



**UNIVERSIDAD DE CHILE
FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS
DEPARTAMENTO DE INGENIERIA INDUSTRIAL**

**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

ANDRÉ CARBONI EWING

PROFESOR GUÍA:
RICARDO MONTOYA MOREIRA

MIEMBROS DE LA COMISIÓN:
LUÍS ABURTO LAFOURCADE
MÁXIMO BOSCH PASSALACQUA
MARCEL GOIC FIGUEROA

SANTIAGO DE CHILE
ENERO 2012

RESUMEN DE LA TESIS PARA OPTAR AL GRADO
DE MAGÍSTER EN GESTIÓN DE OPERACIONES Y
AL TÍTULO DE INGENIERO CIVIL INDUSTRIAL
FECHA: 03/01/2012
PROF. GUÍA: SR. RICARDO MONTOYA

ESTUDIO DE LAS RESPUESTAS DE LOS CONSUMIDORES ANTE UN PROGRAMA DE LEALTAD

La implementación de programas de lealtad ha tenido un importante crecimiento en el retail, sin embargo son escasos los estudios que verifican la existencia de un efecto en el comportamiento de los consumidores. Esta investigación tiene como objetivo estudiar los posibles efectos de un programa de lealtad sobre el comportamiento de compra y canje de sus clientes.

Para ello, se estudia el programa de un importante retail nacional, en el cual los consumidores acumulan puntos según el monto de cada compra, y tienen la posibilidad de canjear un premio al llegar a alguno de los 4 niveles de canje disponibles. Para medir el cambio en la conducta de los consumidores, se estudian cinco posibles efectos: 1) Reducción en los tiempos entre compras sucesivas, a medida que los clientes se acercan al canje de un premio. 2) Aumento en el esfuerzo invertido en cada compra, medido en dinero gastado, al acercarse al canje de un premio. 3) Reinicio del patrón de compras del cliente después de canjear un premio. 4) Relación entre el esfuerzo requerido para el canje y el tipo de premio elegido por el cliente. 5) Magnitud de la aceleración en los tiempos entre compras sucesivas dependiendo del tipo de premio canjeado, hedónico o utilitario.

Utilizando el modelo de hazard proporcional en tiempo discreto, e incorporando heterogeneidad mediante el uso de clases latentes, se verifica la existencia de la hipótesis del gradiente objetivo, que indica la aceleración en las compras a medida que los clientes se acercan al canje de un premio, y se comprueba que existe un efecto de reinicio parcial en el comportamiento de compra después del canje. Además, se determinó que los clientes que eligen premios hedónicos aceleran más que aquellos que eligen premios utilitarios. Por otro lado, mediante modelos lineales generalizados, se encontró que a medida que los clientes se acercan al canje, el esfuerzo invertido, medido en dinero, es efectivamente creciente. Finalmente, mediante una regresión logística binaria e incorporando heterogeneidad con clases latentes, se encontró que a mayor esfuerzo invertido por los clientes, mayor es la probabilidad de que prefieran premios hedónicos por sobre premios utilitarios.

Estos resultados evidencian que este programa de lealtad modifica favorablemente el comportamiento de sus clientes, aumentando la frecuencia de compra y el dinero gastado en cada una de ellas, a medida que se acercan al canje de un premio.

AGRADECIMIENTOS

Quiero agradecer a mi familia y amigos por su constante apoyo durante este largo proceso. A mi profesor guía, Ricardo Montoya, por su gran ayuda en todos los temas de esta tesis y su incalculable paciencia. A todos quienes conocí durante mis años en la Universidad, tanto como alumno como profesor auxiliar. A todos quienes contribuyeron en hacer de todos estos años una experiencia de vida inolvidable. Muchas gracias.

ÍNDICE DE CONTENIDOS

INTRODUCCIÓN	1
I. OBJETIVOS Y ALCANCES.....	2
1.1 Objetivo general.....	2
1.2 Objetivos específicos	2
1.3 Alcances	2
II. RETAIL EN CHILE	3
III. PROGRAMAS DE LEALTAD	4
3.1 Lealtad en el retail	4
3.2 Definición y estructura de los programas de lealtad	5
IV. EFECTOS DE LOS CLUBES DE LEALTAD EN LOS CONSUMIDORES	9
4.1 Aceleración por cercanía a la meta	9
4.2 Preferencias por el tipo de premio según el esfuerzo	10
4.3 Otros efectos	11
4.3.1 Efecto de la ilusión de progreso	11
4.3.2 Efecto del esfuerzo en elecciones con riesgo	12
4.3.3 Ventaja en el esfuerzo como determinante del comportamiento.....	13
4.3.4 Resistencia a las promociones y congruencia esfuerzo-premio.....	14
V. MODELOS	14
5.1 Modelo de hazard proporcional en tiempo continuo	14
5.2 Modelo de hazard proporcional en tiempo discreto	16
5.2.1 Especificación del riesgo base	17
5.2.1.1 Exponencial	17
5.2.1.2 Weibull.....	18
5.2.1.3 Log-logística	18
5.2.1.4 Caso no paramétrico	19
5.2.2 Heterogeneidad con clases latentes	19
5.3 Modelos lineales generalizados.....	21
5.3.1 Regresión Normal	22
5.3.2 Regresión Poisson	23
5.3.3 Regresión logística binaria	23
VI. ANÁLISIS DEL CLUB DE LEALTAD	24
6.1 Estructura de premios	24

6.2 Descripción de los datos	24
6.2.1 Compras y canjes	25
6.2.2 Montos	28
6.2.3 Puntos vencidos	29
6.2.4 Construcción de bases de datos.....	31
VII COVARIABLES DE LOS MODELOS.....	32
7.1 Distancia al canje.....	32
7.2 Logaritmo del número de canjes.....	33
7.3 Distancia al canje suavizada por los canjes.....	34
7.4 Share de compra	34
7.5 Puntos por vencer.....	35
VIII. ESTUDIOS REALIZADOS.....	35
8.1 Simulación	36
8.1.1 Consideraciones generales.....	37
8.1.1.1 Caso distribución exponencial:	38
8.1.1.2 Caso distribución Weibull:	38
8.1.1.3 Caso distribución log-logística:	38
8.1.2 Generando los datos	38
8.1.3 Metodología y resultados	39
8.1.3.1 Datos generados según distribución exponencial.....	39
8.1.3.2 Datos generados según distribución weibull.....	40
8.1.3.3 Datos generados según distribución log-logística.....	41
8.1.4 Simulación con $\beta = 0$	42
8.1.4.1 Datos generados según distribución exponencial.....	42
8.1.4.2 Datos generados según distribución weibull.....	42
8.1.4.3 Datos generados según distribución log-logística.....	43
8.1.5 Simulación con covariable de tiempo	43
8.1.5.1 Datos generados según distribución exponencial.....	44
8.1.5.2 Datos generados según distribución weibull.....	44
8.1.5.3 Datos generados según distribución log-logística.....	45
8.2 Hipótesis 1 – Aceleración de compra	46
8.2.1 Análisis preliminar y motivación	46
8.2.2 Modelos.....	47

8.2.3 Resultados	49
8.2.4 Discusión.....	51
8.2.5 Heterogeneidad.....	52
8.2.5.1 Modelo.....	52
8.2.5.2 Resultados.....	53
8.3 Hipótesis 2 – Esfuerzo de compra pre-canje	61
8.3.1 Análisis preliminar y motivación	61
8.3.2 Modelos.....	62
8.3.3 Resultados	63
8.3.4 Discusión.....	64
8.4 Hipótesis 3 – Comportamiento de compra post-canje	64
8.4.1 Datos.....	65
8.4.2 Análisis preliminar y motivación	65
8.4.3 Modelo	66
8.4.4 Resultados	67
8.4.5 Discusión.....	70
8.4.6 Heterogeneidad.....	71
8.4.6.1 Modelo.....	71
8.4.6.2 Resultados.....	71
8.5 Hipótesis 4 – Preferencias por tipo de premio	77
8.5.1 Datos.....	78
8.5.2 Modelo	79
8.5.3 Resultados	80
8.5.4 Discusión.....	81
8.5.5 Heterogeneidad.....	82
8.5.5.1 Modelo.....	83
8.5.5.2 Resultados.....	83
8.5.5.3 Análisis de los resultados	85
8.6 Hipótesis 5 – Aceleración de compra hedonista versus utilitarista	85
8.6.1 Datos.....	86
8.6.2 Modelo	86
8.6.3 Resultados	87
8.6.4 Discusión.....	88
IX. CONCLUSIONES Y TRABAJOS FUTUROS.....	88

REFERENCIAS.....	92
ANEXOS	95

INDICE DE FIGURAS

Figura nº 1: Participación de mercado entre tiendas por departamento.....	4
Figura nº 2: Ejemplos de la relación entre las compras y el canje de un premio.....	6
Figura nº 3: Tiempo que tarda una rata en recorrer cada sección del camino	9
Figura nº 4: Plantilla para 10 y 12 estampillas, respectivamente	12
Figura nº 5: Formas del hazard según las distintas parametrizaciones.....	19
Figura nº 6: Porcentaje de compras v/s no compras	25
Figura nº 7: Cantidad de compras promedio por mes	26
Figura nº 8: Porcentaje de oportunidades de compra concretadas por mes	27
Figura nº 9: Cantidad de canjes promedio por mes	27
Figura nº 10: Porcentaje de oportunidades de canje concretadas por mes	28
Figura nº 11: Monto promedio gastado por cliente, por mes	29
Figura nº 12: Cantidad de veces que vencen puntos en promedio, por mes	29
Figura nº 13: Porcentaje de veces que vencen puntos sobre el total de ocasiones	30
Figura nº 14: Cantidad de puntos vencidos promedio, por mes	31
Figura nº 15: Tiempo entre compras promedio para los 4 niveles juntos	46
Figura nº 16: Tiempo entre compras promedio para el nivel 1	47
Figura nº 17: Monto de compra promedio v/s distancia al canje	62
Figura nº 18: Tiempo entre compras promedio en función del tramo de la distancia	66
Figura nº 19: Probabilidad de compra promedio en función del tramo del share	68
Figura nº 20: Probabilidad de compra promedio en función del tramo de la distancia	69
Figura nº 21: Gasto necesario para canjear un premio, por nivel.....	78
Figura nº 22: Probabilidad estimada y proporción de canjes hedónicos	82

INDICE DE TABLAS

Tabla nº 1: Estructura de canje por niveles de CMR puntos	8
Tabla nº 2: Estructura de canje por niveles de CMR puntos	24
Tabla nº 3: Estadísticos descriptivos para la frecuencia de compra y canje	25
Tabla nº 4: Estadísticos descriptivos para el monto y los puntos vencidos	31
Tabla nº 5: Clientes, canjes y canjes promedio por nivel	32
Tabla nº 6: Estimación de parámetros para simulación exponencial	40
Tabla nº 7: Estimación de parámetros para simulación Weibull.....	41
Tabla nº 8: Estimación de parámetros para simulación Log-logística	41
Tabla nº 9: Estimación de parámetros caso exponencial	42
Tabla nº 10: Estimación de parámetros caso Weibull	43
Tabla nº 11: Estimación de parámetros caso Log-logística.....	43
Tabla nº 12: Estimación de parámetros caso exponencial	44
Tabla nº 13: Estimación de parámetros caso Weibull	45
Tabla nº 14: Estimación de parámetros caso Log-logística.....	45

Tabla nº 15: Ajuste del modelo 1 según las distintas parametrizaciones	49
Tabla nº 16: Ajuste del modelo 2 según las distintas parametrizaciones	49
Tabla nº 17: Ajuste del modelo 3 según las distintas parametrizaciones	49
Tabla nº 18: Ajuste del modelo 4 según las distintas parametrizaciones	50
Tabla nº 19: Parámetros estimados del modelo 4, con parametrización weibull.....	50
Tabla nº 20: Influencia de la distancia sobre la probabilidad, modelo 4	51
Tabla nº 21: Evaluación del modelo con distinto número de clases latentes	53
Tabla nº 22: Parámetros estimados para el nivel 1, con 9 clases latentes.....	54
Tabla nº 23: Probabilidad de compra para cada grupo, al comienzo y al final	54
Tabla nº 24: Evaluación del modelo con distinto número de clases latentes	56
Tabla nº 25: Parámetros estimados para el nivel 2, con 6 clases latentes.....	56
Tabla nº 26: Probabilidad de compra para cada grupo, al comienzo y al final	57
Tabla nº 27: Evaluación del modelo con distinto número de clases latentes	58
Tabla nº 28: Parámetros estimados para el nivel 3, con 5 clases latentes.....	58
Tabla nº 29: Probabilidad de compra para cada grupo, al comienzo y al final	59
Tabla nº 30: Evaluación del modelo con distinto número de clases latentes	60
Tabla nº 31: Parámetros estimados para el nivel 4, con 3 clases latentes.....	60
Tabla nº 32: Probabilidad de compra para cada grupo, al comienzo y al final	60
Tabla nº 33: MAE y MAPE en la etapa de estimación.....	63
Tabla nº 34: MAE y MAPE en la etapa de validación.....	63
Tabla nº 35: Estimación de parámetros para la normal.....	64
Tabla nº 36: Evaluación del modelo con distinto número de clases latentes	72
Tabla nº 37: Parámetros estimados para el canje 1, con 3 clases latentes.....	72
Tabla nº 38: Probabilidad de compra para cada grupo, al comienzo y al final	72
Tabla nº 39: Evaluación del modelo con distinto número de clases latentes	73
Tabla nº 40: Parámetros estimados para el canje 2, con 2 clases latentes.....	74
Tabla nº 41: Probabilidad de compra para cada grupo, al comienzo y al final	74
Tabla nº 42: Evaluación del modelo con distinto número de clases latentes	75
Tabla nº 43: Parámetros estimados para el canje 3, con 2 clases latentes.....	75
Tabla nº 44: Probabilidad de compra para cada grupo, al comienzo y al final	75
Tabla nº 45: Evaluación del modelo con distinto número de clases latentes	76
Tabla nº 46: Parámetros estimados para el canje 4, con 2 clases latentes.....	76
Tabla nº 47: Probabilidad de compra para cada grupo, al comienzo y al final	76
Tabla nº 48: Ajuste del modelo según las distintas parametrizaciones	67
Tabla nº 49: Parámetros estimados según modelo Hazard, parametrización weibull	67
Tabla nº 50: Intervalos de confianza para la probabilidad de compra, al inicio	70
Tabla nº 51: Intervalos de confianza para la probabilidad de compra, al final.....	70
Tabla nº 52: Cantidad de clientes, canjes y artículos distintos que se canjearon.....	78
Tabla nº 53: Resultados modelo logit binario con las 3 variables independientes	80
Tabla nº 54: Porcentaje de premios hedonistas disponibles por nivel, por catálogo	81
Tabla nº 55: Evaluación del modelo con distinto número de clases latentes	83
Tabla nº 56: Parámetros estimados con 2 clases latentes	84
Tabla nº 57: Clientes, canjes y canjes promedio por nivel, canjes hedonistas.....	86
Tabla nº 58: Clientes, canjes y canjes promedio por nivel, canjes utilitarios	86
Tabla nº 59: Parámetros estimados modelo Hazard, weibull, canjes hedónicos.....	87
Tabla nº 60: Parámetros estimados modelo Hazard, weibull, canjes utilitarios.....	87

Tabla nº 61: Probabilidad de compra, al comienzo y al final, canjes hedonistas.....	88
Tabla nº 62: Probabilidad de compra, al comienzo y al final, canjes utilitaristas	88
Tabla nº 63: Parámetros estimados según distribución exponencial	95
Tabla nº 64: Parámetros estimados según distribución weibull.....	95
Tabla nº 65: Parámetros estimados según distribución log-logística	95
Tabla nº 66: Parámetros estimados según modelo no paramétrico	96
Tabla nº 67: Parámetros estimados según distribución exponencial	97
Tabla nº 68: Parámetros estimados según distribución weibull.....	97
Tabla nº 69: Parámetros estimados según distribución log-logística	97
Tabla nº 70: Parámetros estimados según modelo no paramétrico	98
Tabla nº 71: Parámetros estimados según distribución exponencial	99
Tabla nº 72: Parámetros estimados según distribución weibull.....	99
Tabla nº 73: Parámetros estimados según distribución log-logística	99
Tabla nº 74: Parámetros estimados según modelo no paramétrico	100
Tabla nº 75: Parámetros estimados según distribución exponencial	101
Tabla nº 76: Parámetros estimados según distribución weibull.....	101
Tabla nº 77: Parámetros estimados según distribución log-logística	101
Tabla nº 78: Parámetros estimados según modelo no paramétrico	102
Tabla nº 79: Influencia de la distancia sobre la probabilidad, modelo 1	103
Tabla nº 80: Influencia de la distancia sobre la probabilidad, modelo 2	103
Tabla nº 81: Influencia de la distancia sobre la probabilidad, modelo 3	103
Tabla nº 82: Parámetros estimados según distribución exponencial	104
Tabla nº 83: Parámetros estimados según distribución log-logística	104
Tabla nº 84: Parámetros estimados según modelo no paramétrico	104
Tabla nº 85: Intervalos de confianza para el parámetro beta1	105

INTRODUCCIÓN

La industria del retail en Chile ha tenido un importante crecimiento, con un incremento promedio aproximado del 7,5% en sus utilidades trimestrales en los últimos siete años (Neculmán, 2010). La industria presenta un alto dinamismo, donde las empresas nacionales más importantes incluso se han expandido a otros países de Latinoamérica. De lo anterior, surge la necesidad de las empresas de comprender de mejor manera el comportamiento de los consumidores ante las distintas variables de marketing, para así poder enfrentar con mayor eficacia el mercado altamente competitivo, tanto a nivel nacional como internacional. Con esta información, los retailers pueden desarrollar estrategias adecuadas, que les permitan captar nuevos clientes y retener a los existentes.

Los programas de lealtad suelen ser introducidos por las empresas para fomentar la fidelidad de los clientes, a través de la entrega de premios según las compras que estos realizan. El objetivo principal de los programas es establecer un alto nivel de retención de los clientes de los segmentos más rentables, al producir una mayor satisfacción y entregándoles más valor. Sin embargo, aunque estos programas son ampliamente utilizados en la industria del retail, no existe mucha investigación que indique si los clientes efectivamente perciben un mayor valor o si realmente contribuyen a mejorar la lealtad a la marca (Yi y Jeon, 2003).

La presente investigación tiene como objetivo estudiar las respuestas de los consumidores ante un programa de lealtad. Estudios anteriores de Kivetz et. al. (2006) muestran la existencia de un efecto de aceleración de compras sucesivas a medida que los clientes se acercan al canje de un premio, mientras que Kivetz y Simonson (2002) muestran que existe una preferencia por ciertos tipos de premios según el esfuerzo necesario para canjearlos. Ambos estudios sirven de base para esta investigación, la cual estudiará estos y otros efectos en un programa de lealtad de una cadena de retail chilena.

El programa de lealtad estudiado presenta una estructura de canjes por niveles, donde el cliente acumula puntos según los montos invertidos en cada compra, los cuales puede canjear por premios al alcanzar alguno de estos niveles. Se estudiarán cinco efectos, que permitirán evaluar la aceleración de los clientes tanto en el tiempo entre compras sucesivas como en los montos invertidos en cada una de ellas, así como las preferencias por el tipo de premio.

El contenido de esta tesis se presenta en 9 capítulos. En el capítulo I, se plantean los objetivos y alcances del estudio. Luego, en el capítulo II, se da una breve introducción al retail en Chile. En el capítulo III, se explica el funcionamiento de los programas de lealtad, y se entrega una descripción de los principales programas de las tiendas por departamento nacionales. Posteriormente, en el capítulo IV, se realiza una revisión de la literatura, mostrando distintos efectos que tienen los programas de lealtad sobre el comportamiento de compra y canje de los consumidores. Luego, en el capítulo V, se presenta una explicación detallada de los modelos utilizados en esta tesis.

Después, en el capítulo VI, se entrega una descripción exhaustiva del club de lealtad estudiado y los datos utilizados. En el capítulo VII, se describen las distintas covariables incorporadas en los modelos utilizados. Posteriormente, en el capítulo VIII, se testean las hipótesis planteadas en los objetivos de este trabajo, y se discuten los resultados obtenidos. Finalmente, en el capítulo IX, se entregan las conclusiones generales obtenidas y se plantean posibles trabajos futuros.

I. OBJETIVOS Y ALCANCES

1.1 Objetivo general

Analizar las respuestas de los consumidores ante un programa de lealtad, en función de su comportamiento de compra y canje.

1.2 Objetivos específicos

1. Medir si los clientes presentan un efecto de aceleración en su patrón de compras, a medida que se acercan al canje de un premio.
2. Medir si existe una aceleración en la magnitud de los esfuerzos (dinero invertido en cada compra) a medida que el cliente se acerca al canje de un premio.
3. Evaluar si los clientes del programa de lealtad reinician su patrón de compras después del canje de un premio.
4. Medir la relación entre la cantidad de esfuerzo invertido por los clientes y el tipo de premios que prefieren, hedónico o utilitario.
5. Evaluar si existe una diferencia en la aceleración de compra entre los clientes que canjean premios hedónicos o utilitarios.
6. Medir la efectividad del modelo de hazard proporcional en tiempo discreto en la estimación de parámetros, ante la presencia de datos con características particulares como las usadas en este estudio.

1.3 Alcances

- El estudio se realiza sobre el programa de lealtad de una tienda de retail chilena particular.
- Todos los clientes estudiados utilizan la tarjeta de la tienda como medio de pago.
- Este estudio describe una gran cantidad de efectos que pueden tener los programas de lealtad sobre el comportamiento de los clientes, aunque sólo se probarán aquellos descritos en los objetivos. De todas formas, la descripción de

estos efectos puede servir como apoyo para la toma de decisiones o elaboración de experimentos.

II. RETAIL EN CHILE

El retail es el sector económico dedicado a la comercialización masiva al por menor. Este concepto suele asociarse a grandes cadenas de supermercados y tiendas por departamento, pero en realidad se refiere a todo tipo de tiendas de venta al detalle, desde los pequeños almacenes de barrio hasta las grandes empresas de consumo masivo. El retail cubre gran parte de las necesidades de las personas, entre ellas la alimentación, el vestuario, el equipamiento del hogar, etc.

