 son 0, se ve que no hay un efecto del tratamiento, pues estos sujetos no necesitan de un recordatorio de pago para no caer en mora, ya que pagarán de igual manera al menos el monto mínimo.

Variación de CATE según Cantidad de veces en mora

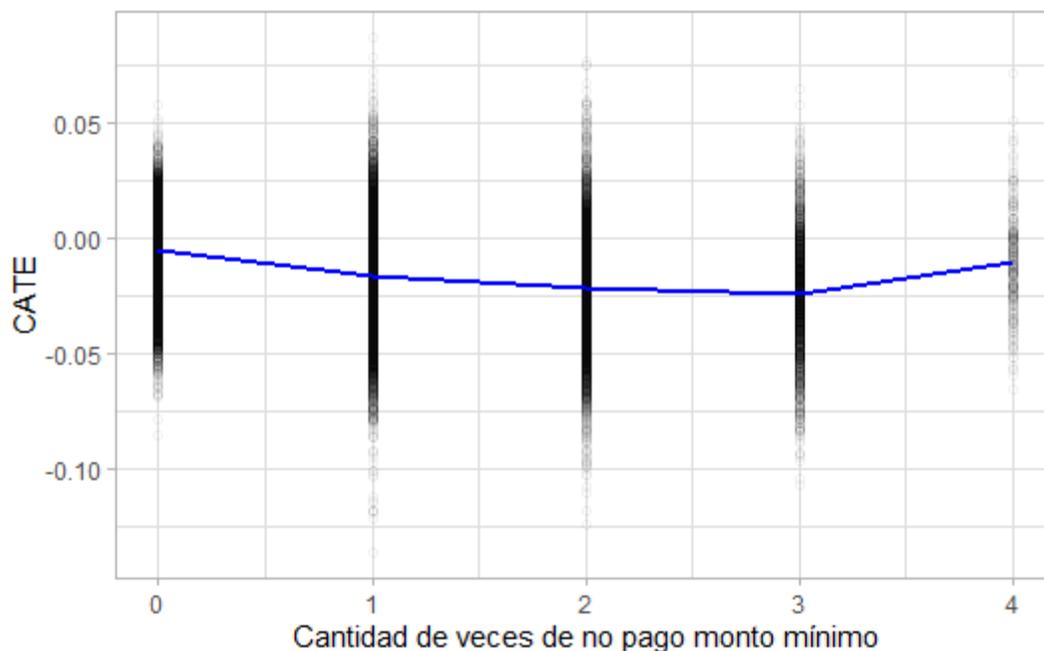


Ilustración 12: variación de CATE según mora en los últimos 4 meses

Los resultados de la búsqueda de heterogeneidad entre covariables, de grupos que presentan heterogeneidad en sus efectos de tratamientos observados, nos permite concluir que existe un importante grupo de sujetos que son sensibles a observar cifras referenciales a su Estado de Cuenta (en contraste a sólo recordar su fecha límite de pago). Además este grupo, que en promedio el tratamiento ha disminuido cerca de un 2% la cantidad de sujetos en mora, presenta características, principalmente conductuales, claramente distintas al resto, pudiéndose caracterizar por presentar pagos dispersos y bajos, junto con haber caído en mora reiteradas veces en el pasado.

v. ¿Qué mensaje enviar y a quién?

En esta última sección de resultados se busca mostrar como poder focalizar el tratamiento, con el fin de obtener la menor mora promedio en los sujetos, por lo que ahora se analizarán los efectos de entregar cada mensaje por separado, comparando nuevamente contra el recordatorio que no muestra algún monto.

Previo a ello, resulta interesante conocer cuál es el efecto promedio de cada recordatorio (contra el recordatorio base), lo cual se muestra en la Tabla 10 y donde se ve como el recordatorio que habla de ambos montos es el de mayor efecto y que todos presentan significancia estadística al 5%.

Recordatorio	ATE	IC al 95%
Ambos montos (1)	-0.87%	[-1.25%, -0.50%]
Monto Mínimo (2)	-0.83%	[-1.20%, -0.46%]
Monto Total Facturado (3)	-0.66%	[-1.03%, -0.28%]

Tabla 10: ATE de los tratamientos 1, 2 y 3 versus el tratamiento 4.