Según Durán y Kremerman (2008), el sector de multitiendas del retail puede dividirse en tres grandes grupos:

1. Tiendas por departamento: Son empresas que cuentan con gran cobertura. En Chile, se ubican principalmente en Santiago y en las ciudades más importantes del país, considerando su tamaño poblacional y nivel de desarrollo económico. Estas empresas cuentan con la importante ventaja de poseer un gran nivel de negociación con los proveedores, debido a su gran tamaño. Ofrecen todo tipo de productos, desde vestuario hasta artículos para el hogar. En esta categoría se encuentran Falabella, Ripley, París y La Polar.
2. Multitiendas: Estas empresas cuentan con una menor gama de productos que las tiendas por departamento y apuntan a segmentos de ingresos más bajos. Tienen menor presencia a nivel nacional y un nivel de ventas menor. En esta categoría se encuentran tiendas como Hites, Corona, Johnson's, etc.
3. Cadenas especialistas: Son empresas dedicadas a satisfacer pequeños nichos de mercado, ofreciendo productos especializados. Típicamente se centran en una sola línea de productos, aunque pueden abarcar más de una. En esta categoría se encuentran tiendas como ABC-Din, Tricot, Bata, etc.

La industria del retail en Chile ha evolucionado exitosamente en los últimos años, donde pocos grandes actores dominan el mercado nacional y apuntan al posicionamiento internacional, principalmente en Argentina, Perú y Colombia (Durán y Kremerman, 2008). Estos grandes actores corresponden principalmente a Falabella y Cencosud, estando ambos entre los retailers más importantes de Latinoamérica. En efecto, un estudio realizado por Deloitte el año 2010 ubicó a Cencosud en el primer lugar en ventas entre los retailers de América Latina. El mismo estudio ubicó a Falabella en el quinto lugar. Por otro lado, este estudio ubicó también a Cencosud en el sexto lugar entre los retailers con mayor crecimiento del año, mientras que Falabella ocupó el decimoquinto lugar.

En cuanto al mercado nacional de las tiendas por departamento, multitiendas y cadenas especialistas, Falabella es el líder del mercado, seguido de cerca por París (de

Cencosud) y Ripley. En la figura nº1 se muestra la participación de mercado de las ventas entre las principales tiendas de estos tres grupos al año 2011.

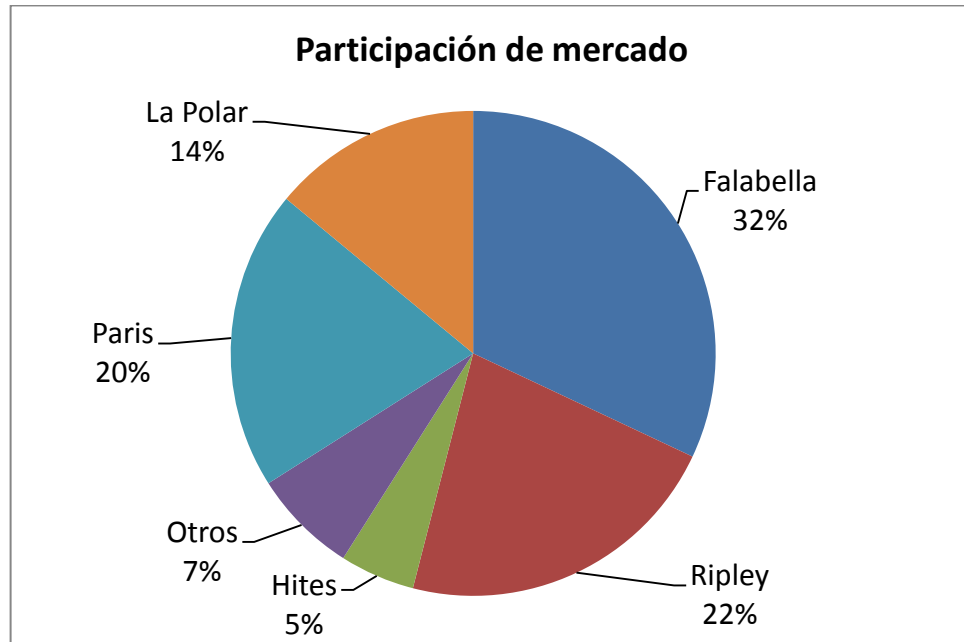


Figura nº 1: Participación de mercado entre tiendas por departamento, multitiendas y cadenas especialistas, Septiembre 2011.
Fuente: <http://www.feller-rate.cl/>.

III. PROGRAMAS DE LEALTAD

3.1 Lealtad en el retail

La lealtad, según la Real Academia Española, es “el cumplimiento de lo que exigen las leyes de la fidelidad y las del honor y hombría de bien”¹. El concepto de lealtad, en el ámbito de la vida cotidiana, se refiere entonces a una creencia en las instituciones, por ejemplo, la fidelidad en el matrimonio, o el compromiso con los amigos. Sin embargo, la lealtad del consumidor no funciona de esta forma. No existe un cliente que vaya a comprar siempre a la misma tienda para satisfacer todas sus necesidades (Humby, Hunt y Phillips, 2007).

La lealtad del consumidor consiste en una relación entre el cliente y la marca, que depende tanto del comportamiento como de la actitud: el cliente consume el producto o servicio regularmente y, además, siente una estrecha relación emocional con éste (Rubinson y Baldinger, 1996). Esta lealtad lleva al consumidor a tener una preferencia por la marca y provocará un efecto de boca-oreja, corriendo la voz sobre las cualidades del producto.

¹ Fuente: <http://buscon.rae.es/drael/SrvltConsulta?LEMA=lealtad> (Junio, 2011)

El problema del retail es cómo cultivar la lealtad de sus clientes, en mercados con una alta cantidad de competidores que venden muchos tipos de productos y de distintas marcas. En este sentido, el retail está en un escenario “always a share”, donde los clientes comparten sus compras entre varias tiendas. Es difícil lograr una relación emocional bajo estas circunstancias, por lo que las empresas deben motivar al cliente de otra forma, por ejemplo, mediante un programa de lealtad, o algún tipo de incentivo a cambio de su preferencia.

3.2 Definición y estructura de los programas de lealtad

Los programas de lealtad son un intento de incrementar el valor de los consumidores, premiándolos de forma proporcional a las compras que han realizado. Estos han sido utilizados en una gran variedad de industrias, como la hotelera, supermercados, tarjetas de crédito, etc. Su objetivo principal es incrementar tanto la tasa de retención de clientes (que el cliente se quede por más tiempo en la firma), como la cantidad de compras realizadas por cada uno de ellos (Blattberg, et. al., 2008).

Uno de los aspectos más relevantes a la hora de construir un programa de lealtad es la decisión en cuanto a la estructura de premios. Como se indicó anteriormente, la idea central es premiar al cliente según las compras que haya realizado, sin embargo, la forma de premiarlo y la cantidad de compras necesarias para entregarle el premio puede variar de acuerdo a la industria o el negocio. Blattberg et. al. (2008) explica que existen dos dimensiones que definen la estructura de los canjes, la linealidad y la continuidad. La linealidad se refiere a si la relación entre la cantidad de compras y la magnitud de los premios es lineal o no, mientras que la continuidad se refiere a si el cliente recibe un premio después de cada compra o sólo hasta alcanzar un cierto nivel de canje. En la figura nº2 se ilustran estas dimensiones.

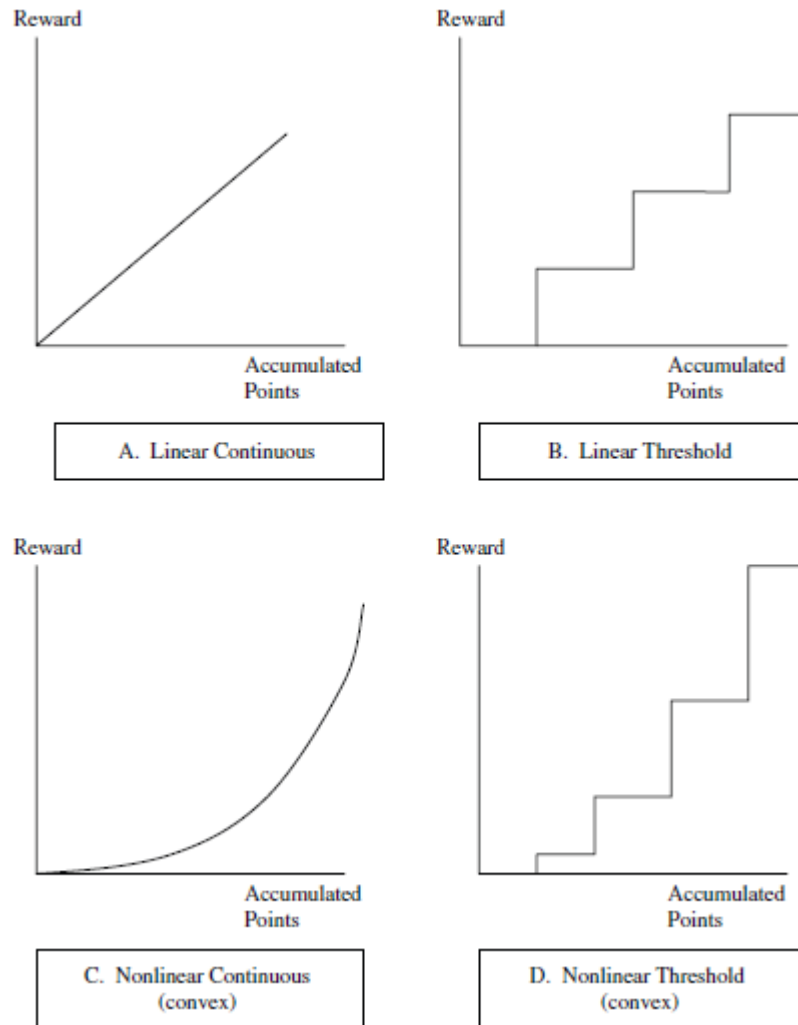


Figura nº 2: Ejemplos de la relación entre las compras y el canje de un premio
 Fuente: Blattberg et. al. (2008)

En la figura, se muestran 4 gráficos con distintos ejemplos de estructuras de premios para el programa de lealtad. En el eje horizontal se muestran los puntos acumulados², mientras que en el eje vertical se muestra el valor de los premios. En la figura nº2.A se expone el caso lineal continuo, el más simple de todos, donde se recibe un premio inmediatamente y la relación entre cantidad comprada y calidad del premio es lineal. Este es el caso típico de algunas tarjetas de crédito, donde se ofrece una milla para alguna línea aérea por cada peso gastado. En la figura nº2.B, se muestra el caso lineal discontinuo, donde el cliente debe alcanzar una cierta cantidad de puntos para canjear (por lo que el gráfico es escalonado), y el premio mejora linealmente con la cantidad de puntos necesarios para recibirlo. Las figuras nº2.C y nº2.D son análogas a las anteriores, solo que ahora la relación entre compras y premio no es lineal, sino

² En general, en muchos programas de lealtad se usa un sistema de puntos, los que se acumulan con cada compra. Estos puntos luego se pueden canjear por un premio.

convexa. Esto implica que a medida que el cliente acumula más puntos, el premio que recibe mejora (por ejemplo) exponencialmente.

La estructura de premios discontinua por niveles (ver por ejemplo figura n°2.B) produce un efecto de presión en el cliente que el caso continuo no provoca. Esto se debe a que cuando el cliente ya ha acumulado una cierta cantidad de puntos, siente la necesidad de comprar para llegar al canje de un premio. De todas formas, se debe notar que ninguna de estas estructuras es implícitamente mejor que las otras, por lo que la selección entre ellas dependerá del comportamiento de los consumidores en el negocio específico, la competencia existente, y otros factores que sean relevantes.

Un último aspecto importante a considerar es el vencimiento de los puntos. Un programa de lealtad donde los puntos vencen después de una cierta cantidad de tiempo provoca una mayor presión en el consumidor y, además, pueden ser usados como discriminador de precios³ (Kopalle y Neslin, 2003). Sin embargo, el problema es que manejar estos datos se hace difícil desde un punto de vista operacional, y el vencimiento podría causar una frustración en los clientes que perjudique al programa.

3.3 Programas de lealtad en Chile

En Chile, los programas de lealtad han tenido un importante crecimiento en los últimos años. El programa de pasajero frecuente LANPASS, de LAN Airlines, fue pionero en el país. Creado en 1996, ofrece un sistema de acumulación de millas similar al de AAdvantage (de American Airlines), ofreciendo la posibilidad de canjearlas por viajes gratuitos alrededor del mundo. Hoy en día, LAN entrega más de 250.000 pasajes gratis al año y cuentan con más de 4.000.000 de socios en Chile, Argentina, Perú, Ecuador y Estados Unidos⁴.

En cuanto a las tiendas por departamento, los principales actores de la industria usan un sistema de acumulación de puntos relativamente similar. En el caso de Falabella, la compañía creó el programa de CMR puntos en el año 2003. Este programa permite acumular puntos al realizar compras con su tarjeta CMR. La cantidad de puntos acumulados depende de dónde se realizó la compra; si la compra fue en Falabella se acumula 1 punto por cada \$150, mientras que por compras en otras tiendas se acumula 1 punto por cada \$300. Además, se entregan puntos por otros usos de la tarjeta, por ejemplo se recibe 1 punto cada \$950 de avance en efectivo o súper avance en efectivo⁵. Estos puntos pueden ser posteriormente utilizados para canjear un premio de alguno de los 12 niveles de canje existentes del programa. A mayor cantidad de puntos, mejores son los premios. En la tabla n°1, se presenta la estructura de canjes por niveles

³ Por ejemplo, un cliente que no compra con mucha frecuencia podría llegar a acumular suficientes puntos para canjear, pero para cuando lo haga, los puntos ya habrán expirado.

⁴ Fuente: http://www.lan.com/es_cl/sitio_personas/lanpass/acerca_de_lanpass/que_es_lanpass.html (Junio, 2011).

⁵ Fuente: <http://www.cmrfalabella.com/> (Junio, 2011)

del programa CMR puntos. Estos puntos tienen una vigencia de 12 meses desde el momento de la compra.

Nivel	Puntos
1	5.000
2	9.000
3	12.000
4	18.000
5	24.000
6	36.000
7	48.000
8	60.000
9	90.000
10	120.000
11	180.000
12	240.000

Tabla nº 1: Estructura de canje por niveles de CMR puntos
Fuente: www.cmrfalabella.com

Otro club de lealtad importante en el país es el Club Néctar, el que reúne a las marcas Jumbo, Santa Isabel, Tarjeta Más, Easy y París. Este programa de lealtad permite acumular puntos al realizar compras en estas tiendas, mostrando la tarjeta Néctar al momento de pagar, o bien indicando el RUT. Se entregan 2 puntos por cada \$200 pesos gastados en Jumbo, y sólo 1 punto por cada \$200 en el resto de las tiendas. Se acumula adicionalmente 1 punto por cada \$400 pagados con la tarjeta Más⁶. Estos puntos tienen una vigencia de 12 meses. Con respecto a los canjes, existen dos posibilidades. El cliente puede cambiar sus puntos por un cheque válido en las tiendas del club, donde un punto se cambia por un peso, así como también puede canjear premios de un catálogo. Este programa no presenta una estructura de niveles preestablecida, por lo que cada producto del catálogo requiere su propia cantidad de puntos.

Finalmente, existe también el club de lealtad de Ripley, Ripley Puntos. Este programa entrega 2 puntos por cada \$200 gastados en tiendas Ripley y pagados con la tarjeta de la tienda y 1 punto por cada \$200 con otro medio de pago. Además, entrega 1 punto por cada \$167 gastados en otras tiendas al pagar con la tarjeta Ripley⁷. Estos puntos también tienen una vigencia de 12 meses. La modalidad de canje consiste en cambiar puntos por cheques canjeables en la tienda, donde 1 punto equivale a 1 peso.

Existen en Chile muchos otros programas de lealtad; aquí sólo se explicaron los más importantes relacionados con las tiendas por departamento más importantes. Otros

⁶ Fuente: <http://www.nectar.cl/dynamic/juntar/juntar/comoJuntar> (Junio, 2011)

⁷ Fuente: http://www.ripley.cl/wcsstore/Ripley/ripleypuntos/como_acumulo.htm (Junio, 2011)

programas de lealtad exitosos son Mi Club Líder, de la cadena de supermercados Líder, el Club de Lectores del diario El Mercurio, Puntos Zona de la empresa de telecomunicaciones Entel, etc.

IV. EFECTOS DE LOS CLUBES DE LEALTAD EN LOS CONSUMIDORES

4.1 Aceleración por cercanía a la meta

Existen diversos estudios que indican que la cercanía o lejanía a una meta específica afecta el comportamiento de las personas. La hipótesis del gradiente objetivo, propuesta por el investigador Clark Hull (1934), propone que la tendencia a acercarse a una meta aumenta con la cercanía a ella. Para demostrarlo, construyó un camino recto para ratas, y realizó un experimento, ubicando a la rata en un extremo del camino y un poco de comida al otro extremo. Como resultado, Hull descubrió que las ratas aumentan su velocidad a medida que se acercan a la comida, como se muestra en la figura nº3.

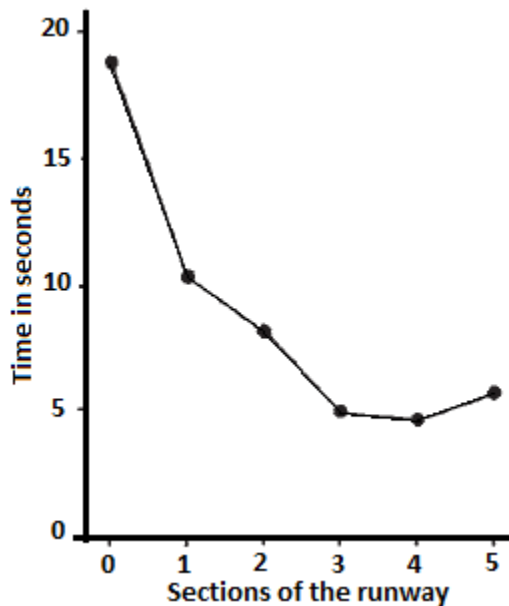


Figura nº 3: Tiempo que tarda una rata en recorrer cada sección del camino
Fuente: Kivetz et. al (2006)

Utilizando estos conceptos, Kivetz et. al. (2006) realizó un estudio de la hipótesis del gradiente objetivo aplicado a los programas de lealtad. Específicamente, analizó cómo la cercanía al canje de un premio afecta el comportamiento de compra de los individuos. En dicho estudio, se generaron proposiciones en base a dos programas de lealtad reales distintos: un programa de lealtad de una cafetería de una universidad, donde la compra de 10 café otorgaba la posibilidad de recibir un premio, y otro de una página web que premiaba a los usuarios por cada 51 canciones evaluadas en el sitio (siendo posible evaluar cualquier número de canciones en cada visita).

Como resultado, en dicho estudio se encontró que los miembros del programa de lealtad de la cafetería compraban café con mayor frecuencia a medida que se acercaban al canje del café gratis (en promedio, los tiempos entre compras se redujeron en 20%, o 7 días, a lo largo del programa). En cuanto al programa de lealtad de la página web, se vieron resultados similares, donde los tiempos entre visitas sucesivas al sitio también disminuían a medida que se acercaban al premio, y la cantidad de canciones evaluadas en cada una de ellas aumentaba. Esto prueba que la cercanía al canje no solo provoca una aceleración en la frecuencia de compra, sino también en la magnitud del esfuerzo que se realiza en cada una de ellas. Finalmente, los investigadores demuestran que los clientes, una vez llegado a la meta, reinician su patrón de compra, volviendo a la frecuencia de compra inicial al comenzar la carrera, con lo que descartan que la aceleración al premio se deba a factores de aprendizaje u otros.

4.2 Preferencias por el tipo de premio según el esfuerzo

La obtención de un premio en un programa de lealtad normalmente requiere que los clientes inviertan un esfuerzo en él. Dicho esfuerzo suele extenderse a lo largo del tiempo, y los premios sólo se entregan una vez cumplida la cantidad de esfuerzo requerida. El esfuerzo percibido se define entonces como cualquier inconveniente o dificultad asociados a cumplir los requisitos del programa, como por ejemplo cuando los clientes hacen el esfuerzo de comprar en una tienda específica o compran una mayor cantidad de un producto de lo que hubieran comprado normalmente. El esfuerzo percibido también incluye cualquier costo de sustitución, es decir, el costo en que incurren los clientes al comprar una marca que de otra forma no hubieran comprado (Blattberg y Neslin, 1990).

Algunas ramas de la psicología han estudiado el efecto en el comportamiento de invertir esfuerzos para alcanzar un premio. Los resultados más directos y robustos son aquellos que indican que los premios pueden ser altamente motivantes. Sin embargo, no existen muchos estudios que muestren la preferencia de los consumidores por un cierto tipo de esfuerzo para alcanzarlos (Kivetz y Simonson, 2002).

Kivetz y Simonson (2002) llevaron a cabo una investigación con el objetivo de mejorar la comprensión general de la preferencia de los consumidores hacia distintos programas de lealtad y hacia ciertos tipos de esfuerzos. Específicamente, investigaron la relación de la magnitud y tipo de esfuerzo realizado por clientes pertenecientes a un programa de lealtad, con los tipos de premio que eligen al momento de canjear un premio. Para ello, dividen primero los tipos de premio en 2 grupos de bienes:

- Bienes hedonistas, es decir, productos o servicios que no son esenciales para el diario vivir, pero que permiten llevar una vida más cómoda o lujosa; en otras palabras, un bien de lujo. Por ejemplo, unas vacaciones en el extranjero.

- Bienes utilitaristas, es decir, productos o servicios necesarios para llevar una vida normal, con un disfrute razonable de ella. Por ejemplo, comida, medicamentos, ropa, etc.

En base a estas definiciones, los investigadores realizaron varios estudios, poniendo a los encuestados frente a distintos programas de lealtad, donde los principales descubrimientos fueron que:

- Aumentar la magnitud de los esfuerzos requeridos para canjear un premio, provoca un incremento en la preferencia de las personas por programas de lealtad que ofrecen premios hedónicos en lugar de premios utilitarios.
- A mayores requerimientos del programa, mayor es el efecto positivo en la preferencia de los consumidores por premios hedonistas por sobre premios utilitaristas, en el caso de los consumidores con una alta tendencia a sentirse culpables al comprar bienes hedónicos (en comparación con consumidores con menor tendencia a sentirse culpables).
- Los consumidores presentan una mayor probabilidad de preferir premios hedónicos por sobre premios utilitarios cuando son premiados por comprar productos o servicios por razones de trabajo, en comparación con el caso en que compran los mismos productos o servicios por razones de placer personal.
- El efecto positivo de los requerimientos del programa en la preferencia por premios hedonistas por sobre utilitarios es más fuerte en el caso de consumo por razones de trabajo comparado con el consumo por razones de placer personal.

4.3 Otros efectos

4.3.1 Efecto de la ilusión de progreso

La hipótesis del gradiente-objetivo⁸ establece que los clientes aceleran sus compras a medida que se acercan al canje de un premio. Esta conceptualización sugiere que la determinante clave de la motivación es la proporción de distancia que queda para llegar al premio.

Kivetz et. al. (2006) propone que, en realidad, es la distancia percibida a la meta la que sirve de motivación y no la distancia real. Ambas pueden diferir cuando se aumenta el esfuerzo total necesario para llegar a la meta, pero se aumenta también (en la misma cantidad) la distancia ya recorrida por el cliente. En tal caso, el consumidor está a la misma distancia real del premio, pero como ya ha hecho un avance, su distancia relativa es menor, generando una ilusión de progreso. Para probarlo, el autor realiza un experimento en la cafetería de una universidad, donde a los clientes se les hace

⁸ Ver capítulo 4.1

entrega de una plantilla que indica que deben comprar 10 o 12 tazas de café. Cada vez que compran un café, se les hace entrega de una estampilla para pegar en su plantilla. Aquellos que reciben la plantilla que exige 12 tazas, reciben un bono inicial de 2 estampillas, como se indica en la figura nº4.

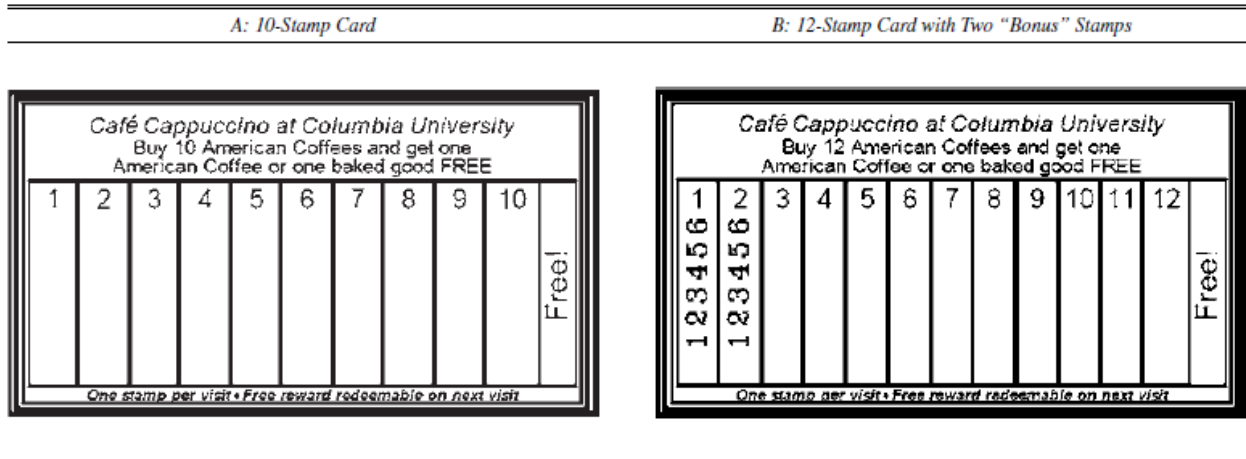


Figura nº 4: Plantilla para 10 y 12 estampillas, respectivamente
 Fuente: Kivetz et. al. (2006)

En ambos casos, el requisito real para llegar al premio es la compra de 10 cafés, sin embargo, la distancia percibida en el segundo caso es menor, debido a que el cliente observa que ya ha avanzado un poco hacia el premio. El autor verifica que, efectivamente, los clientes que recibieron la segunda plantilla aceleraron más rápido que aquellos que recibieron la primera.

4.3.2 Efecto del esfuerzo en elecciones con riesgo

En muchas ocasiones, cuando las personas realizan un esfuerzo que podría llevar a recibir algún tipo de premio, deben realizar un trade-off entre la probabilidad de recibirlo y la magnitud del mismo. Por ejemplo, una persona podría tener la posibilidad de elegir entre un trabajo donde recibe un sueldo estable y otro donde recibe una comisión potencialmente más alta, pero no segura.

Kivetz (2003) propone que, en un programa de lealtad, el nivel de esfuerzo tiene un efecto predecible en las preferencias por la probabilidad y magnitud del premio elegido, por ejemplo, en la preferencia por un premio seguro-pequeño sobre otro incierto-grande⁹. Para probarlo, realizó diversos estudios con casos reales e hipotéticos, los que permitieron concluir que:

⁹ un premio seguro-pequeño podría ser, por ejemplo, un juego de toallas y uno incierto-grande una entrada para participar en una lotería.

- La sola presencia de requerimientos de esfuerzo aumentan la preferencia por premios seguros-pequeños por sobre inciertos-grandes, comparado con el caso en que a las personas se les entrega un premio sin necesidad de esforzarse.
- La preferencia por premios seguros se ve atenuada cuando la actividad de esfuerzo es intrínsecamente motivante, es decir, cuando los requerimientos del programa resultan en una actividad placentera o que la persona disfruta llevar a cabo.
- Aumentar continuamente los requisitos de esfuerzo lleva a un efecto de U-invertida en la preferencia por premios seguros-pequeños sobre inciertos-grandes. Esto se debe a que al aumentar demasiado los esfuerzos requeridos, el premio seguro-pequeño deja de ser atractivo (ya no es suficiente para compensar el esfuerzo invertido) y las personas prefieren tomar el riesgo del incierto-grande.

4.3.3 Ventaja en el esfuerzo como determinante del comportamiento

Una gran cantidad de investigaciones han mostrado que las preferencias de los consumidores son, en muchos casos, inestables y que ellos construyen sus preferencias y evaluaciones al momento de ser enfrentados a la necesidad de tomar una decisión (Bettman et. al., 1998). Así, se hace difícil determinar el valor de una opción individual o una oferta de marketing, pues las personas no siempre tienen un punto de referencia disponible.

Algunos estudios han mostrado que los consumidores suelen buscar señales que indiquen que la oferta es atractiva (Winer, 1986). Por ejemplo, incluso si un consumidor no sabe el valor real de un producto específico, el sólo hecho de que se le indique que está con una rebaja del 20% sirve de señal para saber que el precio es atractivo.

Kivetz y Simonson (2003) plantean que las personas suelen tomar decisiones relacionadas con promociones de marketing, en base a su “ajuste idiosincrático”. En otras palabras, un factor importante en su decisión es si la oferta se ajusta a ellos más que al resto de los consumidores. Si creen que, por ejemplo, una oferta es especialmente buena para ellos, pueden concluir que esa opción es particularmente atractiva, pues sienten que les da una cierta ventaja por sobre otros consumidores¹⁰. Los autores plantean y prueban mediante distintos estudios que, como los programas de lealtad suelen ser difíciles de comparar (debido a las distintas formas de estructurarlos), muchos consumidores basan fuertemente su decisión en su ajuste idiosincrático, lo que puede llevar a decisiones sub-óptimas. En tal caso, una persona

¹⁰ Para ejemplificar el efecto de ajuste idiosincrático, se puede pensar en el programa de lealtad de una estación de gasolina. Si un individuo vive cerca de dicha estación, él puede sentir que se ve especialmente favorecido por el programa, ya que debe realizar menos esfuerzo que alguien que vive más lejos.

podría preferir un programa de lealtad con mayores requisitos que otro equivalente, por la simple razón de que les da una mayor ventaja psicológica con respecto a los demás.

4.3.4 Resistencia a las promociones y congruencia esfuerzo-premio

Las promociones de marketing pueden ser una herramienta muy útil, pero deben implementarse con cuidado. La intuición indica que ellas producen un efecto positivo en las ventas, debido a que los consumidores las ven como una oportunidad de la que pueden sacar provecho. Sin embargo, las mismas promociones podrían ser interpretadas por las personas como un intento de influenciar su comportamiento de compra, lo que provocaría una resistencia a ellas.

Kivetz (2005) sugiere que la resistencia a las promociones es un fenómeno particularmente potente cuando el intento de influencia es evidente, como en un programa de lealtad donde se le exige al cliente comprar para conseguir un premio. El autor plantea y demuestra que los consumidores disminuyen su resistencia a un programa al elegir premios congruentes con el esfuerzo invertido, es decir, premios que son similares al esfuerzo requerido: al hacerlo, los clientes infieren que participan en el programa porque ellos quieren, y no por una influencia externa. Por ejemplo, un programa de lealtad donde se deben evaluar 30 películas presentaría menor resistencia por parte de los clientes si el premio es una película en DVD que un CD de música, pues el primero está relacionado con el esfuerzo requerido y el segundo no.

V. MODELOS

5.1 Modelo de hazard proporcional en tiempo continuo

El modelo de hazard proporcional (Cox, 1972) permite modelar la probabilidad instantánea de que un cliente efectúe una compra, condicional en el tiempo que ha transcurrido desde su compra anterior. A esta probabilidad se le denomina función hazard, la que se define como:

$$h_i(t, X_{it}) = h_i(t) * \Psi_i(X_{it}) \quad (1)$$

Esta función depende de dos factores: la función de riesgo base (baseline hazard) $h_i(t)$, que captura el patrón intertemporal de compra del cliente i , y una función de covariables (*covariate function*) $\Psi_i(X_{it})$, que incorpora el efecto de factores externos al comportamiento de compra. Estos factores externos se representan por X_{it} , y corresponden a variables de marketing como el precio, existencia de promociones, etc. Es importante notar que la función de riesgo base es la que le dará la forma a la función hazard, mientras que la función de covariables sólo afectará en la magnitud de la curva, desplazándola hacia arriba o hacia abajo según corresponda.

La función de covariables se puede representar de muchas maneras, pero una de las formas más comunes está dada por la siguiente expresión (Seetharaman y Chintagunta, 2002):

$$\Psi_i(X_{it}) = e^{X_{it}\beta_i} \quad (2)$$

Donde β_i son los parámetros asociados a las covariables del modelo. Se usa esta forma funcional porque la probabilidad de compra debe ser no negativa. Por otro lado, la función hazard se puede escribir también como:

$$h_i(t, X_{it}) = \frac{f(t, X_{it})}{1 - F(t, X_{it})} \quad (3)$$

Donde $f(t, X_{it})$ es la función de densidad de probabilidad correspondiente a la función hazard del cliente i en el tiempo t , y $F(t, X_{it})$ es su respectiva función de distribución acumulada. Esta expresión se puede interpretar directamente pensando que la función de densidad es la probabilidad incondicional de que el cliente compre el producto en t , mientras que $1 - F(t, X_{it})$ es la probabilidad de que el cliente no haya comprado hasta t .

De las expresiones (1), (2) y (3) se deduce que:

$$\frac{f_i(t, X_{it})}{1 - F_i(t, X_{it})} = h_i(t) * e^{X_{it}\beta_i} \quad (4)$$

Y definiendo la función de sobrevivencia $S(t, X_{it}) = 1 - F(t, X_{it})$ (la probabilidad de que el cliente sobreviva al evento de compra hasta t), se obtiene:

$$f_i(t, X_{it}) = h_i(t) * e^{X_{it}\beta_i} * S(t, X_{it}) \quad (5)$$

Para simplificar esta última expresión y facilitar las estimaciones, se puede hacer un paso intermedio, reescribiendo la ecuación (4) como se muestra a continuación:

$$\int_0^{F_i(t, X_{it})} \frac{dF_i(u, X_{iu})}{1 - F_i(u, X_{iu})} = \int_0^t h_i(u) * e^{X_{iu}\beta_i} du \quad (6)$$

Se debe notar que en esta última ecuación, el tiempo igual a cero corresponde al período en que ocurrió la compra anterior del cliente i . Resolviendo (6) se obtiene:

$$F_i(t, X_{it}) = 1 - e^{-\int_0^t h_i(u) * e^{Xiu\beta_i} du} \quad (7)$$

De las expresiones (5) y (7) y recordando que $S(t, X_t) = 1 - F(t, X_t)$, se llega finalmente a que:

$$f_i(t, X_t) = h_i(t) * e^{Xit\beta_i} * e^{-\int_0^t h_i(u) * e^{Xiu\beta_i} du} \quad (8)$$

Esta expresión representa la densidad de probabilidad asociada a que ocurra el evento de compra del cliente i en el instante t . Sólo queda señalar que la función de riesgo base puede ser especificada según distintas parametrizaciones, lo que se explicará en profundidad en el capítulo 5.2.1. Para estimar los parámetros del modelo de hazard proporcional en tiempo continuo, se puede proceder por el método de máxima verosimilitud.

5.2 Modelo de hazard proporcional en tiempo discreto

Una limitante importante del modelo de hazard en tiempo continuo, es que las acciones del cliente se miden en intervalos de tiempo discretos, por ejemplo mensualmente. Por esta razón, es necesario derivar una versión discreta del modelo.

Como las compras ocurren a intervalos discretos, la expresión (7) se puede reescribir en tiempo discreto, como se muestra a continuación:

$$F_i(t, X_{it}) = 1 - \exp\left(-\sum_{u=1}^t e^{Xiu\beta_i} \int_{u-1}^u h_i(w) dw\right) \quad (9)$$

En esta expresión, la integral se reescribe parcialmente como una sumatoria, donde los tiempos u son discretos (por ejemplo, medido en meses). De lo anterior, la probabilidad de que un cliente i compre en el tiempo discreto t se puede escribir como:

$$P_i(t, X_{it}) = 1 - \frac{S(t, X_{it})}{S(t-1, X_{i,t-1})} = 1 - \frac{\exp\left(-\sum_{u=1}^t e^{Xiu\beta_i} \int_{u-1}^u h_i(w) dw\right)}{\exp\left(-\sum_{u=1}^{t-1} e^{Xiu\beta_i} \int_{u-1}^u h_i(w) dw\right)} =$$

$$P_i(t, X_{it}) = 1 - \exp\left(-e^{Xit\beta_i} \int_{t-1}^t h_i(u) du\right) \quad (10)$$

A esta probabilidad se le conoce como hazard en tiempo discreto.

Para estimar los parámetros, se supondrá en primera instancia homogeneidad de los clientes, es decir, se asumirá que todos los clientes presentan el mismo comportamiento inter-temporal de compra¹¹. Bajo este supuesto, la función de verosimilitud queda entonces dada por:

$$L = \prod_{i=1}^N \prod_{t=1}^{T_i} P_i(t, X_{it})^{\delta_{it}} * (1 - P_i(t, X_{it}))^{1-\delta_{it}} \quad (11)$$

Donde δ_{it} es un parámetro binario que indica si el cliente i compra en el período t , T_i es el total de períodos observados del cliente i y N es el número de clientes. Esta definición de la verosimilitud considera tanto cuando el cliente compró como cuando no lo hizo. En otras palabras, el tiempo t corresponde a una posibilidad de compra, mientras que el parámetro δ_{it} determina si efectivamente la compra se concretó o no. Ambos casos aportan información y, por ende, deben incluirse en la verosimilitud. Finalmente, para estimar los parámetros, sólo queda maximizar la expresión (11)¹².

5.2.1 Especificación del riesgo base

El riesgo base es parte fundamental de la función hazard, pues es el que determina la forma de esta curva, es decir, cómo varía la probabilidad a medida que pasa el tiempo. El riesgo podría ser creciente, constante o decreciente en el tiempo, e incluso en casos más complicados puede ser primero creciente y luego decreciente, por ejemplo. Por otro lado, la estimación de los parámetros asociados a las variables externas depende de la especificación del riesgo base, pues dependiendo de la forma de la curva, su valor podría cambiar para ajustarse a ella.

Algunas parametrizaciones típicas para el modelo hazard proporcional en tiempo discreto son las siguientes.

5.2.1.1 Exponencial

En la parametrización Exponencial, el riesgo base queda dado por:

$$h(t) = \lambda \quad (12)$$

Con $\lambda > 0$. Esto implica que la probabilidad condicional de que ocurra un evento es constante en el tiempo.

¹¹ La heterogeneidad de los clientes será abordada en el capítulo 5.2.2

¹² En realidad, lo que se suele maximizar es el logaritmo natural de la verosimilitud, lo que facilita los cálculos computacionales.

5.2.1.2 Weibull

En la parametrización Weibull, el riesgo base queda dado por:

$$h(t) = \gamma\alpha(\gamma t)^{\alpha-1} \quad (13)$$

Con $\gamma, \alpha > 0$. Este modelo es más general que la exponencial y, por ende, más flexible. Permite que la función hazard sea no constante, pero monótonica. Es un modelo de 2 parámetros, donde γ es el parámetro de localización y α el parámetro de forma, es decir, el que determina finalmente si la función es creciente, constante o decreciente. Notar que si $\alpha = 1$ se recupera el modelo exponencial, si $\alpha < 1$ el riesgo base es monótonicamente decreciente y si $\alpha > 1$ es monótonicamente creciente en el tiempo. Esta es la especificación más popular en la literatura del modelo hazard proporcional (Seetharaman y Chintagunta, 2002).

5.2.1.3 Log-logística

En la parametrización Log-logístico, el riesgo base queda dado por:

$$h(t) = \frac{\gamma\alpha(\gamma t)^{\alpha-1}}{1 + (\gamma t)^\alpha} \quad (14)$$

Con $\gamma, \alpha > 0$. Tal como en el modelo Weibull, en este caso el riesgo base presenta dos parámetros, donde γ es el parámetro de localización y α el parámetro de forma. La log-logística permite que el riesgo base sea monótonicamente decreciente o bien presente una forma de "U invertida".

En la figura nº5, se muestra un bosquejo de las curvas que se pueden obtener con las distintas parametrizaciones del hazard, usando parámetros arbitrarios.

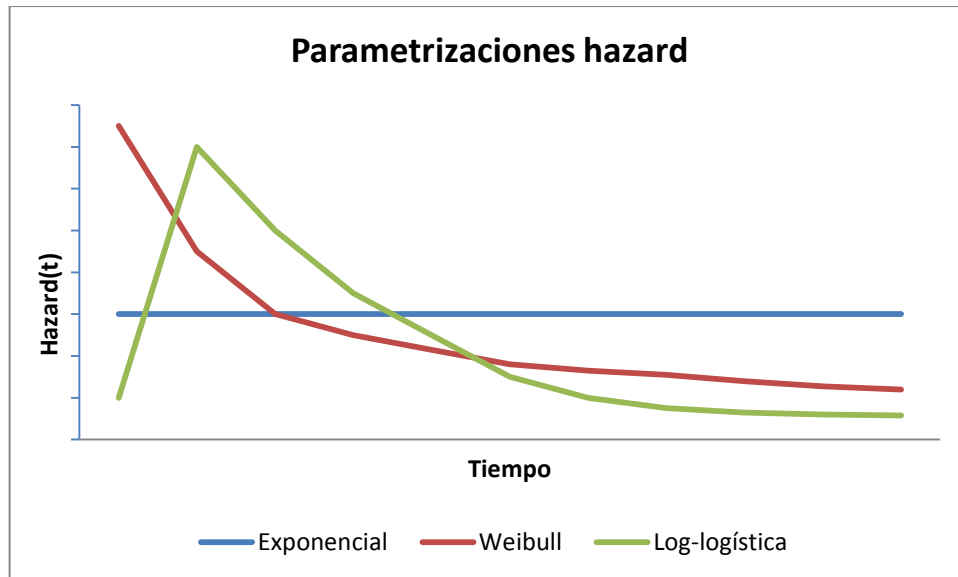


Figura nº 5: Formas del hazard según las distintas parametrizaciones
Fuente: Elaboración propia

5.2.1.4 Caso no paramétrico

Como en muchos casos es difícil especificar una parametrización particular para el riesgo base, ya sea porque se desconoce su comportamiento o este varía mucho en el tiempo, se puede usar también una versión no paramétrica del modelo. Un riesgo base no paramétrico es equivalente a estimar un parámetro para cada tiempo de falla posible (es decir, para cada período en que se produce una compra). Una ventaja de esta aproximación es que los datos determinan la forma de la función hazard que mejor se ajusta a ellos. La desventaja es que la forma de la curva depende mucho de los datos del estudio y podría no servir para otros con características diferentes, es decir, incluso si la curva se ajusta bien a los datos, es improbable que ella represente el proceso intrínseco que los genera. Otra desventaja es que si la cantidad de períodos en estudio es grande (por ejemplo, muchos meses), la cantidad de parámetros a estimar se hace muy grande.

Como no existe una regla que determine cuál de las dos aproximaciones es mejor (modelos paramétricos versus no paramétricos), en este estudio se aplicarán ambos y se verá empíricamente cuál se ajusta mejor al problema particular.

5.2.2 Heterogeneidad con clases latentes

Un aspecto importante a considerar es la heterogeneidad no observada de los individuos. Hasta ahora, se han descrito los modelos de hazard proporcional en tiempo continuo y discreto, considerando que existe homogeneidad de los individuos de la población estudiada, es decir, considerando que todos ellos presentan el mismo comportamiento. Sin embargo, esto podría ser un supuesto demasiado fuerte, pues los resultados obtenidos para la población como un todo podrían no ser aplicables para

cada una de las personas de la muestra en forma individual (Lancaster, 1979). Esto podría llevar a conclusiones erróneas que no representan bien la realidad.

Para abordar el problema de la heterogeneidad, existen técnicas estadísticas de clasificación que detectan patrones similares entre los individuos, de tal forma de agruparlos en segmentos distintos, cada uno con sus propias características. Una de estas técnicas es el análisis de clases latentes, que permite modelar la relación entre las distintas variables observadas, suponiendo que la estructura de la relación subyacente se explica por una variable latente categórica no observada (Monroy et. al., 2009). Este análisis permite identificar el menor número de grupos (o clases) que describe a la población, asignando a cada individuo en uno de estos grupos con cierta probabilidad, la que se debe estimar. Cada grupo cuenta con características diferentes (cada uno tiene valores distintos de los parámetros estimados), por lo que al terminar el análisis se contará con un número idealmente reducido de grupos que representan de buena forma los distintos comportamientos de los individuos de la población estudiada.

Para aplicar el análisis de clases latentes al modelo hazard proporcional en tiempo discreto, se deben escribir primero las ecuaciones para un solo individuo y considerar el grupo al que este pertenece. Observando la ecuación (11), la verosimilitud para el individuo i que pertenece al grupo g queda dada por.

$$L_{ig} = \prod_{t=1}^{T_i} P_{ig}(t, X_{it})^{\delta_i} * (1 - P_{ig}(t, X_{it}))^{1-\delta_i} \quad (15)$$

donde

$$P_{ig}(t, X_{it}) = 1 - \exp\left(-e^{X_{it}\beta_{ig}} \int_{t-1}^t h_{ig}(u) du\right) \quad (16)$$

donde además se debe considerar que h_{ig} se modela según las parametrizaciones vistas en el capítulo 5.2.1, pero agregando un índice g a los parámetros. En otras palabras, cada parámetro estimado en el hazard proporcional en tiempo discreto ahora cuenta con un índice que representa el grupo al que pertenece el individuo. Sea π_g la probabilidad de que el individuo pertenezca al grupo g . La verosimilitud del individuo i queda entonces dada por la siguiente expresión.

$$L_i = \sum_{g=1}^G L_{ig} \pi_g \quad (17)$$

donde G es el total de grupos. Finalmente, la verosimilitud para la muestra global queda como se muestra a continuación.

$$L = \prod_{i=1}^N \left(\sum_{g=1}^G \left(\prod_{t=1}^{T_i} P_{ig}(t, X_{it})^{\delta_{it}} * (1 - P_{ig}(t, X_{it}))^{1-\delta_{it}} * \pi_g \right) \right) \quad (18)$$

Se debe notar que π_g es un conjunto de parámetros adicionales a estimar.