Los resultados de la Tabla 10 son producto de predicciones tipo OOB y honestas. Se ha utilizado el 80% de los sujetos (mencionados en el Capítulo 2, sección iii) para estimar y construir los modelos anteriores (cada cual según el tratamiento que haya recibido). El 20% restante de los sujetos serán utilizados para estimar cuál es el valor de focalizar el tratamiento, por lo que para el individuo i se predice su $CATE_{ij}$ para los 3 modelos (j) de la Tabla 10. Entonces, para focalizar el tratamiento y gracias al hecho que un modelo de Causal Forest que evalúa el efecto del tratamiento j entrega una estimación $CATE_{ij}$ para el sujeto i , o sea, un Individual Treatment Effect (ITE), se procede así: primero, estimar el $CATE_{ij}$ para todos los individuos i en cada tratamiento j ; segundo, para cada recordatorio j seleccionar el grupo de individuos que muestren un efecto positivo frente a la mora y estadísticamente significativo (con el 90% de nivel de confianza); tercero, del grupo seleccionado, algunos recibieron efectivamente el recordatorio j y otros algún otro recordatorio (esto por diseño experimental, Capítulo 2, sección i), por lo que quienes efectivamente recibieron el tratamiento j tienen su tratamiento “optimizado”, mientras que el resto no; cuarto, verificar el efecto promedio de focalizar el tratamiento en la mora, la dimensión y su significancia. Esto se puede ver así en la Tabla 11 como ejemplo:

ID	Value	Treatment 1	Treatment 2	Treatment 3
1	\widehat{CATE}_{1j}	-0.028**	0.013	-0.006
	SE	(0.012)	(0.048)	(0.021)
2	\widehat{CATE}_{2j}	-0.027**	-0.020	0.018
	SE	(0.013)	(0.031)	(0.023)
3	\widehat{CATE}_{3j}	0.002	0.027	0.071
	SE	(0.015)	(0.053)	(0.080)
...

Tabla 11: Ejemplo de selección de tratamiento

De la Tabla 11, se ve que los sujetos 1 y 2 muestran un $CATE_{i1}$ significativo, por lo que habrá que verificar quién, aleatoriamente, recibió el tratamiento 1 y quién algún otro tratamiento (no óptimo). El sujeto 3 no mostró un $CATE_{ij}$ significativo, por lo que no se considera para el análisis de focalización, y así con todos los sujetos del test set.

Dado que se están comparando sujetos similares a los ojos del $CATE_{ij}$, es esperable que aquellos que recibieron el tratamiento óptimo presenten una menor mora promedio que los que recibieron otro tratamiento (cuando debían recibir el óptimo). En efecto, en la Tabla 12 se ve como el ATE de focalizar el tratamiento j muestra una gran mejora en la mora promedio y es estadísticamente significativa en los tratamientos que resultaron mostrar un mayor efecto.

Tratamiento	ATE optimizando tratamiento j	P-value
1	-7.18%	0.036 (*)
2	-5.88%	0.0076 (**)
3	-2.09%	0.57

Tabla 12: Resultado de optimizar el tratamiento en la mora promedio.

Los resultados de la tabla anterior se pueden interpretar como la ganancia de focalizar los diversos tratamientos en gente que es sensible a la personalización de éstos. Por lo que dejan en claro las magnitudes de cuan beneficioso es personalizar la comunicación (en concreto, los tratamientos) de la organización hacia sus clientes (en contraste de asignarlos de una manera no óptima).

Capítulo 4. Conclusión

En este trabajo de fin de grado se ha implementado un experimento de campo aleatorizado de recordatorio de pago y se ha examinado la heterogeneidad de la respuesta de pago en la consecuencia de morosidad de los sujetos mediante el algoritmo de Causal Forest.

Como primera conclusión, se puede ver que existe un efecto positivo y significativo en contra de la mora tras exponer a sujetos a sus montos referenciales de sus pagos de tarjetas de crédito, en contraste con sólo recordarle cuándo vence su deuda. En promedio, se ve que la reducción de la mora promedio es cerca de 1%.