Un último aspecto a considerar es que el número de grupos a utilizar depende de las características de la muestra. En general, la metodología utilizada es comenzar estimando los parámetros con dos grupos, y luego ir incrementándolos de a uno, hasta encontrar el menor número de grupos que representa mejor a los datos¹³. También se debe notar que si se usa un solo grupo, se recupera el modelo del caso homogéneo, visto en el capítulo 5.2.

5.3 Modelos lineales generalizados

Originalmente desarrollados por Nelder y Wedder (1972), los modelos lineales generalizados son una extensión de los familiares modelos de regresión, como la regresión lineal o los modelos de Tobit y Probit (McCullagh y Nelder, 1989). Un modelo lineal generalizado se estructura en tres partes:

1. Una componente aleatoria, que especifica la distribución condicional de la variable dependiente Y_i dados los valores de la i -ésima observación de las variables independientes del modelo. La distribución de Y_i debe pertenecer a la familia exponencial¹⁴, como la normal, binomial, poisson, gamma, etc.
2. Un predictor lineal, es decir, una función lineal de regresores, como se muestra a continuación:

$$\eta_i = \beta_0 + \sum_{k=1}^K \beta_k X_{ik} \quad (19)$$

donde, al igual que en la regresión lineal, las variables independientes X_{ik} pueden ser variables explicativas cuantitativas, regresores polinomiales, variables dummy, etc.

¹³ Para determinar el mejor modelo, una opción es utilizar el criterio de información bayesiana. El criterio de información bayesiano es un criterio de selección de modelos que mide el ajuste en base al logaritmo de la verosimilitud, penalizando por el número de parámetros. Se calcula como $BIC = -2*LL + k*Ln(N)$, donde k es el número de parámetros y N el tamaño de la muestra. Aquel modelo con un menor BIC se debe preferir por sobre los demás.

¹⁴ La familia exponencial se define como aquellas distribuciones cuya función de densidad de probabilidad se puede reescribir como $f(y_i) = \exp\left(\frac{y_i \theta_i - b(\theta_i)}{a_i(\phi)}\right) + c(y_i, \phi)$ donde θ_i y ϕ son parámetros y

$a_i(\phi)$, $b(\theta_i)$ y $c(y_i, \phi)$ son funciones conocidas.

3. Una “función link” $g(\cdot)$, continua e invertible, que transforma la esperanza de la variable dependiente ($\mu_i \equiv E(Y_i)$) en el predictor lineal:

$$g(\mu_i) = \eta_i = \beta_0 + \sum_{k=1}^K \beta_k X_{ik} \quad (20)$$

La especificación de la función link dependerá principalmente de la distribución que se elija para la regresión. Existe una función link canónica para cada distribución, pero esto no significa que no se pueda elegir otra forma si resulta mejor para el estudio particular.

Debido a que la función link es invertible, se puede reescribir la expresión anterior como $\mu_i = g^{-1}(\beta_0 + \sum_{k=1}^K \beta_k X_{ik})$, por lo que los modelos lineales generalizados pueden ser pensados como un modelo lineal para una transformación de los valores esperados de la variable de respuesta Y_i o como un modelo no-lineal para la respuesta Y_i . Estos modelos pueden ser estimados por el método de máxima verosimilitud.

En las secciones siguientes, se mostrarán distintas regresiones de interés que se pueden realizar utilizando modelos lineales generalizados, especificando la distribución a usar y su función link canónica.

5.3.1 Regresión Normal

Una variable aleatoria distribuida como una Normal, tiene una función de densidad de probabilidad dada por:

$$f(Y_i) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(Y_i - \mu_i)^2}{2\sigma^2}\right) \quad (21)$$

Se puede verificar que la Normal pertenece a la familia exponencial, desarrollando el término que está al cuadrado en la exponencial en (21). Con esto y un poco de álgebra se llega a:

$$f(Y_i) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(\frac{Y_i\mu_i - 2Y_i^2 - 2\mu_i^2}{\sigma^2}\right) \quad (22)$$

Observando el coeficiente que acompaña a Y_i , en (22), se obtiene inmediatamente que $\theta_i = \mu_i$, de donde se concluye que función link canónica para la Normal es la que se denomina como “identidad” (no se aplica ninguna transformación a la esperanza de la variable dependiente). Es importante notar que al usar una distribución normal y su función link canónica, se recupera la regresión lineal habitual.

5.3.2 Regresión Poisson

Una variable aleatoria distribuida como una Poisson, tiene una función de densidad de probabilidad dada por:

$$f(Y_i) = \frac{e^{-\mu_i} \mu_i^{Y_i}}{Y_i!} \quad (23)$$

Con $y_i \in \{0,1,2,\dots\}$ y $\mu_i > 0$. Se puede verificar que la Poisson pertenece a la familia exponencial, aplicando el logaritmo a la expresión (23).

$$\log(f(Y_i)) = Y_i \log(\mu_i) - \mu_i - \log(Y_i!) \quad (24)$$

Y aplicando la exponencial a (24), se obtiene:

$$f(Y_i) = \frac{\exp\left(\frac{Y_i \log(\mu_i) - \mu_i}{1}\right)}{Y_i!} \quad (25)$$

Observando el coeficiente que acompaña a Y_i , en (25), se obtiene que $\theta_i = \log(\mu_i)$ de donde se concluye que el logaritmo es la función *link* canónica para la Poisson.

5.3.3 Regresión logística binaria

Esta regresión es de particular importancia, pues permite modelar casos donde la variable dependiente es dicotómica (0/1, sí/no, compra/no compra, etc.) y donde se desea saber cuál es la probabilidad de elegir una de las dos alternativas por sobre la otra. Si se supone que la variable dependiente Y_i del modelo es una variable binaria, se puede especificar su función de densidad de probabilidad como una Bernoulli:

$$f(y_i) = \pi^{y_i} (1 - \pi)^{1-y_i} \quad (26)$$

donde π_i representa la probabilidad de obtener un 1 en la variable y $1 - \pi_i$ la probabilidad de obtener un 0. Haciendo un poco de trabajo algebraico, se puede llegar a:

$$f(y_i) = (1 - \pi) \left(\frac{\pi}{1 - \pi}\right)^{y_i} = (1 - \pi) \exp\left(y_i \log\left(\frac{\pi}{1 - \pi}\right)\right) \quad (27)$$

De aquí se observa que la expresión que acompaña a Y_i en (27) es $\theta_i = \log\left(\frac{\pi}{1-\pi}\right)$ de donde se concluye que la función *link* para la regresión logística binaria es un logit.

VI. ANÁLISIS DEL CLUB DE LEALTAD

6.1 Estructura de premios

El presente estudio tiene como objetivo medir los efectos de un programa de lealtad real nacional en el comportamiento de los consumidores. El programa de lealtad estudiado presenta una estructura de premios convexa y discontinua¹⁵, existiendo cuatro niveles de canje, donde es necesario alcanzar la cantidad de puntos requerida por alguno de ellos para canjear un premio. En cada compra, se acumula 1 punto por cada \$120 invertidos en la tienda. En la tabla nº2, se muestran los 4 niveles de canje, junto con la cantidad de puntos necesarios para canjear un premio en cada uno de ellos.

Nivel	Puntos
1	5.000
2	9.000
3	12.000
4	24.000

Tabla nº 2: Estructura de canje por niveles del programa
Fuente: Elaboración propia

Se debe notar que el programa de lealtad presentaba más niveles de canje, los que requerían una mayor cantidad de puntos, sin embargo estos no fueron considerados para efectos de este estudio debido a la escasa cantidad de canjes en comparación con los cuatro primeros niveles. Esta falta de canjes se puede explicar por la gran cantidad de dinero que se debe invertir para canjear en los niveles más altos. Por ejemplo, para canjear en el nivel 4 se debe gastar al menos \$2.880.000, por lo que son pocos los clientes que logran llegar a ese objetivo.

6.2 Descripción de los datos

Para la realización de este estudio, se utilizaron datos de 20.697 clientes de una tienda de retail nacional y que pertenecen al programa de lealtad de la misma. Cada cliente en la base de datos presenta múltiples ocasiones de compra, entre los meses Marzo del año 2003 y Abril del año 2008. Se cuenta con datos de compras y canjes, montos invertidos, puntos recibidos y vencidos, todo a nivel mensual y para cada cliente. A continuación se procederá a describirlos.

¹⁵ ver capítulo 3.2 para más detalles sobre las estructuras de premios típicas de los programas de lealtad.

6.2.1 Compras y canjes

En la base de datos, se tienen 652.823 ocasiones de compra en total entre los distintos clientes, de las cuales se concretan 535.592¹⁶. En la figura nº6 se muestra el porcentaje de ocasiones de compra concretadas (donde hubo una compra) contra las no concretadas (donde no hubo compra).

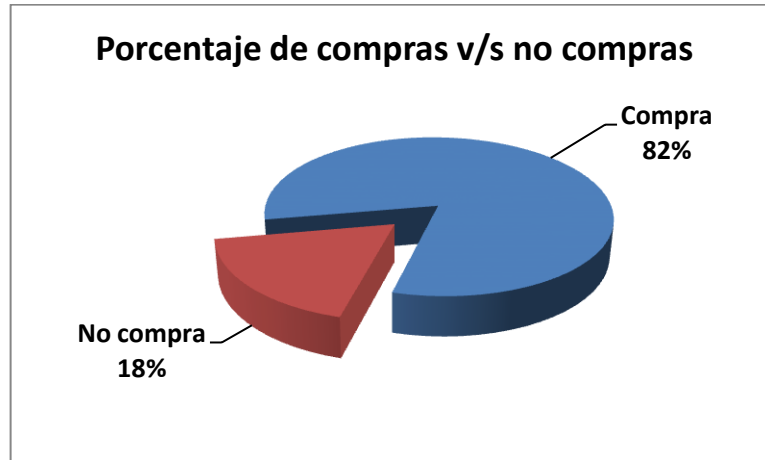


Figura nº 6: Porcentaje de compras v/s no compras
Fuente: Elaboración propia

En la tabla nº3 se muestran algunos estadísticos descriptivos adicionales sobre la frecuencia de compra y canje de los clientes.

	Media	Desv. Estándar	Percentil 5%	Percentil 95%
Frec. Compra	0,824	0,205	0,398	0,999
Frec Canje	0,105	0,099	0,0215	0,247

Tabla nº 3: Estadísticos descriptivos para la frecuencia de compra y canje
Fuente: Elaboración propia

De lo anterior, se puede observar que existe una gran cantidad de ocasiones de compra concretadas en comparación con las no concretadas. Esto se puede explicar por el hecho de contar con datos agregados a nivel mensual. A su vez, el número de canjes es relativamente reducido, con una frecuencia de 0,105. Era esperable que la cantidad de canjes tuviera una frecuencia relativamente baja, pues la cantidad requerida de puntos para canjear un premio es alta, incluso en el nivel de canje más

¹⁶ Una ocasión de compra (o de canje) se define para efectos de este estudio como un mes donde un cliente tuvo la oportunidad de comprar (y por ende, de canjear). Cada mes del período de estudio representa una ocasión de compra, aunque no siempre se concreta dicha compra pues los clientes no compran necesariamente todos los meses.

bajo, por lo que es esperable que se necesiten varias compras (y por ende, varios meses) para alcanzar los puntos necesarios¹⁷.

Para mostrar de mejor forma los datos, a continuación se realizarán algunos análisis a nivel mensual. En la figura n°7, se muestra la cantidad de compras promedio mensuales en un año.

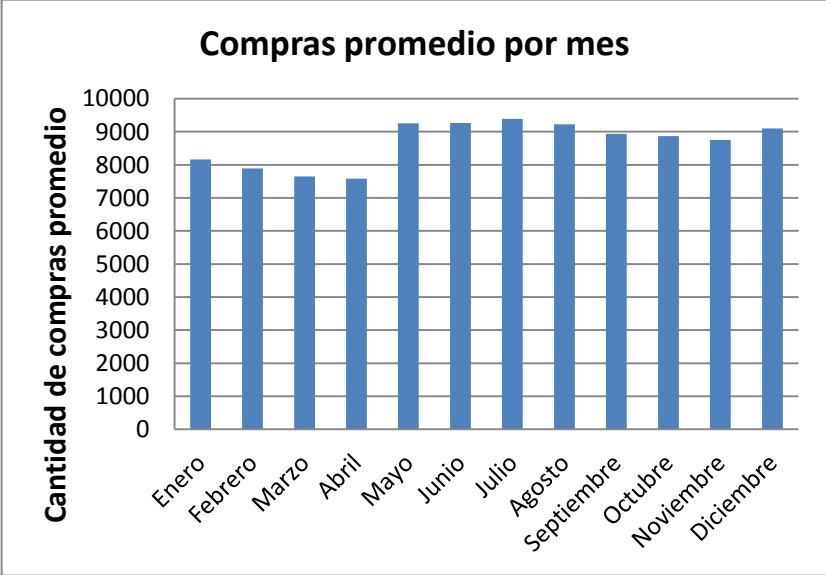


Figura n° 7: Cantidad de compras promedio por mes
Fuente: Elaboración propia

El gráfico anterior, aun cuando es útil para estudiar los datos, puede llevar a conclusiones erróneas pues hay meses que cuentan con más ocasiones de compra que otros y, por ende, la comparación del promedio de compras entre los distintos meses no es directa. Esto se debe a que hay clientes que entran al programa en forma tardía (después de Marzo del 2003) y también a la censura de datos. Por esta razón, en la figura n°8 se muestra el porcentaje de compras concretadas a nivel mensual sobre el total de ocasiones de compra promedio en un mes.

¹⁷ Por ejemplo, el nivel de canje más bajo requiere de 5.000 puntos, que equivalen a \$600.000.

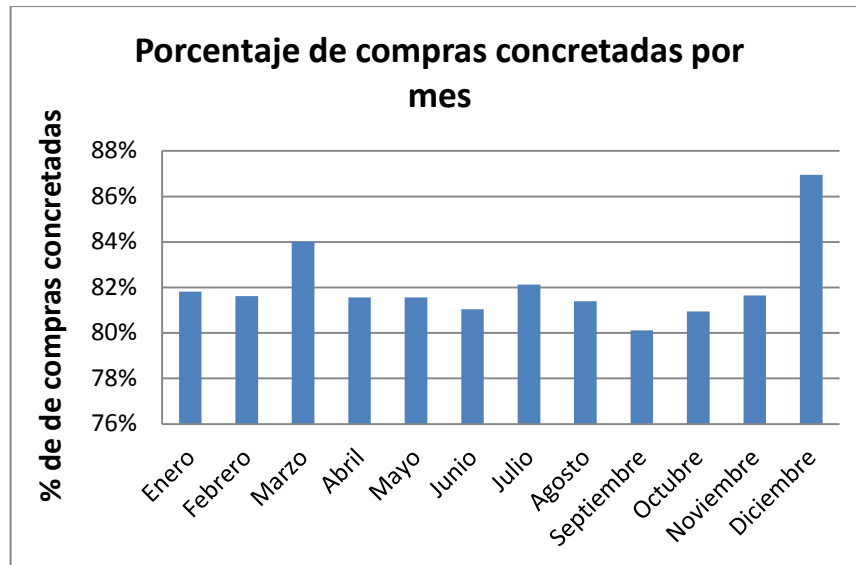


Figura nº 8: Porcentaje de oportunidades de compra concretadas por mes
Fuente: Elaboración propia

Del gráfico anterior se puede observar claramente que el mes donde se concreta el mayor porcentaje de ocasiones de compra es en diciembre, seguido del mes de marzo. Esto era esperable, pues corresponden a los meses de navidad y comienzo del año escolar, respectivamente, donde la cantidad de compras que realizan las personas es mayor.

Aun cuando es el comportamiento de compra lo que se desea estudiar, los canjes son también muy importantes en este estudio, pues estos son el motivo por el cual el cliente debería acelerar sus tiempos entre compras sucesivas. En la figura nº9 se muestra el número de canjes promedio a nivel mensual.

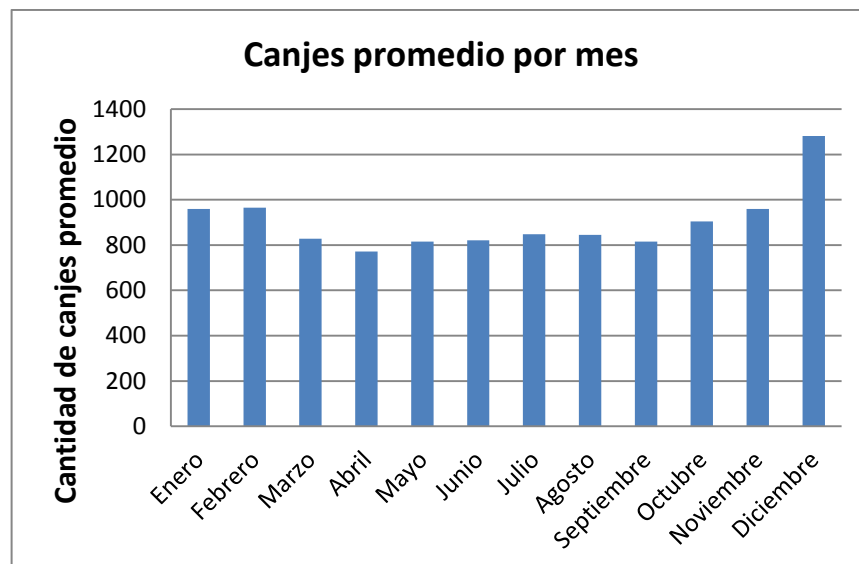


Figura nº 9: Cantidad de canjes promedio por mes
Fuente: Elaboración propia

Del gráfico anterior, se puede observar que es en diciembre donde se realiza la mayor cantidad de canjes promedio, sin embargo, y al igual que en el caso de las compras, estos resultados no deberían compararse directamente entre sí. Por ello, en la figura nº10 se muestra el porcentaje de canjes concretadas a nivel mensual sobre el total de ocasiones para canjear en un mes.

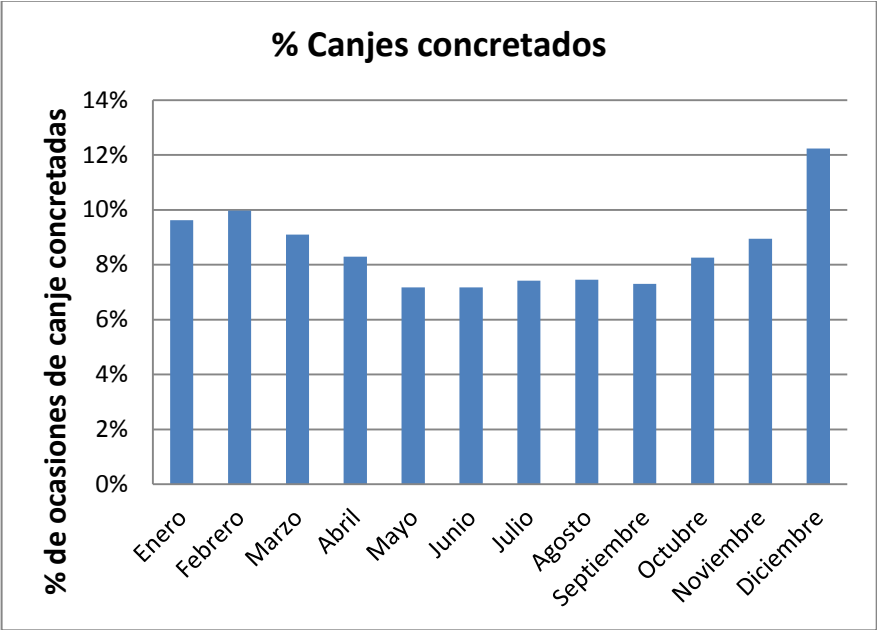


Figura nº 10: Porcentaje de oportunidades de canje concretadas por mes
Fuente: Elaboración propia

De este último gráfico se observa que la actividad de los canjes tiene un comportamiento muy marcado, siendo los meses de verano donde se concretan más canjes (de diciembre a marzo) y en invierno donde se concretan menos (de mayo a agosto). Era esperable que en diciembre se concretara un mayor porcentaje de canjes, pues en dicho mes los clientes deben realizar sus compras navideñas, y al utilizar los puntos para canjear pueden suavizar un poco los montos gastados. Además, siendo que es en diciembre donde el promedio de compras concretadas es mayor (ver gráfico nº4), se puede suponer que acumularán muchos puntos, los que pueden gastar inmediatamente.

6.2.2 Montos

Además de las compras y canjes, en la base de datos se cuenta con los montos invertidos por cada cliente a nivel mensual. En la figura nº11 se muestra el monto promedio gastado por cliente mensualmente, dado que realizaron una compra.

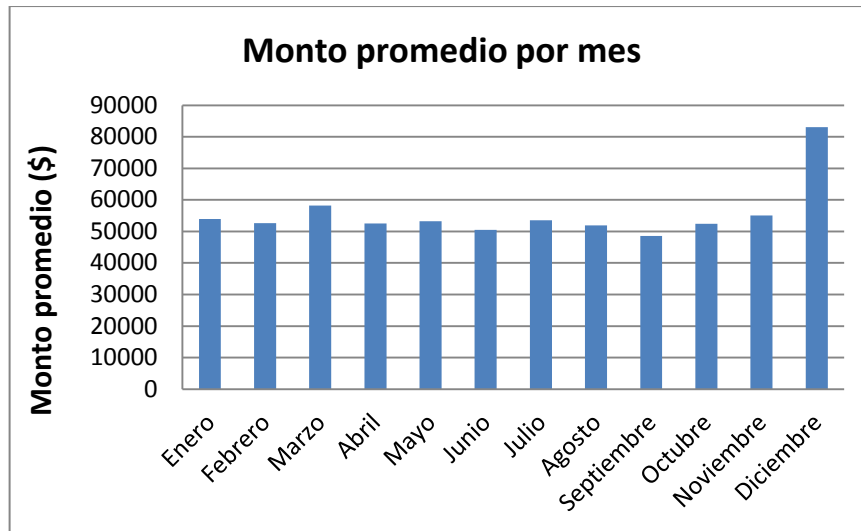


Figura nº 11: Monto promedio gastado por cliente, por mes
Fuente: Elaboración propia

Consistente con lo visto anteriormente, se observa que el monto promedio gastado en diciembre es considerablemente mayor al resto de los meses, seguido por el mes de marzo, lo que nuevamente se puede explicar por navidad y el inicio del año escolar. Estos meses son de gastos importantes para las personas, en general.

6.2.3 Puntos vencidos

Los puntos vencidos son puntos que el cliente pierde después de un año de haberlos ganado. En cada mes, se cuenta con el dato de cuántos puntos vencieron para cada cliente. En la figura nº12 se muestra la cantidad de veces que vencen puntos en promedio, mes a mes.

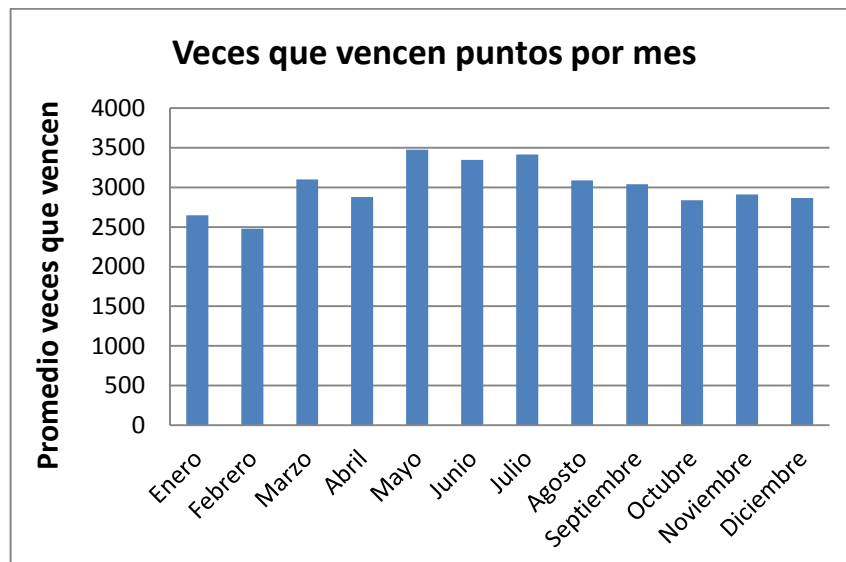


Figura nº 12: Cantidad de veces que vencen puntos en promedio, por mes
Fuente: Elaboración propia

Se puede ver que es a mediados de año cuando ocurre el mayor número de vencimientos de puntos en la base de datos. Al igual que con las compras y los canjes, en el figura nº13 se muestra el porcentaje de veces que vencieron puntos sobre el total de ocasiones en que pudo haber vencimientos en un mes, recordando que no todos los meses tienen la misma cantidad de ocasiones de compra.

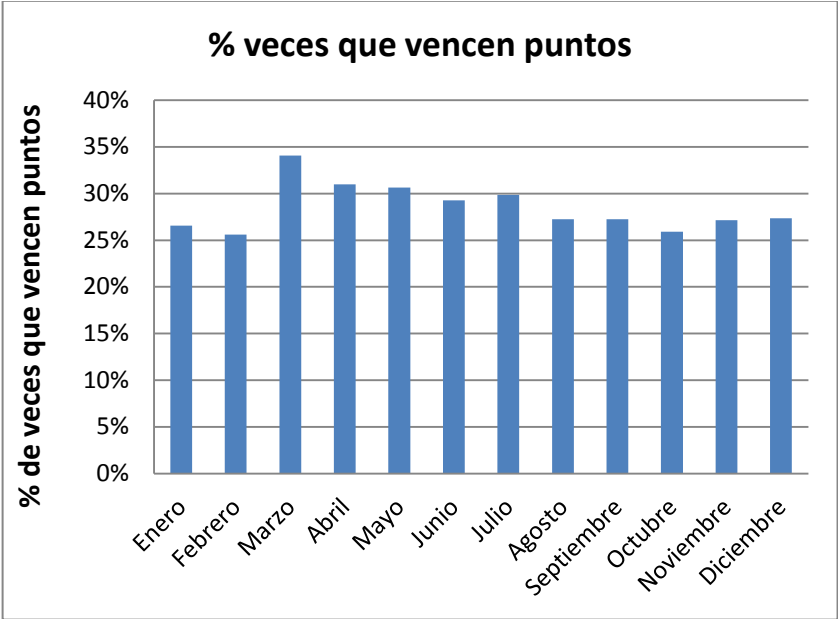


Figura nº 13: Porcentaje de veces que vencen puntos sobre el total de ocasiones
Fuente: Elaboración propia

De aquí se puede observar que es en marzo cuando se da el mayor porcentaje de vencimiento de puntos sobre el total de ocasiones de puntos vencidos, donde en cerca del 35% de las oportunidades de compra al cliente le vencen puntos. Para medir la magnitud de los vencimientos, en la figura nº14 se muestra la cantidad de puntos perdidos en promedio, mensualmente.

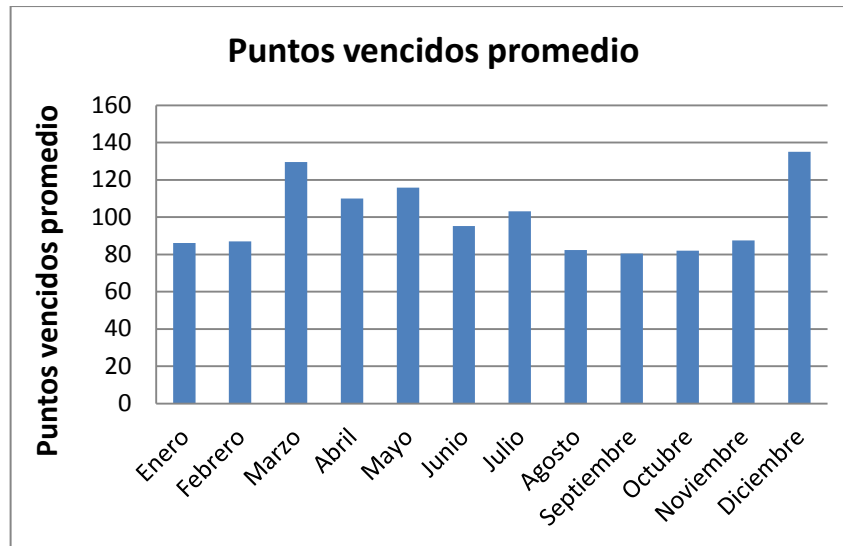


Figura nº 14: Cantidad de puntos vencidos promedio, por mes
Fuente: Elaboración propia

Aun cuando la mayor cantidad de vencimientos se daba a mitad de año y era en marzo cuando el porcentaje de vencimiento de puntos sobre el total de ocasiones de vencimiento era mayor, del gráfico anterior se puede observar que es nuevamente en diciembre donde la cantidad de puntos vencidos es más grande. Esto tiene sentido, pues si es en diciembre donde se gastan los mayores montos promedio (ver gráfico nº11), es en ese mes donde se acumula la mayor cantidad de puntos y, por ende, es el mes donde hay mayor cantidad de puntos potenciales para perder, incluso si la cantidad de gente que pierde esos puntos es menor que en otros meses.

Para finalizar esta sección, se muestran algunos estadísticos adicionales sobre los montos gastados y cantidad de puntos vencidos en la base de datos, en la tabla nº4. Estos corresponden a la media de los montos y puntos vencidos promedio de los clientes, es decir, se calculó primero los montos promedio per cápita y puntos vencido promedio per cápita y a partir de esto se calculó la media, desviación estándar, percentil 5% y percentil 95%.

	Media	Desv. Estándar	Percentil 5%	Percentil 95%
Monto (\$)	75.633	10.861	22.398	179.524
Puntos vencidos	358	548	11	1.035

Tabla nº 4: Estadísticos descriptivos para el monto y los puntos vencidos
Fuente: Elaboración propia

6.2.4 Construcción de bases de datos

A partir de los datos analizados anteriormente, se construyeron cuatro bases de datos, según el nivel del premio que llegaron a canjear, con el fin de determinar si existen diferencias en el comportamiento de compra según el nivel del premio canjeado.

Para aquellos clientes que canjearon en más de un nivel distinto durante el período de estudio, se tomaron los datos asociados a cada nivel de canje y se asociaron con la base de datos respectiva. Se debe notar que sólo se tomaron datos hasta que el cliente llegó a acumular puntos suficientes para canjear, lo que no necesariamente coincide con el momento del canje mismo (podría canjear sus puntos en meses posteriores). Esto se debe a que se desea capturar el efecto de aceleración de compra al premio; una vez que el cliente llega al nivel deseado se dice que llegó a la meta. En la Tabla nº5 se muestran la cantidad de clientes, canjes realizados y el promedio de canjes por cliente para cada una de las bases de datos.

	Clientes	Canjes	Canjes promedio
Nivel 1	17.827	35.114	1,97
Nivel 2	7.011	10.605	1,51
Nivel 3	5.170	8.746	1,69
Nivel 4	879	1.328	1,51

Tabla nº 5: Clientes, canjes y canjes promedio por nivel
Fuente: Elaboración propia

VII COVARIABLES DE LOS MODELOS

Para este estudio, se utilizan distintos modelos, donde existe un factor en común; las covariables Dado que los datos se tienen a nivel mensual, se considera que cada mes es una ocasión de compra, donde esta se puede concretar o no. Por lo tanto, cada fila de las bases de datos generadas corresponden a datos de un cliente particular en un mes específico, donde se cuenta con la información de si en dicho mes se concreta el evento de compra (parámetro binario δ_t , vale 1 si el cliente compra el mes t y 0 sino) y el tiempo en meses transcurrido desde el la última compra del cliente (recency de compra).

En cada fila se cuenta, además, con distintos datos que son utilizados como covariables (variables independientes del modelo), las cuales dependen del cliente y del tiempo. Estas covariables son:

1. Distancia al canje
2. Logaritmo del número de canjes
3. Distancia al canje suavizada por los canjes
4. Share de compra
5. Puntos por vencer
6. Puntos vencidos

7.1 Distancia al canje

Esta covariable representa la distancia psicológica del cliente a un premio, y está basada en el modelo de gradiente-objetivo (*goal-gradient model*), propuesto por Kivetz

et. al. (2006). Este modelo se basa en la idea de que los clientes consideran espontáneamente su distancia a un objetivo, incorporando la distancia total como un punto de referencia, lo que los lleva a una evaluación relativa de la distancia recorrida hasta la meta. Así, la distancia-objetivo se modela como una distancia psicológica que indica la proporción de la distancia que queda por recorrer, con respecto a la distancia total. En otras palabras, la distancia se modela como:

$$d_t = \frac{(r - n_t)}{r} \quad (28)$$

donde:

- d_t : Distancia psicológica al premio en el mes t
- r : Distancia total al objetivo
- n_t : Distancia ya recorrida hasta el mes t

Se debe notar que, por construcción, esta distancia psicológica toma valores entre 0 y 1, donde 1 indica que el consumidor está en el punto de partida (progreso nulo) y 0 indica que ya ha llegado a la meta.

En este estudio, r es la distancia total a un nivel de premios, es decir, la cantidad de puntos total necesarios para canjear un premio, mientras que n_t corresponde a la cantidad de puntos acumulados en el mes t. Se debe notar que el valor específico de r depende de la base de datos que se está considerando (el nivel de canje que se está estudiando), y por ende un cliente al que le falta una cierta cantidad de puntos en un nivel no está necesariamente a la misma distancia psicológica al premio que uno con la misma cantidad de puntos en otro nivel. Por ejemplo, si se considera que un cliente está acumulando puntos para canjear un premio del nivel 1 y lleva acumulados 2.000 puntos, su distancia al premio es $(5.000-2.000)/5.000 = 0,6$. Por otro lado, si otro cliente está juntando puntos para un premio del nivel 2 y ya ha acumulado 6.000 puntos, su distancia al premio es $(9.000-6.000)/9.000 = 0,33$. De lo anterior, se observa que ambos clientes están a 3.000 puntos de cumplir el objetivo, sin embargo la distancia psicológica del primer individuo es mayor que la del segundo debido a que su punto de referencia (el total de puntos necesarios para llegar a la meta) es distinto.

7.2 Logaritmo del número de canjes

Además del efecto psicológico de la distancia al objetivo, existe un factor adicional en el comportamiento de compra del cliente y que el modelo de distancia explicado anteriormente no considera: La experiencia del canje. Se ha estudiado que la motivación de un cliente se ve afectada positivamente por la experiencia del canje. Dicho de otra forma, mientras más veces canjee un cliente, más motivado estará con el programa de lealtad, aunque dicha motivación crece a tasa decreciente con el número de canjes. Esto se explica porque cuando una persona canjea un premio por primera

vez se emociona mucho, mientras que la segunda vez que canjea se emociona un poco menos, y así sucesivamente. Así, esta covariable queda representada por la siguiente expresión:

$$Ln_canjes_t = \ln(1 + canjes_t) \quad (29)$$

donde:

- $canjes_t$: Número de canjes hasta el mes t

El logaritmo en la expresión da cuenta del impacto creciente a tasa decreciente de los canjes.

7.3 Distancia al canje suavizada por los canjes

Esta covariable es una fusión de las 2 covariables anteriores. Corresponde a la distancia al canje, suavizada por el logaritmo del número de canjes. El objetivo es captar ambos efectos en una sola covariable para ver si el modelo se ajusta mejor o peor con esta que con ambas por separado. La covariable queda de la siguiente forma:

$$D_t = \frac{d_t}{1 + \ln(1 + canjes_t)} \quad (30)$$

donde:

- D_t : Distancia psicológica suavizada en el mes t
- d_t : Distancia psicológica al premio en el mes t
- $canjes_t$: Número de canjes hasta el mes t

De lo anterior, cuando un cliente no ha canjeado ninguna vez, la distancia D_t es igual a la distancia psicológica d_t . En cambio, cuando el cliente ha canjeado una o más veces, la distancia D_t es menor que d_t , y por ende el cliente está más cerca del premio (está más motivado, producto de los canjes acumulados con anterioridad).

7.4 Share de compra

Como se ha mostrado en la descripción de los datos, hay meses del año que tienen mayor actividad que otros: en diciembre, por ejemplo, suele haber más compras y canjes debido a la navidad y el mayor volumen de compras que esta fecha implica.

El Share de compra corresponde al porcentaje de compras concretadas por todos los clientes en un mes sobre el total de oportunidades de compra existentes. Se calcula este estadístico para cada mes del período de estudio (62 meses en total) y se utiliza

como covariable para controlar por el efecto de la estacionalidad en el comportamiento de compra de los clientes descrito anteriormente.

7.5 Puntos por vencer

Los puntos que acumulan los clientes tienen una fecha de vencimiento: duran 12 meses. Por lo tanto, en cada instante del período de estudio existe la posibilidad de que un cliente vaya a perder los puntos acumulados hace 12 meses atrás.

Esta covariable indica si un cliente tiene puntos por vencer al final del mes actual, como se muestra a continuación:

$$POR_VENCER_t = \begin{cases} 1 & \text{Si tiene puntos por vencer en mes } t \\ 0 & \text{sino} \end{cases} \quad (31)$$

Se esperaría que un cliente que tiene puntos por vencer tenga una motivación adicional para comprar y canjear un premio, ya que si no lo hace perderá parte de la distancia recorrida hacia él.

7.6. Puntos vencidos

Esta covariable indica si un cliente perdió puntos al final del mes anterior, como se muestra a continuación:

$$VENCIDOS_t = \begin{cases} 1 & \text{Si vencieron puntos en mes } t-1 \\ 0 & \text{sino} \end{cases} \quad (32)$$

. Existe la posibilidad de que los clientes que pierden puntos se sientan defraudados y esto afecte negativamente su comportamiento de compra futuro, lo que justifica el uso de esta covariable.

VIII. ESTUDIOS REALIZADOS

Los programas de lealtad que reconocen y premian la frecuencia de compra de los clientes se han convertido en una de las herramientas de marketing más utilizadas para retener a los clientes y estimular el uso de distintos productos y servicios. En particular, la evidencia de que retener a un cliente es más económico que adquirir uno nuevo y de que hay clientes más rentables que otros, ha llevado a las compañías a establecer programas de lealtad que no sólo premian a los clientes, sino también hacen que sea más difícil para ellos cambiarse de marca (Kivetz y Simonson, 2002)..

A pesar del alto crecimiento y gran popularidad de estos programas, aún no se sabe mucho del impacto real que tienen en las percepciones de los consumidores y cuáles son los factores que impactan en mayor o menor medida en el comportamiento

de compra de sus clientes. Muchos programas de lealtad son creados simplemente para seguir la tendencia de la competencia, sin un estudio previo que acredite su real utilidad para el negocio. Ante esta situación surgen varias preguntas con respecto a si son realmente rentables o no, o si realmente logran impactar favorablemente en el comportamiento de compra de los clientes, lo que se traduciría en una mayor lealtad.

De acuerdo a lo estudiado en el capítulo IV, se plantean una serie de hipótesis que permitirán estudiar el comportamiento de los clientes de dicho programa de lealtad. Estas hipótesis se presentan a continuación.

- H₁: Los clientes del programa de lealtad disminuyen el tiempo transcurrido entre compras sucesivas, a medida que se acercan al nivel de canje de un premio.
- H₂: Los clientes aumentan el esfuerzo invertido en el programa de lealtad, medido en dinero gastado en las compras, a medida que se acercan al nivel de canje de un premio.
- H₃: Los clientes del programa de lealtad reinician su patrón de compras después del canje de un premio.
- H₄: La preferencia por premios hedónicos aumenta al aumentar el esfuerzo necesario para el canje de un premio.
- H₅: Los clientes que canjean premios hedonistas aceleran más que aquellos que canjean premios utilitarios

Cabe destacar que, aun cuando el foco de esta investigación es estudiar las respuestas de los consumidores ante el programa de lealtad, es necesario realizar antes que todo una simulación, para verificar que el modelo de hazard proporcional en tiempo discreto se comporta bien ante las características particulares de los datos utilizados. De esta forma, se podrá tener confianza de que los resultados se pueden interpretar directamente y estos no se deben simplemente a las características especiales de los datos. Por lo tanto, se comienza este estudio con dicha simulación y, posteriormente, se analizarán cada una de las hipótesis planteadas.

8.1 Simulación

Los estudios de simulación son una herramienta importante para investigar el desempeño, propiedades y adecuación de modelos estadísticos, tests estadísticos y técnicas de estimación. El modelo de Cox, en particular, suele ser intensivamente investigado por medio de estudios de simulación de datos, para obtener información acerca de posibles sesgos y la eficiencia de los estimadores de la regresión, en una variedad de situaciones diferentes (Bender, Augustin, Blettner, 2003).

Para efectos de este estudio, resulta necesario cuestionarse si el modelo de Cox es adecuado, dadas las características de los datos. Como se señaló en el capítulo 6.2.1, los clientes del programa de lealtad compran con una frecuencia considerablemente alta, lo que se traduce en una gran cantidad de datos con valor igual a 1 en la columna de compras en la base de datos. Por esta razón, existe la posibilidad de que el modelo no se comporte bien bajo esta condición y se hace importante simular datos con estas características para asegurarse de que la estimación de los parámetros se realiza correctamente. Por otro lado, se desea verificar también el comportamiento del modelo ante situaciones especiales que se indicarán posteriormente.

8.1.1 Consideraciones generales

Lo que se desea simular son los tiempos entre compras sucesivas de los clientes. Para ello, se debe desarrollar una expresión que permita generarlos, según las distintas parametrizaciones utilizadas en capítulos anteriores (exponencial, weibull y log-logística). Como se ha señalado en el capítulo 5.2, la función de distribución acumulada del modelo de Cox proporcional en tiempo discreto está dada por:

$$F(t | x) = 1 - \exp(-H_0(t) * \exp(\beta * X_t)) \quad (33)$$

donde:

$$H_0(t) = \int_0^t h(t) dt \quad (34)$$

Se define Y , una variable aleatoria con función de distribución F , tal que $U=F(Y)$ sigue una distribución uniforme(0,1). Dado que U sigue una uniforme(0,1), entonces $(1-U)$ también sigue una uniforme(0,1), por lo tanto:

$$U = \exp(-H_0(t) * \exp(\beta * X_t)) \sim Uniforme(0,1) \quad (35)$$

De la expresión (35), si $h(t) > 0$ (para todo t) de tal forma que H sea invertible, se puede obtener una expresión para los tiempos de sobrevivencia:

$$T = H_0^{-1}(-\log(U) * \exp(-\beta * X_t)) \quad (36)$$

Esta expresión se utilizará para simular los tiempos entre compras sucesivas. Como se puede observar, al ser una función de H_0^{-1} , los tiempos dependen de la parametrización elegida, por lo que a continuación se desarrollarán los 3 casos respectivos.

8.1.1.1 Caso distribución exponencial:

En este caso,

$$H_0(t) = \int_0^t \gamma dt = \gamma t \Rightarrow H_0^{-1}(t) = \gamma^{-1}t \quad (37)$$

y reemplazando en (36) se obtiene

$$T = -\frac{\log(U)}{\gamma * \exp(-\beta * X_t)} \quad (38)$$

8.1.1.2 Caso distribución Weibull:

En este caso,

$$H_0(t) = \int_0^t \gamma \alpha (\gamma t)^{\alpha-1} dt = (\gamma t)^\alpha \Rightarrow H_0^{-1}(t) = \gamma^{-1}t^{1/\alpha} \quad (39)$$

y reemplazando en (36) se obtiene

$$T = \frac{1}{\gamma} * \left(-\frac{\log(U)}{\exp(-\beta * X_t)} \right)^{1/\alpha} \quad (40)$$

8.1.1.3 Caso distribución log-logística:

En este caso:

$$H_0(t) = \int_0^t \frac{\gamma \alpha (\gamma t)^{\alpha-1}}{(1 + (\gamma t)^\alpha)} dt = \log(1 + (\gamma t)^\alpha) \Rightarrow H_0^{-1}(t) = \gamma^{-1}(\exp(t) - 1)^{1/\alpha} \quad (41)$$

y reemplazando en (36) se obtiene

$$T = \frac{1}{\gamma} * \left(\exp\left(-\frac{\log(U)}{\exp(-\beta * X_t)} \right) - 1 \right)^{1/\alpha} \quad (42)$$

8.1.2 Generando los datos

Para generar los datos, se siguen los siguientes pasos para cada una de las parametrizaciones del modelo de Cox utilizadas:

1. Se fijan los valores para los parámetros beta, gamma y alfa arbitrariamente (si corresponde).
2. Se simula una variable u , uniforme(0,1)
3. Se calcula la distancia al canje normalizada, suponiendo que el cliente acumula 1000 puntos por cada compra.
4. Se evalúa la expresión (36) para el tiempo T , con los datos calculados anteriormente (U y distancia al canje) y según la parametrización elegida.
5. Se discretiza T , aplicándole cajón superior.
6. Esto se repite N veces, siendo N el número de compras que se desea simular.