Con fin de encontrar patrones de heterogeneidad de este efecto, se han revisado las variables que más variaron junto con las que sugiere el método de Causal Forest, las cuales coincidían con las hipótesis iniciales y lo que indica la literatura para este tipo de tratamiento: las variables de comportamiento y conducta del pasado son las que explican de mejor forma un efecto mayor del tratamiento. En particular, sujetos que se caracterizaban por presentar pagos bajos (que pagan en promedio bajo el 50% de su deuda), irregulares (la dispersión del porcentaje de pago de la deuda sobre 0,45) y que hayan caído previamente reiteradas veces en mora (entre 5 y 10 veces en el último año), presentan los mayores efectos en promedio en contra de la mora (disminución cerca de un 2,5% estadísticamente significativo al 95%). Desde la Tabla 7 a la Tabla 9 se pueden ver los claros patrones mencionados anteriormente, donde se puede caracterizar al grupo tanto por su efecto predicho como por sus observables (con altos niveles prácticos y estadísticamente significativos) y gráficamente, los patrones de heterogeneidad se pueden observar (cómo es que en promedio varía la predicción del efecto según las covariables mencionadas) desde la Ilustración 9 a la Ilustración 12.

A nivel del conjunto de recordatorios también se examina la heterogeneidad del efecto. Se verifica que estos provocan distintos efectos en promedio y también a nivel individual. Estas diferencias se pueden deber a los distintos comportamientos de pagos de las personas. En específico, exponer a sujetos a ambos montos y sus consecuencias de pago presenta un efecto promedio muy similar que el recordatorio que enfatiza con el monto mínimo (ambos presentando una magnitud cercana al 1,0% en contra de la mora), en contraste a la notificación que destaca sólo al monto total facturado y presenta un efecto cercano al 0.7% (valores inferidos al comparar contra el recordatorio que sólo recuerda la fecha del pago).

Gracias a los resultados y conclusiones anteriores, se pueden validar y afirmar las hipótesis postuladas en el Capítulo 2. Primero, se ve cómo es que está presente el sesgo econométrico de Anclaje y Ajuste, ya que las personas expuestas a cifras numéricas presentaron un comportamiento distinto a las que no (reflejándose en el efecto contra la mora), según lo planteado. Segundo, se afirma la hipótesis en que las notificaciones que presentaban el Monto Mínimo de pago junto a sus consecuencias presentarían un mayor efecto que no mostrarlo, en efecto y como se dijo, fueron donde se observaron mayores disminuciones de mora en promedio. Esto se puede explicar por lo restrictivo y severo que es el castigo (interpretado por el sujeto) de la consecuencia del pago mínimo (en contraste con la del Monto Total Facturado).

Para concluir el cumplimiento de los objetivos y verificar si han sido llevados a cabo de manera satisfactoria, es que se pone a prueba el desempeño de la metodología implementada junto con sus resultados de dos formas. Primero, se utiliza el test Best Linear Predictor, el cual ha indicado que el modelo ha sido capaz de detectar el efecto promedio y los patrones de heterogeneidad de manera satisfactoria. Segundo, una forma práctica (que busca responder qué recordatorio de pago enviar y a quién) de focalizar el tratamiento en grupos específicos que el modelo prediga que los sujetos presentarán un gran efecto dados sus observables (dividiendo la muestra en un set de entrenamiento y otro de testeo), la cual muestra una importante reducción de la mora promedio (cerca de un 7%) en estos grupos que identifica, apuntando entonces a que el método es capaz de reconocer la heterogeneidad de los efectos de tratamiento.

Capítulo 5. Trabajos futuros

Gracias a los resultados obtenidos se pueden contestar preguntas que los gestores de políticas de cobranza preventiva, de la organización donde se desarrolló el trabajo de fin de grado, les resultaba interesante conocer.

Una de estas preguntas es el conocer qué tipo de personas son más sensibles a los recordatorios y consecuencias que se les expone. De esta manera, se podrá utilizar la caracterización encontrada con fin de focalizar el efecto en quienes exista un efecto positivo, optimizando esfuerzo y tiempo implementando tratamientos.

Además, otra pregunta es el conocer qué mensaje enviar y a quién, lo cual se ha logrado mediante una metodología que es capaz de indicar qué mensaje entrega un mayor efecto de tratamiento para un individuo dado (además de ser una metodología altamente generalizable y aplicable a distintas campañas, tratamientos o intervenciones).