Con esto, para cada una de las N compras se simula un tiempo T que corresponde al recency de las mismas. Así, se construye la base de datos respectiva y se corre con el modelo Cox según las distintas parametrizaciones. Para efectos de la simulación realizada en este estudio, se calibró el valor de N y se encontró que $N=100.000$ era suficientemente grande para evitar problemas por el tamaño de la muestra.

8.1.3 Metodología y resultados

Una vez generados los datos, se corre el modelo de hazard proporcional en tiempo discreto para ver si este modelo es capaz de estimar los valores de los parámetros utilizados para generarlos. Se espera que si, por ejemplo, se generan mediante una distribución Weibull, sea esta parametrización del hazard la que se ajuste mejor a los datos, pues con ella fueron generados.

8.1.3.1 Datos generados según distribución exponencial

Los datos se generaron con los siguientes valores arbitrarios, pero elegidos de tal forma de que se obtengan muchas compras concretadas:

$$\beta = -1,2$$
$$\gamma = 3,7$$

Con esto se obtuvo una base de datos con un total de 120.975 ocasiones de compra, donde el 80.7% de ellas se concretan, muy similar al 82,4% de las ocasiones concretadas en los datos originales. En la tabla nº6, se presentan los resultados de la estimación de los parámetros, según las distintas parametrizaciones del modelo de Cox:

	β	γ	α	LL	BIC
Exponencial	-1,190 (0,038)	3,658 (0,010)	-	-56.355	112.733
Weibull	-1,180 (0,024)	4,029 (0,180)	0,935 (0,074)	-56.294	112.623
Log-logística	-1,147 (0,048)	1,939 (0,161)	5,434 (0,089)	-56.791	113.617

Tabla nº 6: Estimación de parámetros para simulación exponencial
Fuente: Elaboración propia

Se puede observar que el mejor modelo para estimar los parámetros es el de parametrización Weibull, aunque por una diferencia pequeña con respecto a al exponencial. Esto se puede explicar por el hecho de que la exponencial es un caso particular de la Weibull, donde $\alpha = 1$. Por lo tanto, es esperable que la Weibull sea capaz de ajustarse igual o mejor a los datos que la exponencial. Se puede observar que, de hecho, α es cercano a 1. La log-logística, en cambio, resultó peor que las otras dos distribuciones.

Finalmente, en el caso de la mejor estimación (Weibull), no existe una diferencia significativa entre el β simulado y el β estimado, según test t, $P > 0,1$.

8.1.3.2 Datos generados según distribución weibull

Los datos se generaron con los siguientes valores arbitrarios, pero elegidos de tal forma de que se obtengan muchas compras concretadas:

:

$$\beta = -0,9$$

$$\gamma = 26$$

$$\alpha = 0,44$$

Con esto se obtuvo una base de datos con un total de 123.005 ocasiones de compra, donde el 81.3% de ellas se concretan. En la tabla nº7, se presentan los resultados de la estimación de los parámetros, según las distintas parametrizaciones del modelo de Cox:

	β	γ	α	LL	BIC
Exponencial	-1,371 (0,059)	4,256 (0,151)	-	-53.962	107.947
Weibull	-0,912 (0,015)	25,79 (0,120)	0,431 (0,068)	-41.775	83.585
Log-logística	-0.930 (0,056)	5,430 (0,032)	2,618 (0,159)	-42.071	84.177

Tabla nº 7: Estimación de parámetros para simulación Weibull
Fuente: Elaboración propia

Como era de esperarse, la Weibull tiene el mejor ajuste, seguido de cerca por la Log-logística, mientras que la exponencial presenta el peor ajuste. Se observa que en el caso de la mejor estimación (Weibull), no existe una diferencia significativa entre el β simulado y el β estimado según test t, $P > 0.1$.

8.1.3.3 Datos generados según distribución log-logística

Los datos se generaron con los siguientes valores arbitrarios, pero elegidos de tal forma de que se obtengan muchas compras concretadas:

$$\beta = -0,8$$

$$\gamma = 5$$

$$\alpha = 2.4$$

Con esto se obtuvo una base de datos con un total de 125.045 ocasiones de compra, donde el 79,9% de ellas se concretan. En la tabla nº8, se presentan los resultados de la estimación de los parámetros, según las distintas parametrizaciones del modelo de Cox:

	β	γ	α	LL	BIC
Exponencial	-1,241 (0,105)	3,696 (0,017)	- (-)	-58.065	116.153
Weibull	-0.873 (0,004)	29,33 (0,166)	0.419 (0,000)	-45.130	90.295
Log-logística	-0,815 (0,016)	4,992 (0,063)	2,510 (0,032)	-45.045	90.125

Tabla nº 8: Estimación de parámetros para simulación Log-logística
Fuente: Elaboración propia

Como era de esperarse, la log-logística resulta ser la de mejor ajuste. Se puede notar que la de peor ajuste es nuevamente la exponencial. Se observa que en el caso

de la mejor estimación (Log-logística), no existe una diferencia significativa entre el β simulado y el β estimado, según test t, $P > 0,1$.

8.1.4 Simulación con $\beta = 0$

Ahora se desea estudiar si el modelo es capaz de estimar correctamente los parámetros cuando los tiempos entre compras se generan con un $\beta = 0$. Es decir, en este caso, al generar los datos, la distancia al canje no jugará ningún rol y se desea estudiar si luego el modelo de hazard proporcional en tiempo discreto es capaz de detectar esta situación, o si entregará estimadores incorrectos.

8.1.4.1 Datos generados según distribución exponencial

Se simulan los tiempos entre compras utilizando los siguientes parámetros:

$$\beta = 0$$

$$\gamma = 2$$

Con esto se obtuvo una base de datos con un total de 115.149 ocasiones de compra, donde el 86.7% de ellas se concretan. En la tabla nº9, se muestran los resultados para la parametrización exponencial del modelo de Cox.

	β	γ	α	LL	BIC
Exponencial	0,003* (0,417)	1,989 (0,592)	-	-46.043	92.109

*No son significativos, $P > 0,01$ test t de dos colas.

Tabla nº 9: Estimación de parámetros caso exponencial

Fuente: Elaboración propia

Se puede observar que no existe diferencia significativa entre el β estimado y el β simulado, según test t, $P > 0,01$.

8.1.4.2 Datos generados según distribución weibull

Se simulan datos con los siguientes parámetros:

$$\beta = 0$$

$$\gamma = 10$$

$$\alpha = 0,4$$

Con esto se obtuvo una base de datos con un total de 118.428 ocasiones de compra, donde el 84,4% de ellas se concretan. En la tabla nº10, se muestran los resultados para la parametrización Weibull del modelo de Cox.

:	β	γ	α	LL	BIC
Weibull	0,003* (0,158)	10,03 (0,079)	0,436 (0,031)	-45.627	91.289

*No son significativos, $P > 0.01$ test t de dos colas.

Tabla nº 10: Estimación de parámetros caso Weibull

Fuente: Elaboración propia

Se puede observar que no existe diferencia significativa entre el β estimado y el β simulado, según test t, $P > 0.01$.

8.1.4.3 Datos generados según distribución log-logística

Se simulan datos con los siguientes parámetros:

$$\beta = 0$$

$$\gamma = 4$$

$$\alpha = 1,5$$

Con esto se obtuvo una base de datos con un total de 125.433 ocasiones de compra, donde el 79,7% de ellas se concretan. En la tabla nº11, se muestran los resultados para la parametrización Log-logística del modelo de Cox.

	β	γ	α	LL	BIC
Log-logística	0,015* (0,141)	3,840 (0,071)	1,517 (0,031)	-55.754	111.543

*No son significativos, $P > 0.01$ test t de dos colas.

Tabla nº 11: Estimación de parámetros caso Log-logística

Fuente: Elaboración propia

Se puede observar que no existe diferencia significativa entre el β estimado y el β simulado, según test t, $P > 0.01$.

De todo lo anterior, se concluye que el modelo de Cox utilizado sí es capaz de predecir los parámetros correctamente cuando la distancia al canje no juega un rol en el comportamiento de compra de los clientes.

8.1.5 Simulación con covariable de tiempo

En esta sección, se utilizará el tiempo que lleva el cliente en el programa de lealtad como covariable para generar los datos, en lugar de usar la distancia al canje normalizada, como se ha hecho hasta ahora. Se espera con esto que, al estimar los parámetros con el modelo de Cox habitual, se encuentre que el beta asociado a la distancia al canje sea no significativo (pues los datos fueron simulados en base a una covariable distinta). Se decide usar esta covariable para la simulación porque ella

representa el efecto de aprendizaje del cliente. A medida que un consumidor va comprando, cada vez se hace más familiar con el programa por lo que en teoría podría comenzar a comprar en forma más frecuente simplemente porque aprendió a hacerlo y no por la distancia recorrida hacia el premio. Se desea saber si el modelo de Cox proporcional en tiempo discreto se comporta adecuadamente ante esta situación.

8.1.5.1 Datos generados según distribución exponencial

Se simulan los tiempos entre compras utilizando los siguientes parámetros:

$$\beta = -1.2$$

$$\gamma = 2$$

Con estos parámetros se simuló una base de datos de 115.624 ocasiones de compra, donde el 86,4% de ellas se concretan. En la tabla nº12, se muestran los resultados para la parametrización exponencial del modelo de Cox.

	beta	gamma	alfa	LL	BIC
Exponencial	-0,100* (0,417)	2,126 (0,592)	-	-46.043	92.109

*No son significativos, $P > 0.01$ test t de dos colas.

Tabla nº 12: Estimación de parámetros caso exponencial

Fuente: Elaboración propia

Se observa que el parámetro de la distancia al canje estimado es no significativo según test t, $p > 0,01$.

8.1.5.2 Datos generados según distribución weibull

Se simulan los tiempos entre compras utilizando los siguientes parámetros:

$$\beta = -0,8$$

$$\gamma = 5$$

$$\alpha = 0,5$$

Con estos parámetros se simuló una base de datos de 118.471 ocasiones de compra, donde el 84,4% de ellas se concretaron. En la tabla nº13, se muestran los resultados para la parametrización Weibull del modelo de Cox.

	beta	gamma	alfa	LL	BIC
Weibull	-0,082* (0,158)	5,388 (0,079)	0,516 (0,004)	-45.627	91.289

*No son significativos, $P > 0.1$ test t de dos colas.

Tabla nº 13: Estimación de parámetros caso Weibull

Fuente: Elaboración propia

Se observa que el parámetro de la distancia al canje estimado es no significativo según test t, $p > 0,1$.

8.1.5.3 Datos generados según distribución log-logística

Se simulan los tiempos entre compras utilizando los siguientes parámetros:

$$\beta = -0.8$$

$$\gamma = 5$$

$$\alpha = 1.2$$

Con estos parámetros se simuló una base de datos de 131.228 ocasiones de compra, donde el 76.2% de ellas se concretaron. En la tabla nº14, se muestran los resultados para la parametrización Log-logística del modelo de Cox.

	beta	gamma	alfa	LL	BIC
Log-logística	-0,072* (0,141)	4,425 (0,071)	1,411 (0,017)	-55.754	111.543

*No son significativos, $P > 0.1$ test t de dos colas.

Tabla nº 14: Estimación de parámetros caso Log-logística

Fuente: Elaboración propia

Se observa que el parámetro de la distancia al canje estimado es no significativo según test t, $p > 0,1$.

Ahora, dado que se ha mostrado que el modelo hazard proporcional en tiempo discreto entrega resultados correctos dadas las particularidades de los datos utilizados en esta investigación, se procede a continuación a probar cada una de las hipótesis planteadas con anterioridad.

8.2 Hipótesis 1 – Aceleración de compra

H_1 : Los clientes del programa de lealtad disminuyen el tiempo transcurrido entre compras sucesivas, a medida que se acercan al nivel de canje de un premio.

Esta hipótesis tiene como objetivo determinar si existe un efecto de aceleración de compra en los clientes, en la medida que su distancia psicológica a un determinado premio del programa de lealtad disminuye. Esta aceleración se traduce en una disminución en los tiempos entre compras sucesivas.

8.2.1 Análisis preliminar y motivación

Antes de probar la hipótesis 1, se realizan algunos análisis gráficos con el fin de determinar si existe realmente evidencia de aceleración en los esfuerzos de compra. En la figura nº15, se muestra el tiempo entre compras sucesivas promedio para los 4 niveles de canje, uno al lado del otro. Se puede observar que, a medida que los clientes van acumulando puntos, la frecuencia de compra aumenta.

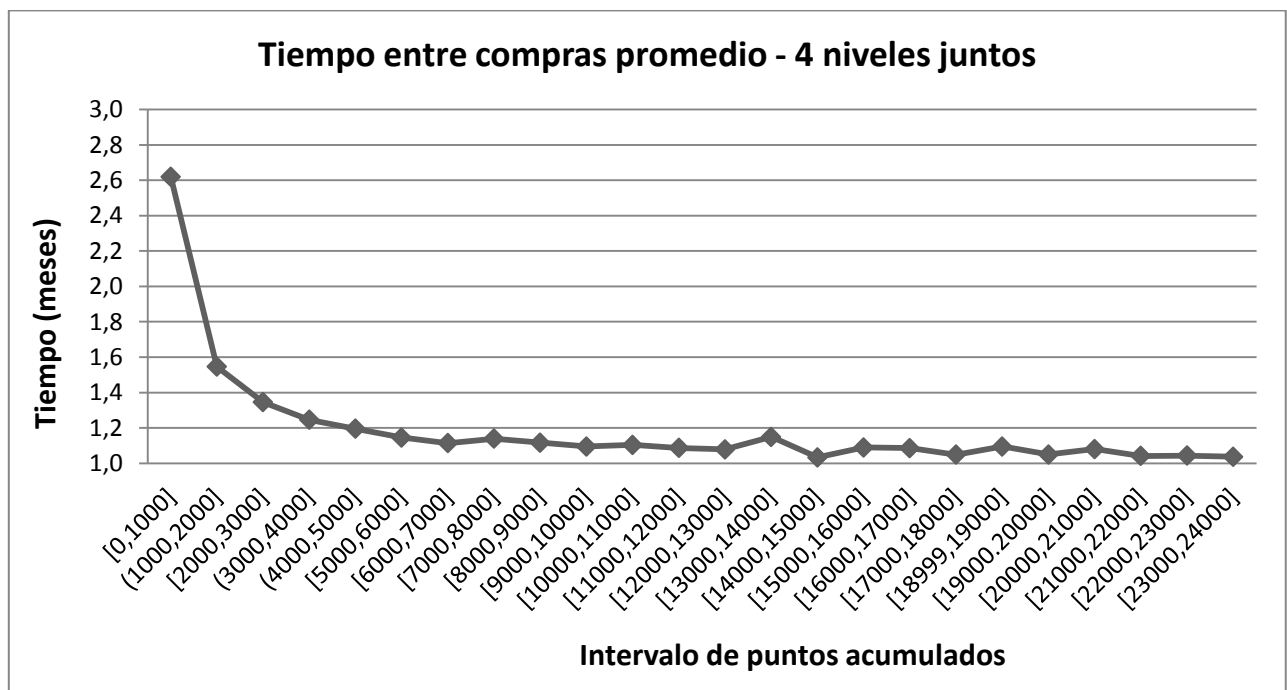


Figura nº 15: Tiempo entre compras promedio para los 4 niveles juntos
Fuente: Elaboración propia

Para observar mejor el fenómeno, en la figura nº16 se muestran los tiempos entre compras promedio, sólo para el nivel 1. De esta se puede observar claramente que la frecuencia va aumentando a medida que se tienen más puntos, lo que podría indicar que hay un efecto de aceleración de compra. En efecto, una persona que tiene entre 0 y 1000 puntos compra en promedio cada 2,6 meses, mientras que una persona entre 1000 y 2000 compra en promedio cada 1,5 meses, y así sucesivamente. La diferencia

entre una persona en el rango de puntos más bajo (entre 0 y 1000) y una persona en el rango más alto (entre 4000 y 5000) es de 219%.



Figura nº 16: Tiempo entre compras promedio para el nivel 1
Fuente: Elaboración propia

De los análisis anteriores se puede concluir, a primera vista, que podría existir un efecto de aceleración de compra en los clientes del programa de lealtad. Las personas compran más seguido a medida que acumulan más puntos y se acercan al nivel de canje. Este efecto es particularmente visible en el nivel 1, aunque no es muy claro para el resto de los niveles. Todo esto sirve como motivador para estudiar más a fondo la hipótesis 1, para lo cual se realizará un estudio donde se utilizará un modelo hazard proporcional en tiempo discreto. Con el modelo, se podrán medir distintos efectos externos que podrían estar ocasionando este comportamiento, con lo que se espera validar estos primeros resultados.

8.2.2 Modelos

El modelo utilizado para testear la hipótesis 1 es el modelo de hazard proporcional en tiempo discreto¹⁸. Para correrlo, se van incorporando distintas combinaciones de las covariables definidas en el capítulo 7. El objetivo es ir agregando distintos efectos que podrían incidir en el comportamiento del consumidor y ver si el ajuste penalizado del modelo (según el criterio de información bayesiana) es mejor o peor según las distintas combinaciones.

¹⁸ Ver capítulo 5.2 para ver una explicación detallada del modelo

Se prueban 4 combinaciones diferentes de las covariables, lo que genera cuatro modelos diferentes.

Modelo 1: Distancia al canje y logaritmo del número de canjes

El modelo hazard se corre primero con las covariables de distancia al canje y el logaritmo del número de canjes (ver capítulo 7.1 y 7.2 respectivamente). La función de probabilidad del hazard queda como se muestra a continuación:

$$\Pr(t, X_t) = 1 - \exp\left(-e^{(\beta_1 d_t + \beta_2 \ln_canjes_t)} \int_{t-1}^t h(s) ds\right) \quad (43)$$

Modelo 2: Distancia al canje suavizada por los canjes

El segundo modelo incorpora una sola covariable: La distancia al canje suavizada por el número de canjes (ver capítulo 7.3). Se desea estudiar si es mejor juntar ambas variables en una sola, o bien mantenerlas por separado, recordando que el criterio de información bayesiana, utilizado para comparar los modelos, penaliza por la cantidad de covariables utilizadas. La función de probabilidad queda como se muestra a continuación:

$$\Pr(t, X_t) = 1 - \exp\left(-e^{(\beta D_t)} \int_{t-1}^t h(s) ds\right) \quad (44)$$

Modelo 3: Distancia, número de canjes, share de compra, puntos por vencer y puntos vencidos

En el tercer modelo se decide agregar las otras covariables al modelo 1 para determinar si éstas explican también el comportamiento de compra del cliente, o bien si éste sólo acelera debido a la distancia psicológica que lo separa del canje. La función de probabilidad queda como se muestra a continuación:

$$\Pr(t, X_t) = 1 - \exp\left(-e^{(\beta_1 d_t + \beta_2 \ln_canjes_t + \beta_3 SHARE_t + \beta_4 VENCIDOS_t + \beta_5 POR_VENCER)} \int_{t-1}^t h(s) ds\right) \quad (45)$$

Modelo 4: Distancia al canje suavizada, share de compra, puntos por vencer y puntos vencidos

Para finalizar, se decide correr el modelo 3, pero cambiando la distancia al canje d_t y el número de canjes \ln_canjes_t por la covariable de distancia suavizada D_t . La función de probabilidad queda como se muestra a continuación:

$$\Pr(t, X_t) = 1 - \exp \left(- e^{(\beta_1 D_t + \beta_2 SHARE_t + \beta_3 VENCIDOS_t + \beta_4 POR_VENCER)} \int_{t-1}^t h(s) ds \right) \quad (46)$$

8.2.3 Resultados

Para estimar los parámetros, se probaron distintas parametrizaciones del modelo hazard proporcional en tiempo discreto (funciones exponencial, weibull y log-logística), además del modelo no paramétrico, para cada uno de los cuatro modelos planteados anteriormente. El ajuste se midió en base al logaritmo natural de la verosimilitud (LL) y el criterio de información bayesiano (BIC). En las tablas nº15, 16, 17 y 18 se muestran tanto el ajuste como el BIC para los modelos 1, 2, 3 y 4, respectivamente.

	Exponencial		Weibull		Log-logística		No paramétrico	
	LL	BIC	LL	BIC	LL	BIC	LL	BIC
Nivel 1	-204.888	409.815	-177.028	354.108	-178.241	356.534	-180.050	360.803
Nivel 2	-40.212	80.460	-33.726	67.498	-33.905	67.856	-34.494	69.533
Nivel 3	-25.871	51.776	-21.087	42.218	-21.184	42.413	-21.802	44.125
Nivel 4	-2.544	5.115	-2.039	4.115	-2.041	4.119	-2.096	4.656

Tabla nº 15: Ajuste del modelo 1 según las distintas parametrizaciones, para todos los niveles
Fuente: Elaboración propia.

	Exponencial		Weibull		Log-logística		No paramétrico	
	LL	BIC	LL	BIC	LL	BIC	LL	BIC
Nivel 1	-204.536	409.098	-177.090	354.219	-178.392	356.823	-178.821	358.332
Nivel 2	-39.973	79.969	-33.695	67.424	-33.895	67.825	-34.026	68.586
Nivel 3	-25.775	51.572	-21.104	42.241	-21.214	42.463	-21.571	43.650
Nivel 4	-2.524	5.066	-2.029	4.086	-2.042	4.112	-2.075	4.604

Tabla nº 16: Ajuste del modelo 2 según las distintas parametrizaciones, para todos los niveles
Fuente: Elaboración propia.

	Exponencial		Weibull		Log-logística		No paramétrico	
	LL	BIC	LL	BIC	LL	BIC	LL	BIC
Nivel 1	-202.037	404.152	-174.763	349.617	-174.998	350.087	-179.494	359.730
Nivel 2	-39.947	79.964	-33.490	67.062	-33.545	67.171	-34.037	68.654
Nivel 3	-25.756	51.580	-20.998	42.075	-21.043	42.166	-21.369	43.291
Nivel 4	-2.542	5.140	-2.031	4.128	-2.035	4.136	-2.062	4.615

Tabla nº 17: Ajuste del modelo 3 según las distintas parametrizaciones, para todos los niveles
Fuente: Elaboración propia.

	Exponencial		Weibull		Log-logística		No paramétrico	
	LL	BIC	LL	BIC	LL	BIC	LL	BIC
Nivel 1	-201.157	402.379	-174.448	348.974	-174.683	349.444	-179.616	359.961
Nivel 2	-39.602	79.262	-33.376	66.770	-33.424	66.918	-33.954	68.475
Nivel 3	-25.608	51.273	-20.949	41.965	-20.983	42.033	-21.328	43.199
Nivel 4	-2.843	5.732	-2.016	4.087	-2.034	4.123	-2.058	4.597

Tabla nº 18: Ajuste del modelo 4 según las distintas parametrizaciones, para todos los niveles
Fuente: Elaboración propia.

Se puede observar que, para todos los modelos, se encontró que la parametrización Weibull es la de mejor BIC en todos los niveles de canje. Por otro lado, al comparar los cuatro modelos entre sí (según la Weibull), se observa que es el modelo 4 el de mejor BIC en todos los niveles y, por lo tanto, el mejor de todos. En la tabla nº19 se muestran los resultados para este modelo, con parametrización Weibull¹⁹.

	β_1 (Dt)	β_2 (SHAREt)	β_3 (VENCIDost)	β_4 (X_VENCERT)	γ	α	LL	BIC
Nivel 1	-0,814 (0,060)	1,952 (0,183)	-0,039* (0,032)	0,053* (0,045)	0,445 (0,033)	0,500 (0,004)	-174.448	348.974
Nivel 2	-0,898 (0,000)	1,729 (0,007)	-0,038* (0,031)	0,061* (0,056)	0,800 (0,016)	0,461 (0,000)	-33.376	66.770
Nivel 3	-0,954 (0,000)	1,572 (0,011)	0,013* (0,011)	0,059* (0,056)	1,304 (0,084)	0,430 (0,000)	-20.949	41.965
Nivel 4	-0,979 (0,004)	1,938 (0,158)	-0,055* (0,044)	0,017* (0,015)	0,796 (0,026)	0,417 (0,000)	-2.016	4.087

* No son estadísticamente distintos de cero, según test t, P>0.1

Tabla nº 19: Parámetros estimados del modelo 4, con parametrización weibull

Fuente: Elaboración propia

Se observa que el parámetro estimado asociado a la distancia al canje es menor que cero para todos los niveles, lo que implica que, a medida que la distancia al premio disminuye, la probabilidad de compra aumenta. Esto es consistente con H1, pues se traduce en una aceleración de compra por parte del cliente a medida que acumula más puntos y se acerca al canje de un premio. Por otro lado, el estimador asociado al share de compra es positivo, es decir, a medida que aumenta el porcentaje de compras concretadas en un mes, mayor es la probabilidad de compra. Este es un resultado relevante pues el que este parámetro sea significativo indica que la covariable del share está capturando efectivamente un efecto de estacionalidad. De todas formas, la covariable de distancia sigue siendo importante en el modelo y de hecho su parámetro no varía mucho con respecto al modelo 2²⁰.

¹⁹ Para ver las estimaciones de los parámetros de los cuatro modelos, según todas las parametrizaciones señaladas, ver anexos A, B, C y D.

²⁰ Ver anexo B para comparar con el modelo 2, parametrización Weibull

Tanto los parámetros estimados asociados a si hay o no puntos por vencer o puntos vencidos resultaron no ser significativos, por lo que no se puede sacar conclusiones al respecto.

8.2.4 Discusión

De los resultados anteriores, se puede concluir que la distancia al canje explica la aceleración de compra de los individuos, a través de todos los niveles de canje del programa. A medida que los consumidores se acercan al canje de un premio, la probabilidad de compra aumenta, lo que es consistente con la hipótesis 1. El share de canje resultó ser un factor importante para explicar el comportamiento, capturando la estacionalidad, mientras que los puntos vencidos y por vencer no resultaron ser significativos en ninguno de los modelos donde se incorporaron, para ninguno de los niveles. El mejor modelo resultó ser el modelo 4, con 4 covariables, utilizando la distancia suavizada por los canjes en lugar de mantener ambas por separado.

Un análisis que queda por realizar es la comparación de los resultados a través de los niveles. Para ello, se medirá el impacto de una variación en la distancia al canje en la probabilidad de compra. Se calcularán las probabilidades al comenzar una carrera (distancia igual a 1) y al finalizar (distancia igual a 0), y se verá el incremento relativo en ella. En la tabla nº20, se muestra la probabilidad de compra al inicio y al final de la carrera, así como la variación relativa, para todos los niveles según el modelo 4²¹.

	$P(D_t=1)$	$P(D_t=0)$	Incremento %
Nivel 1	0,77	0,96	25,6%
Nivel 2	0,81	0,98	21,8%
Nivel 3	0,82	0,99	20,0%
Nivel 4	0,87	1,00	14,5%

Tabla nº 20: Influencia de la distancia sobre la probabilidad, modelo 4
Fuente: Elaboración propia

De la tabla anterior se puede observar que el incremento relativo en la probabilidad entre un nivel y el siguiente es decreciente. Esto se traduce en que el impacto psicológico de la distancia al canje es menor en los clientes de niveles superiores. Esto tiene sentido y era de esperarse, pues en niveles más altos, la cantidad de dinero requerida para canjear es muy grande. Por lo tanto, los clientes que canjean en estos niveles están acostumbrados a comprar constantemente, y el programa de lealtad no representa una motivación tan grande como para alguien que no compra tan seguido.

²¹ Para el mismo análisis en los otros modelos, ver el anexo E. Los resultados son análogos.

8.2.5 Heterogeneidad

Un aspecto importante que se debe considerar, es el sesgo por heterogeneidad. En efecto, al realizar una estimación de parámetros homogéneos, es posible que se observe un efecto aparente de aceleración, pero que dicho efecto en realidad no exista entre la mayoría de los clientes individualmente. Para evitar este sesgo, se debe calibrar el modelo de forma tal que se permita la heterogeneidad en los parámetros a través de los individuos.

Para incorporar el efecto de la heterogeneidad, se hará un análisis de clases latentes. El análisis de clases latentes permite detectar la heterogeneidad de una población, identificando el menor número de grupos en los que se puede agrupar a sus miembros. Cada uno de estos grupos tendrá entonces sus propias características y, por ende, sus propios parámetros a estimar. Además, cada grupo tendrá una cierta probabilidad de que un individuo de la población pertenezca a él²².

8.2.5.1 Modelo

Para efectos de este análisis, se considerará que todos los parámetros del modelo pueden variar a través de los individuos (tanto los parámetros asociados a las covariables como a los de la línea base del hazard). Debido a que, en el caso homogéneo, el modelo 4 fue el que entregó mejores resultados, será este al que se incorporará el análisis de heterogeneidad. Para este modelo, al considerar las clases latentes, la función de probabilidad del hazard proporcional en tiempo discreto queda dada por la expresión que se muestra a continuación:

$$P_{ig}(t, X_t) = 1 - \exp\left(-e^{(\beta_{1g}D_t + \beta_{2g}SHARE_t + \beta_{3g}VENCIDOS_t + \beta_{4g}POR_VENCER_t)} \int_{t-1}^t h_g(s) ds\right) \quad (47)$$

En esta expresión, cada uno de los parámetros a estimar varía con respecto a un índice “g”, que representa los grupos o clases a considerar. Luego, la función de verosimilitud queda dada por

$$L = \prod_{i=1}^N \left(\sum_{g=1}^G \left(\prod_{t=1}^{T_i} P_{ig}(t, X_{it})^{\delta_{it}} * (1 - P_{ig}(t, X_{it}))^{1-\delta_{it}} * \pi_g \right) \right) \quad (48)$$

Es importante recordar que la probabilidad π_g es también un parámetro a estimar en este modelo.

²² Ver capítulo 5.2.2 para más detalles del análisis de clases latentes.

8.2.5.2 Resultados

Para estimar los parámetros, se maximiza la verosimilitud indicada anteriormente, usando la parametrización Weibull, que fue la que consistentemente dio los mejores resultados en todos los modelos del caso homogéneo. Este análisis se realizó para cada uno de los 4 niveles de estudio.

Como primer paso, se estima el modelo con una sola clase latente, con lo que se recuperan los resultados del modelo 4 homogéneo. Después, se evalúan los modelos incrementando el número de clases latentes, de uno en uno. Finalmente, se decide cuál es el número de clases latentes que mejor describe a la muestra, usando el criterio de información bayesiana.

Nivel 1

En la tabla nº21, se encuentra la evaluación de los modelos con distinta cantidad de grupos, mostrando para cada uno de ellos el logaritmo de la verosimilitud, así como el estadístico del criterio de información bayesiana (BIC) y la diferencia entre el BIC para un modelo y el anterior.

Nº Grupos	LL	BIC	Diferencia
1	348.896	348.987	-
2	321.016	321.198	27.788,9
3	316.262	316.535	4.662,9
4	314.326	314.691	1.844,9
5	313.056	313.512	1.178,9
6	311.842	312.389	1.122,9
7	311.072	311.710	678,9
8	310.660	311.389	320,9
9	310.514	311.334	54,9
10	310.452	311.363	-29,1

Tabla nº 21: Evaluación del modelo con distinto número de clases latentes
Fuente: Elaboración propia

El modelo con 9 clases latentes es el que presenta un mejor BIC. Los parámetros estimados para este modelo se muestran en la tabla nº 22.

	β_1 (Dt)	β_2 (SHAREt)	β_3 (VENCIDOST)	β_4 (X_VENCERT)	γ	α	π
Grupo 1	-4,394 (0,308)	8,349 (0,392)	-0,198 (0,065)	0,077* (0,065)	0,031 (0,401)	0,868 (0,037)	0,067 (0,116)
Grupo 2	-11,890 (1,167)	9,217 (0,790)	-0,366 (0,161)	0,296* (0,253)	111,280 (0,817)	0,842 (0,057)	0,104 (0,103)
Grupo 3	-1,691 (0,078)	5,284 (0,536)	-0,094 (0,022)	0,201 (0,022)	0,038 (0,406)	0,971 (0,02)	0,094 (0,131)
Grupo 4	-1,667 (0,044)	1,364 (0,180)	0,043* (0,044)	0,096 (0,024)	0,438 (0,172)	0,891 (0,014)	0,069 (0,087)
Grupo 5	-0,918 (0,080)	7,878 (0,608)	0,013* (0,034)	0,059* (0,136)	4,588 (0,699)	0,561 (0,038)	0,059 (0,141)
Grupo 6	-0,436 (0,033)	1,274 (0,167)	-0,061* (0,124)	0,020* (0,024)	3,209 (0,421)	0,343 (0,031)	0,188 (0,087)
Grupo 7	-2,766 (0,863)	9,112 (1,073)	0,203* (0,794)	-1,182* (0,802)	0,547 (0,995)	0,007 (0,655)	0,154 (0,083)
Grupo 8	-0,977 (0,057)	5,425 (0,226)	-0,074 (0,022)	0,099 (0,022)	0,023 (0,21)	0,783 (0,026)	0,166 (0,101)
Grupo 9	-0,764 (0,059)	0,457 (0,182)	-0,081 (0,020)	0,182 (0,020)	2,125 (0,227)	0,669 (0,043)	0,095 (-)**

* No son estadísticamente significativos, según test t de dos colas, $P > 0.1$

** Este es el nivel base de la probabilidad, por lo que la desviación estándar es nula.

Tabla nº 22: Parámetros estimados para el nivel 1, con 9 clases latentes

Fuente: Elaboración propia

Por último, en la tabla nº23 se presenta la probabilidad de compra de los clientes de cada grupo, evaluado en la distancia al canje inicial y final (1 y 0 respectivamente). Con ello, se puede calcular el incremento porcentual en la probabilidad de compra, lo que permite observar a qué grupos afecta en mayor o menor medida el programa de lealtad.

	$P(D_t=1)$	$P(D_t=0)$	Incremento %
Grupo 1	0,380	1,000	163,25%
Grupo 2	0,073	1,000	1.272,69%
Grupo 3	0,427	0,951	122,88%
Grupo 4	0,246	0,777	215,20%
Grupo 5	0,561	0,873	55,56%
Grupo 6	0,929	0,983	5,83%
Grupo 7	0,980	1,000	2,06%
Grupo 8	0,791	0,984	24,48%
Grupo 9	0,686	0,917	33,70%

Tabla nº 23: Probabilidad de compra para cada grupo, al comienzo y al final de la carrera

Fuente: Elaboración propia

Análisis de resultados nivel 1

El primer aspecto importante a destacar es que la distancia al canje normalizada es significativa para todos los grupos, lo que confirma nuevamente la presencia de la aceleración de compra para el nivel 1 a medida que los clientes se acercan al premio.

Se debe notar que la probabilidad de que un cliente pertenezca a cada grupo es relativamente similar a través de los grupos, lo que se traduce en que el tamaño de los grupos es parecido, no hay un grupo considerablemente predominante. Observando la tabla n°23, se encuentra que el grupo 2 es por lejos el grupo en el que el programa de lealtad tiene un mayor impacto, produciendo un incremento en la probabilidad de compra del 1.272,69% desde que comienza una carrera hacia un premio hasta lograr el objetivo. Por otro lado, los grupos 1, 3 y 4 presentan un incremento importante mayor al 100% en la probabilidad, mientras que los grupos 5, 6, 7, 8 y 9 presentan un incremento menor al 100%, siendo los grupos donde el programa tiene un menor impacto.

Un aspecto interesante a revisar con respecto al caso homogéneo es que ahora los puntos por vencer y puntos vencidos sí presentan estimadores significativos para algunos grupos. Específicamente, los grupos 1 y 2 presentan un valor negativo para los puntos vencidos, lo que indica que a mayor cantidad de puntos vencidos en un mes, menor es la probabilidad de compra de los clientes de estos grupos. Esto se puede interpretar como que a los clientes les molesta el perder puntos, con lo que pierden interés en el programa. Por otro lado, el grupo 4 presenta un valor significativo para el parámetro asociado a los puntos por vencer y este es positivo. Esto se refleja en un aumento en la probabilidad de compra, lo que se puede interpretar como que, cuando tienen puntos por vencer en un mes, los clientes se ven motivados a ir comprar para tratar de alcanzar la meta antes de que les venzan los puntos. Finalmente, para los grupos 3, 8 y 9, tanto el parámetro de los puntos vencidos como el de los puntos por vencer son significativos, en el mismo sentido descrito anteriormente (el parámetro asociado a los puntos vencidos es negativo y a los puntos por vencer es positivo).

Por último, es importante destacar que el parámetro asociado al share de compra es significativo para todos los grupos, lo que muestra nuevamente que parte de la tendencia de compra por estacionalidad se captura mediante este parámetro.

Nivel 2

En la tabla n°24, se encuentra la evaluación de los modelos con distinta cantidad de grupos, mostrando para cada uno de ellos el logaritmo de la verosimilitud, así como el BIC y la diferencia entre el BIC para un modelo y el anterior.

Grupos	LL	BIC	Diferencia
1	-33.376	66.833	
2	-30.302	60.765	6.068,2
3	-29.814	59.870	894,6
4	-29.514	59.351	519,0
5	-29.391	59.187	164,0
6	-29.290	59.067	120,0
7	-29.264	59.096	-29,0

Tabla nº 24: Evaluación del modelo con distinto número de clases latentes
Fuente: Elaboración propia

El modelo con 6 clases latentes es el que presenta un mejor BIC. Los parámetros estimados para este modelo se muestran en la tabla nº 25.

	β_1 (Dt)	β_2 (SHAREt)	β_3 (VENCIDOST)	β_4 (X_VENCERT)	γ	α	π
Grupo 1	-2,296 (0,171)	6,770 (0,496)	-0,122 (0,05)	0,172 (0,046)	0,025 (0,412)	0,926 (0,075)	0,160 (0,099)
Grupo 2	-13,860 (1,112)	10,713 (1,369)	-0,042* (0,156)	0,224 (0,131)	615,170 (1,416)	0,652 (0,045)	0,168 (0,110)
Grupo 3	-0,465 (0,062)	3,079 (0,360)	0,022* (0,043)	0,022* (0,041)	0,023 (0,790)	0,401 (0,036)	0,141 (0,137)
Grupo 4	-1,789 (0,117)	4,655 (0,921)	0,025* (0,09)	-0,084* (0,084)	0,016 (0,382)	0,922 (0,022)	0,032 (0,173)
Grupo 5	-1,414 (0,063)	3,956 (0,369)	-0,092 (0,033)	-0,119 (0,031)	0,054 (0,760)	0,832 (0,875)	0,140 (0,086)
Grupo 6	-0,270 (0,074)	1,162 (0,380)	0,060* (0,058)	-0,090* (0,057)	5,516 (0,034)	0,457 (0,131)	0,356 (-)**

* No son estadísticamente significativos, según test t de dos colas, $P > 0.1$

** Este es el nivel base de la probabilidad, por lo que la desviación estándar es nula.

Tabla nº 25: Parámetros estimados para el nivel 2, con 6 clases latentes

Fuente: Elaboración propia

Por último, en la tabla nº26 se presenta la probabilidad de compra de los clientes de cada grupo, evaluado en la distancia al canje inicial y final, y el incremento porcentual de dicha probabilidad.

	$P(D_t=1)$	$P(D_t=0)$	Incremento %
Grupo 1	0,663	1,000	50,84%
Grupo 2	0,455	1,000	119,93%
Grupo 3	0,856	0,954	11,48%
Grupo 4	0,177	0,689	288,69%
Grupo 5	0,445	0,911	104,85%
Grupo 6	0,988	0,997	0,90%

Tabla nº 26: Probabilidad de compra para cada grupo, al comienzo y al final de la carrera
Fuente: Elaboración propia

Análisis de resultados nivel 2

La distancia al canje normalizada es, nuevamente, significativa para todos los grupos, lo que confirma la presencia de la aceleración de compra para el nivel 2 a medida que los clientes se acercan al premio.

Para este nivel, se puede notar que el grupo 6 predomina en tamaño, teniendo un 35,6% de probabilidad de que un cliente pertenezca a él. Este grupo tiene una característica especial, ya que el impacto del programa de lealtad es muy pequeño en ellos (el aumento porcentual de la probabilidad de compra es de 0.90%), sin embargo tanto al comenzar como al terminar la carrera, la probabilidad de compra es muy alta. Esto indica que los clientes de este grupo están muy habituados a comprar, por lo que el programa de lealtad no sirve realmente de incentivo, pues ellos de todas formas irán a comprar casi mensualmente.

Por otro lado, el grupo 4 es considerablemente más pequeño que el resto, con apenas un 3,2% de los clientes del nivel 2. Este grupo presenta la particularidad de que la probabilidad de compra de los clientes en la meta (al llegar al premio, distancia igual a cero) es mucho más baja que para el resto, lo que indicaría que este es un grupo de clientes reacios a comprar. De todas formas, el programa tiene un gran impacto en ellos, de hecho es el grupo donde el aumento en la probabilidad de compra es mayor.

El resto de los grupos es de tamaño relativamente similar. Los grupos 1 y 3 presentan un aumento relativamente pequeño en la probabilidad, mientras que los grupos 2 y 5 presentan un aumento mayor al 100%.

Al igual que para el nivel 1, los puntos por vencer y puntos vencidos presentan estimadores significativos para algunos grupos. El grupo 2 presenta un beta positivo asociado a los puntos por vencer, mientras que los grupos 1 y 5 presentan parámetros significativos tanto para los puntos vencidos como para los puntos por vencer, y ambos en el mismo sentido que se ha explicado con anterioridad (positivo para los puntos por vencer y negativo para los puntos vencidos).

Por último, es importante destacar que el parámetro asociado al share de compra es significativo para todos los grupos, lo que muestra nuevamente que parte de la tendencia de compra por estacionalidad se captura mediante este parámetro.

Nivel 3

En la tabla n°27, se encuentra la evaluación de los modelos con distinta cantidad de grupos, mostrando para cada uno de ellos el logaritmo de la verosimilitud, así como el BIC y la diferencia entre el BIC para un modelo y el anterior.

Grupos	LL	BIC	Diferencia
1	-20.949	41.977	-
2	-19.078	38.313	3.663,5
3	-18.798	37.833	479,9
4	-18.622	37.560	273,1
5	-18.523	37.442	118,7
6	-18.486	37.447	-4,9

Tabla n° 27: Evaluación del modelo con distinto número de clases latentes
Fuente: Elaboración propia

El modelo con 5 clases latentes es el que presenta un mejor BIC. Los parámetros estimados para este modelo se muestran en la tabla n° 28.

	β_1 (Dt)	β_2 (SHAREt)	β_3 (VENCIDOST)	β_4 (X_VENCERT)	γ	α	π
Grupo 1	-1,388 (0,078)	2,859 (0,450)	0,0076* (0,048)	-0,038* (0,044)	0,080 (0,562)	0,716 (0,031)	0,093 (0,261)
Grupo 2	-1,725 (0,188)	5,514 (0,586)	-0,103 (0,052)	0,114 (0,046)	0,023 (0,535)	0,840 (0,043)	0,146 (0,381)
Grupo 3	-7,428 (0,681)	8,004 (0,720)	0,1829* (0,118)	0,220 (0,091)	2,040 (0,630)	0,560 (0,059)	0,259 (0,272)
Grupo 4	-0,176 (0,102)	1,605 (0,532)	0,0213* (0,079)	-0,104* (0,073)	1,889 (0,758)	0,639 (0,167)	0,370 (0,262)
Grupo 5	-0,515 (0,106)	2,154 (0,797)	-0,136 (0,055)	0,0671* (0,063)	0,211 (1,457)	0,431 (0,114)	0,130 (-)**

* No son estadísticamente significativos, según test t de dos colas, $P > 0.1$

** Este es el nivel base de la probabilidad, por lo que la desviación estándar es nula.

Tabla n° 28: Parámetros estimados para el nivel 3, con 5 clases latentes
Fuente: Elaboración propia

Por último, en la tabla n°29 se presenta la probabilidad de compra de los clientes de cada grupo, evaluado en la distancia al canje inicial y final, y el incremento porcentual de dicha probabilidad.

	$P(D_t=1)$	$P(D_t=0)$	Incremento %
Grupo 1	0,391	0,863	120,60%
Grupo 2	0,609	0,995	63,29%
Grupo 3	0,657	1,000	52,24%
Grupo 4	0,993	0,997	0,42%
Grupo 5	0,863	0,964	11,70%

Tabla nº 29: Probabilidad de compra para cada grupo, al comienzo y al final de la carrera
Fuente: Elaboración propia

Análisis de resultados nivel 3

Nuevamente, la distancia al canje normalizada es significativa para todos los grupos, lo que confirma la presencia de la aceleración de compra para el nivel 3 a medida que los clientes se acercan al premio, consistente con la hipótesis 1.

Para este nivel, se puede notar que el grupo 4 predomina en tamaño, teniendo un 37,04% de probabilidad de que un cliente pertenezca a él. Al igual que en el nivel 2, este grupo de mayor tamaño agrupa a los clientes que presentan una probabilidad de compra muy alta tanto al comenzar la carrera hacia el premio como al alcanzarlo, lo que indica que están muy habituados a comprar. El programa de lealtad para estos clientes no sirve realmente de gran incentivo pues son clientes que de todas formas irán mensualmente a comprar. El resto de los grupos es de tamaño relativamente similar. El grupo 1, que es el más pequeño, agrupa a clientes donde el impacto del programa es más alto (120,60%), mientras que en los grupos 2,3 y 5 el impacto es relativamente pequeño.