Ligado a lo anterior, es altamente recomendable realizar nuevos experimentos aleatorios con el fin de testear los modelos propuestos y así obtener resultados robustos, para conocer cuan positivo es el hecho de focalizar. Ejemplos de lo anterior, son comparar el tratamiento focalizado con un grupo que reciba algún otro tratamiento aleatorio (lo que se muestra en el Capítulo 3, sección v) o con el tratamiento que obtuvo mayor efecto promedio (que sería el tratamiento 1, mostrado también en el Capítulo 3, sección 5).

Además, se recomiendan otras variables dependientes a observar, como por ejemplo el porcentaje de pago de las personas, el pago o no de la deuda total, el no haber pagado absolutamente nada, cambios en el consumo de la persona (transacciones, rubros, montos), utilización de canales digitales (tiempo de uso, veces de ingreso), entre otros.

También se recomienda utilizar otros canales en el envío de mensaje. Esto implica necesariamente variar el mensaje (sumando un desafío a la hora de componer el mensaje) dadas las diversas restricciones que contienen los distintos canales, por ejemplo, cantidad de caracteres en un SMS, espacio de pixeles en un banner de alguna plataforma digital, restricciones legales en una carta, entre otras restricciones.

Finalmente, se recomienda utilizar la misma metodología para evaluar otro tipo de mensajes y/o sesgos, que tengan que ver con el giro de la organización, como por ejemplo mostrar las consecuencias negativas de la mora a largo plazo, mostrar específicamente los montos que evitará pagar de más si no cae en mora, mostrar beneficios de pagos automáticos, etc.

Bibliografía

- [1]. Comisión para el Mercado Financiero. Informe de Tarjetas de Crédito. Junio 2019.
- [2]. Superintendencia de Bancos e Instituciones Financieras. Informe de Endeudamiento 2018. 2018.
- [3]. Superintendencia de Bancos e Instituciones Financieras. Informe de Tarjetas de Crédito. Junio 2018.
- [4]. A. Belloni, V. Chernozhukov, & C. Hansen. High-dimensional methods and inference on structural and treatment effects. *The Journal of Economic Perspectives*, 28(2):29–50. 2014.
- [5]. S. Athey, G. W. Imbens, & S. Wager. Approximate residual balancing: Debiased inference of average treatment effects in high dimensions. 2016.
- [6]. V. Chernozhukov, D. Chetverikov, M. Demirer, E. Duflo, C. Hansen, & W. Newey. Double/Debiased/Neyman Machine Learning of Treatment Effects. 2017.
- [7]. S. Athey & G. Imbens. Recursive partitioning for heterogeneous causal effects. *Proceedings of the National Academy of Sciences*, 113(27):7353–7360. 2016.
- [8]. S. Wager & S. Athey. Estimation and inference of heterogeneous treatment effects using random forests. *Journal of the American Statistical Association*. 2017.
- [9]. S. Athey. *The Impact of Machine Learning on Economics*. 2018.
- [10]. D. Navarro-Martinez, L. Salisbury, K. Lemon, N. Stewart, W. Matthews & A. Harris. Minimum Required Payment and Supplemental Information Disclosure Effects on Consumer Debt Repayment Decisions. *Journal of Marketing Research*. 2011.
- [11]. B. Keys & J. Wang. Minimum Payments and Debt Paydown in Consumer Credit Cards. *Journal of Financial Economics*. 2018.
- [12]. D. Rubin. Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of Educational Psychology*, 66(5), 688-701. 1974.
- [13]. S. Athey, J. Tibshirani & S. Wager. Generalized random forests. *The Annals of Statistics*. 2019.
- [14]. P. Robinson. Root-N-Consistent Semiparametric Regression. *Econometrica*. 1988.
- [15]. D. Kahneman & A. Tversky. The psychology of preferences. *Scientific American*, 246(1), 160–173. 1982.

- [16]. X. Nie & S. Wager. Quasi-Oracle Estimation of Heterogeneous Treatment Effects. 2017.
- [17]. V. Chernozhukov, M. Demirer, E. Duflo, & I. Fernandez-Val. Generic machine learning inference on heterogeneous treatment effects in randomized experiments. Technical report, National Bureau of Economic Research, 2018b.
- [18]. E. Duflo, R. Glennerster & M. Kremer. Using Randomization in Development Economics Research: A Toolkit. T. Schultz and John Strauss, eds., Handbook of Development Economics. Vol. 4. Amsterdam and New York: North Holland. 2007.
- [19]. G. Imbens & D. Rubin. Causal Inference for Statistics, Social and Biomedical Sciences. 2015.
- [20]. Gerber, A. S., & Green, D. P. (2012). Field experiments: Design, analysis, and interpretation. WW Norton.
- [21]. E. O'Neill & M. Weeks. Causal Forest Estimation of Heterogeneous Household Response to Time-Of-Use Electricity Pricing Schemes. 2019.
- [22]. Jonathan M.V. Davis & Sara B. Heller, 2017. "Using Causal Forests to Predict Treatment Heterogeneity: An Application to Summer Jobs," American Economic Review, American Economic Association, vol. 107(5), pages 546-550, May.
- [23]. Suk, Y., Kim, J., & Kang, H. (2019, October 23). Hybrid Causal Forests: Using Random Forests and Finite Mixture Models for Estimating Heterogeneous Treatment Effects in Latent Classes.
- [24]. M. Lechner. Modified Causal Forests for Estimating Heterogeneous Causal Effects. 2018.
- [25]. S. Athey & S. Wager. Estimating Treatment Effects with Causal Forests: An Application. 2019.

Anexo y apéndices

Anexo A. Histograma de Sexo de los cuatro grupos.

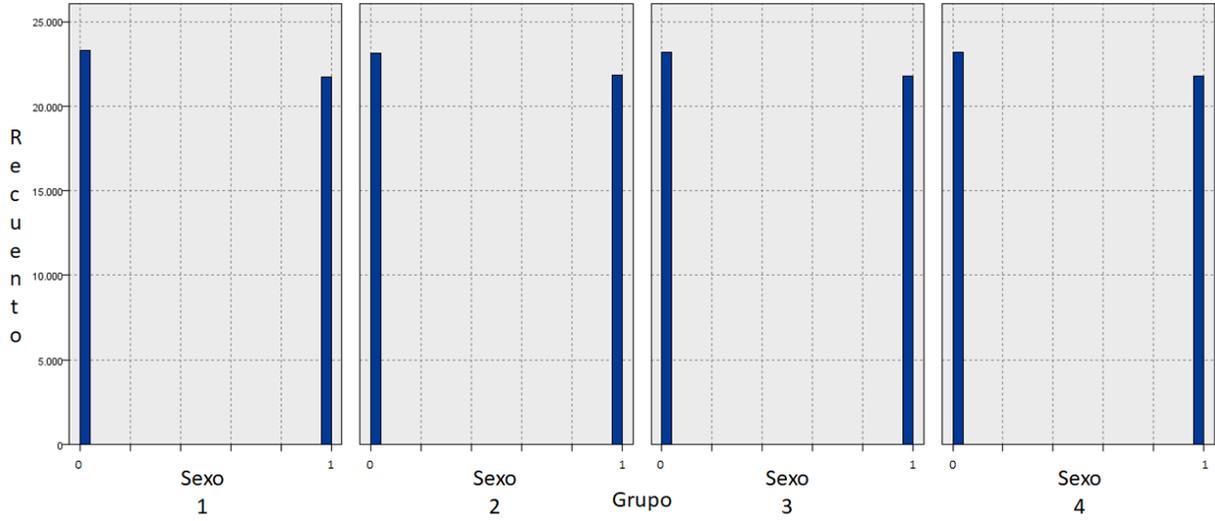


Ilustración 13: Histograma de Sexo entre los cuatro grupos (Anexo).

Anexo B. Tabla que contiene los valores ajustados del método y los métodos predeterminados.

Parámetro	Valor predeterminado	Valor ajustado
Sample fraction	0.500	0.500
Mtry	28	28
Min. Node size	5	3
Honesty fraction	0.500	0.669
Honesty prune leaves	TRUE	FALSE
Alpha	0.500	0.156
Imbalance penalty	0.000	0.002

Tabla 13: Ajustes de parámetros de modelos Causal Forest (Anexo).