Al igual que en los niveles 1 y 2, los puntos por vencer y puntos vencidos presentan estimadores significativos para algunos grupos. El grupo 5 presenta un parámetro negativo asociado a los puntos vencidos, el grupo 3 presenta un parámetro positivo para los puntos por vencer y el grupo 2 presentan ambos parámetros significativos, con signo negativo para los vencidos y positivo para los por vencer. Todo esto es consistente con la intuición y los resultados que se han visto hasta ahora, en cuanto a la interpretación de estos valores.

Por último, es importante destacar que el parámetro asociado al share de compra es significativo para todos los grupos, lo que muestra nuevamente que parte de la tendencia de compra por estacionalidad se captura mediante este parámetro.

Nivel 4

En la tabla nº30, se encuentra la evaluación de los modelos con distinta cantidad de grupos, mostrando para cada uno de ellos el logaritmo de la verosimilitud, así como el BIC y la diferencia entre el BIC para un modelo y el anterior.

Grupos	LL	BIC	Diferencia
1	-2.016	4.096	-
2	-1.848	3.824	271,9
3	-1.809	3.813	11,6
4	-1.782	3.824	-10,9

Tabla nº 30: Evaluación del modelo con distinto número de clases latentes
Fuente: Elaboración propia

El modelo con 3 clases latentes es el que presenta un mejor BIC. Los parámetros estimados para este modelo se muestran en la tabla nº 31.

	β_1 (Dt)	β_2 (SHAREt)	β_3 (VENCIDOST)	β_4 (X_VENCERT)	γ	α	π
Grupo 1	-2,553 (0,150)	5,560 (0,811)	0,024* (0,084)	-0,001* (0,08)	1,110 (1,049)	1,711 (0,060)	0,610 (0,162)
Grupo 2	-1,265 (0,178)	1,879 (0,926)	-0,001* (0,151)	0,005* (0,140)	1,237 (1,526)	1,834 (0,069)	0,062 (0,274)
Grupo 3	-1,350 (0,734)	5,133 (0,941)	-0,013* (0,393)	-0,007* (0,382)	1,010 (0,022)	1,926 (0,434)	0,326 (-)**

* No son estadísticamente significativos, según test t de dos colas, $P > 0.1$

** Este es el nivel base de la probabilidad, por lo que la desviación estándar es nula.

Tabla nº 31: Parámetros estimados para el nivel 4, con 3 clases latentes
Fuente: Elaboración propia

Por último, en la tabla nº32 se presenta la probabilidad de compra promedio de los clientes de cada grupo, evaluado en la distancia al canje inicial y final, y el incremento porcentual de dicha probabilidad.

	$P(D_t=1)$	$P(D_t=0)$	Incremento %
Grupo 1	0,977	1,000	2,38%
Grupo 2	0,460	0,888	92,87%
Grupo 3	0,751	0,995	32,61%

Tabla nº 32: Probabilidad de compra para cada grupo, al comienzo y al final de la carrera
Fuente: Elaboración propia

Análisis de resultados nivel 4

Nuevamente, la distancia al canje normalizada es significativa para todos los grupos, lo que confirma la presencia de la aceleración de compra para el nivel 4 a medida que los clientes se acercan al premio.

En este nivel se tienen sólo 3 grupos. El grupo 1 es el más grande, agrupando a más de la mitad de los individuos. Este grupo es, nuevamente, el grupo donde los clientes presentan una gran probabilidad de compra tanto cuando están al comienzo como al final de la carrera, al igual como ha ocurrido en los otros niveles, donde el

grupo más grande agrupa a este tipo de clientes. El grupo 2 es el más pequeño, agrupando a sólo el 6,22% de los individuos. Este grupo es el que presenta mayor impacto por el programa de lealtad. Por último, el grupo 3 tiene también un tamaño importante, pero el impacto del programa en ellos es relativamente pequeño.

Se debe notar que en este nivel es el único donde ni los puntos vencidos ni por vencer juegan ningún rol: Ambos son no significativos para los 3 grupos. Por último, es importante destacar que el parámetro asociado al share de compra es significativo para todos los grupos, lo que muestra nuevamente que parte de la tendencia de compra por estacionalidad se captura mediante este parámetro.

8.3 Hipótesis 2 – Esfuerzo de compra pre-canje

H₂: Los clientes aumentan el esfuerzo invertido en el programa de lealtad, medido en dinero gastado en las compras, a medida que se acercan al nivel de canje de un premio.

El estudio anterior ha examinado cómo la distancia al canje afecta los tiempos entre compras sucesivas. En muchos casos, sin embargo, los clientes pueden acelerar sus esfuerzos incrementando la cantidad de puntos acumulados en cada interacción con el programa. En otras palabras, así como un cliente puede disminuir los tiempos entre compras sucesivas, es posible que, además, la cantidad de dinero invertido en cada una de ellas vaya aumentando a medida que se acerca al premio.

8.3.1 Análisis preliminar y motivación

A modo de motivación, en la figura nº17 se muestran los montos de compra promedio versus la distancia al canje.

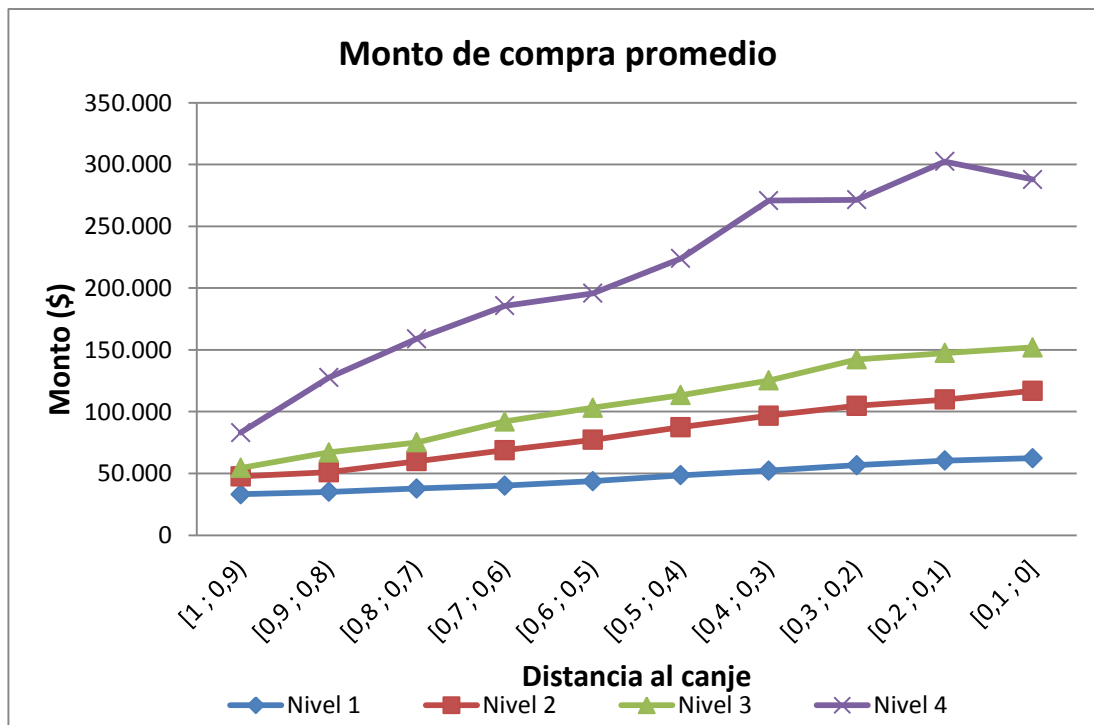


Figura n° 17: Monto de compra promedio v/s distancia al canje
Fuente: Elaboración propia

Se observa una clara tendencia en todos los niveles, donde el cliente incrementa el monto promedio de compra a medida que se va acercando al premio. Dado lo anterior, se procede a realizar un estudio que permita determinar si la aceleración es efectiva.

8.3.2 Modelos

Para modelar la cantidad de esfuerzo invertido en cada una de las compras, se usará por comodidad los puntos recibidos en cada oportunidad, lo que es equivalente a la cantidad de dinero invertido multiplicado por un factor. Para esto, se aplicará un modelo lineal generalizado de distribución Normal²³ y Poisson²⁴. El uso de modelos lineales generalizados es directo para aplicaciones como esta. Se usan porque se desea estudiar la tendencia de los montos a medida que el cliente se acerca al canje de un premio. Al usar una distribución normal, se recupera la regresión lineal típica, mientras que la distribución de Poisson presenta características de interés para el estudio, al ser una función de distribución de probabilidad discreta.

En la aplicación de los modelos señalados con anterioridad, se utilizarán las cuatro covariables que han demostrado ser más relevantes en los estudios anteriores: distancia al canje normalizada, share, puntos vencidos y por vencer, y además se

²³ Ver capítulo 5.3.1

²⁴ Ver capítulo 5.3.2

incluye una constante β_0 . La variable dependiente serán los puntos acumulados en las compras. Estos modelos permiten generalizar el modelo del gradiente-objetivo al dominio de las cantidades de esfuerzo invertido.

Para la estimación se utilizará la técnica de validación simple, dividiendo cada base de datos en 2 sub-muestras. La primera considera el 90% de los clientes y será la muestra utilizada para estimar los parámetros de los modelos. La segunda, con el resto de los clientes, será la muestra que se utilizará para validar el modelo y medir sus capacidades predictivas.

8.3.3 Resultados

En las tablas n°33 y n°34 se muestra el MAE y MAPE²⁵ de las etapas de estimación y validación, respectivamente, para cada uno de los 2 modelos en los 4 niveles de canje

	Normal		Poisson	
	MAE	MAPE	MAE	MAPE
Nivel 1	411,36	11,58%	411,54	11,59%
Nivel 2	728,95	13,38%	729,71	13,39%
Nivel 3	941,69	13,56%	943,10	13,59%
Nivel 4	1.792,10	14,45%	1.793,70	14,52%

Tabla n° 33: MAE y MAPE en la etapa de estimación.

Fuente: Elaboración propia

	Normal		Poisson	
	MAE	MAPE	MAE	MAPE
Nivel 1	409,37	12,83%	409,73	12,85%
Nivel 2	710,30	16,09%	711,48	16,12%
Nivel 3	971,55	13,91%	974,45	14,11%
Nivel 4	1.522,30	9,50%	1.528,00	9,51%

Tabla n° 34: MAE y MAPE en la etapa de validación.

Fuente: Elaboración propia

Se puede observar de las dos tablas anteriores que el modelo lineal generalizado con distribución normal es el que presenta mejores MAE y MAPE, tanto en la etapa de estimación como en la de validación. De todas formas, el modelo de Poisson presenta

²⁵ El MAE (mean absolute error o error medio absoluto) y MAPE (mean absolute percentage error o error medio porcentual absoluto) son medidas de error que miden la precisión general de los modelos para

asegurar su buen desempeño. $MAE = \frac{1}{N} \sum_{i=1}^N |y_i - f_i|$ y $MAPE = \frac{1}{N} \sum_{i=1}^N \left| \frac{y_i - f_i}{y_i} \right|$ donde y_i es el valor real y

f_i el valor predicho de la variable independiente.

errores muy similares a la Normal. En la tabla nº 35, se presentan los resultados de la estimación según la regresión Normal.

	β_1 (Dt)	β_2 (SHAREt)	β_3 (VENCIDost)	β_4 (X_VENCert)	β_0
Nivel 1	-219,6 (4,656)	716,3 (17,99)	-70,1 (3,493)	108,6 (3,393)	59,6 (15,18)
Nivel 2	-449,6 (15,430)	1.589,2 (79,54)	-170,4 (13,32)	89,1 (12,44)	-211,1 (70,36)
Nivel 3	-612,6 (111,1)	2.335,1 (22,57)	-266,5 (122,1)	61,0 (20,04)	-542,0 (18,20)
Nivel 4	-1.413,4 (112,8)	4651,5 (794,5)	-480,7 (11,35)	122,5* (95,58)	-1.208,9 (74,87)

*No son significativos, $P > 0.1$ test t de dos colas.

Tabla nº 35: Estimación de parámetros para la normal

Fuente: Elaboración propia

8.3.4 Discusión

Consistente con la hipótesis 2, se puede observar que el parámetro asociado a la distancia al canje es negativo para todos los niveles, lo que indica que los clientes reciben más puntos por visita (gastan más dinero) a medida que se acercan a la meta. Se puede observar también un efecto incremental del impacto de esta covariable a través de los niveles, pero esto podría explicarse simplemente por el hecho de que en niveles superiores se requieren más puntos y, por ende, la cantidad de dinero gastado y puntos recibidos tiene que ser mayor. Por otro lado, el parámetro asociado al share es positivo, lo que indica que este parámetro está capturando efectos de estacionalidad.

Con respecto a los parámetros asociados a puntos por vencer y puntos vencidos, estos son significativos en la mayoría de los casos. Los puntos vencidos tienen signo negativo y su valor es significativo para todos los niveles, lo que indica que cuando los puntos vencen, los clientes gastan menos dinero en el programa, lo que podría explicarse por una desmotivación o enojo de su parte. Además, este valor es decreciente (más negativo) a través de los grupos, lo que indicaría que el impacto es mayor en niveles más altos, donde los clientes necesitan más puntos para canjear. Por último, los puntos por vencer son sólo significativos en los primeros tres niveles, y presentan en estos casos signo positivo. Esto indica que al tener puntos por vencer, se sienten motivados a gastar más dinero, presuntamente para acumular la cantidad de puntos necesaria para canjear antes de que se les venzan.

8.4 Hipótesis 3 – Comportamiento de compra post-canje

H_3 : *Los clientes del programa de lealtad reinician su patrón de compras después del canje de un premio.*

De acuerdo con la hipótesis de aceleración de compra H1, la motivación para esforzarse aumenta a medida que el cliente avanza hacia el canje de un premio. Una pregunta natural que surge de este efecto es si, después de canjear, el cliente mantiene su alta frecuencia de compras permanentemente, o bien exhibe una disminución en su frecuencia de compra, volviendo a un estado inicial similar al del inicio del canje anterior.

Si los consumidores disminuyeran su frecuencia de compra, se esperaría que volvieran a comenzar un nuevo patrón de compras con aceleración, como plantea la hipótesis 1, para alcanzar un nuevo premio. A este efecto se le denomina “reinicio post-premio”²⁶.

8.4.1 Datos

Para testear esta hipótesis, se utiliza el grupo de datos del Nivel 1, dividido en 4 subgrupos. Para dividirlos, se han considerado solo aquellos clientes cuyos primeros 4 canjes sucesivos fueron en este nivel. Se eligieron los clientes de esta forma pues se necesita tener canjes consecutivos para revisar las diferencias en la frecuencia de compras al comienzo y al final de una carrera hacia un premio. Además, como el nivel 1 presenta la mayor cantidad de canjes, se logra encontrar una base de clientes mayor para el estudio que eligiendo otros niveles. Es necesario considerar canjes del mismo nivel para evitar sesgos por cambios en las probabilidades de compra simplemente porque un canje fue en un nivel y el siguiente en otro.

Una vez seleccionados los clientes, se agrupan los datos en 4 bases de datos distintas, una para cada canje de los clientes (primer, segundo, tercer y cuarto canje). Los canjes posteriores al cuarto, de existir, no se consideran en esta parte del estudio.

En total, existen 980 clientes que cumplen esta condición, lo que da un total de 3.920 canjes.

8.4.2 Análisis preliminar y motivación

Primero, se realizará un análisis previo de los datos. En la figura nº18 se muestra el tiempo entre compras promedio (en meses) en función de la distancia al canje (d_i), para cada uno de los 4 grupos mencionados anteriormente.

²⁶ Ver capítulo 4.1 para explicación del efecto de reinicio post premio

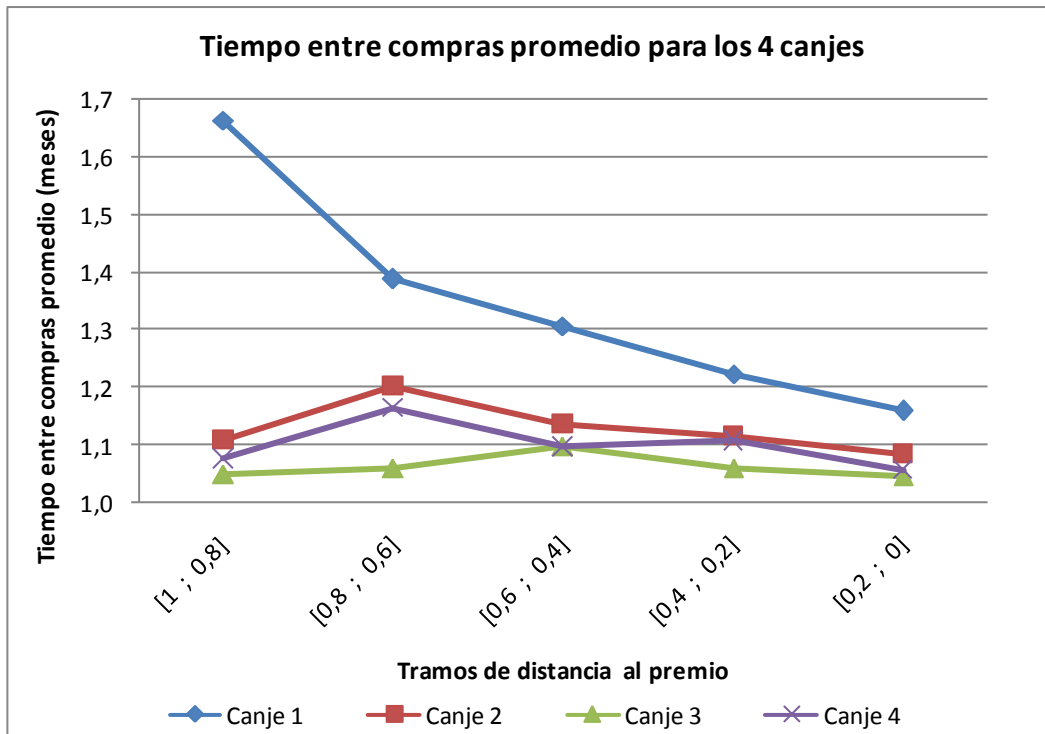


Figura n° 18: Tiempo entre compras promedio en función del tramo de la distancia al canje
Fuente: Elaboración propia.

De aquí se puede observar que, para el primer canje, los tiempos entre compras decrecen entre un tramo y el siguiente, a medida que el cliente se va acercando al premio. Sin embargo, para el resto de los canjes, las compras se realizan casi mes a mes, independientemente de la distancia al premio. En particular, en los canjes 2 y 3, se observa que el tiempo entre compras promedio aumenta y luego decrece a medida que pasan a los tramos de distancia siguiente. El canje 4, en cambio, presenta un patrón irregular. Esto estaría indicando que el primer canje aumenta efectivamente la motivación para comprar, en efecto en el primer canje se observa que los tiempos entre compras decrecen a medida que el cliente se acerca al canje, y en los canjes siguientes los clientes compran prácticamente todos los meses en forma constante. Sin embargo, producto de esta motivación que genera el primer canje, no se ve claramente que exista a priori un reseteo post premio. Es necesario realizar el estudio correspondiente para ver si es así.

8.4.3 Modelo

Para estudiar el fenómeno se utiliza un modelo hazard proporcional en tiempo discreto. De acuerdo a los resultados de H1, se usan las covariables de distancia al canje d_t , el share de compra, puntos vencidos y por vencer, que mostraron ser las covariables más relevantes. Es importante destacar que no se agregó al modelo el efecto del número de canjes, porque este es constante en las bases de datos generadas; todos los clientes tienen el mismo número de canjes en cada una de ellas (1, 2, 3 o 4 canjes). La función de probabilidad queda como sigue:

$$\Pr(t, X_t) = 1 - \exp\left(-e^{(\beta_1 d_t + \beta_3 \text{SHARE}_t + \beta_4 \text{VENCIDOS}_t + \beta_5 \text{POR_VENCER}_t)} \int_{t-1}^t h(s) ds\right) \quad (49)$$

8.4.4 Resultados

Para estimar los parámetros, se probaron distintas parametrizaciones del modelo hazard proporcional en tiempo discreto (funciones exponencial, weibull y log-logística), además del modelo no paramétrico. El ajuste se midió en base al logaritmo natural de la verosimilitud (LL) y el criterio de información bayesiano (BIC).

En la tabla nº48, se muestra el ajuste de las distintas parametrizaciones.

	Exponencial		Weibull		Log-logística		No paramétrico	
	LL	BIC	LL	BIC	LL	BIC	LL	BIC
Canje 1	-6.145	12.318	-5.628	11.295	-5.629	11.297	-5.741	11.665
Canje 2	-2.395	11.295	-2.169	4.375	-2.171	4.378	-2.191	4.555
Canje 3	-1.331	2.689	-1.223	2.481	-1.227	2.490	-1.241	2.606
Canje 4	-1.565	3.157	-1.362	2.760	-1.379	2.793	-1.453	3.030

Tabla nº 36: Ajuste del modelo según las distintas parametrizaciones, para todos los canjes
Fuente: Elaboración propia.

La que mostró mejor ajuste fue la parametrización Weibull, para la cual el BIC es menor en todos los canjes comparado con el resto de las parametrizaciones. Los resultados con esta parametrización, para todos los niveles, se muestran en la tabla nº49²⁷.

	β_1 (dt)	β_2 (SHAREt)	β_3 (VENCIDOS _t)	β_4 (X_VENCERT)	γ	α	LL	BIC
Canje 1	-0,765 (0,043)	3,327 (0,598)	0,003* (0,037)	-0,031* (0,033)	1,064 (0,056)	0,571 (0,012)	-5.628	11.295
Canje 2	-0,450 (0,066)	3,647 (0,572)	0,069* (0,052)	0,110* (0,086)	1,017 (0,02)	0,515 (0,038)	-2.169	4.375
Canje 3	-0,144 (0,061)	4,028 (0,094)	0,005* (0,080)	0,062* (0,051)	1,006 (0,01)	0,498 (0,03)	-1.223	2.481
Canje 4	-0,210 (0,063)	4,972 (0,08)	-0,001* (0,081)	0,001* (0,054)	1,001 (0,031)	0,460 (0,02)	-1.362	2.760

* No son estadísticamente significativos, según test t de dos colas, P>0.1

Tabla nº 37: Parámetros estimados según modelo Hazard, parametrización weibull
Fuente: Elaboración propia

²⁷ Para ver las estimaciones de los parámetros del modelo, con todas las parametrizaciones de la función hazard (y de la no paramétrica), ver anexo F

En la figura nº19, se muestra la probabilidad de compra promedio por tramos, en función del share de compra, para cada uno de los grupos de interés.

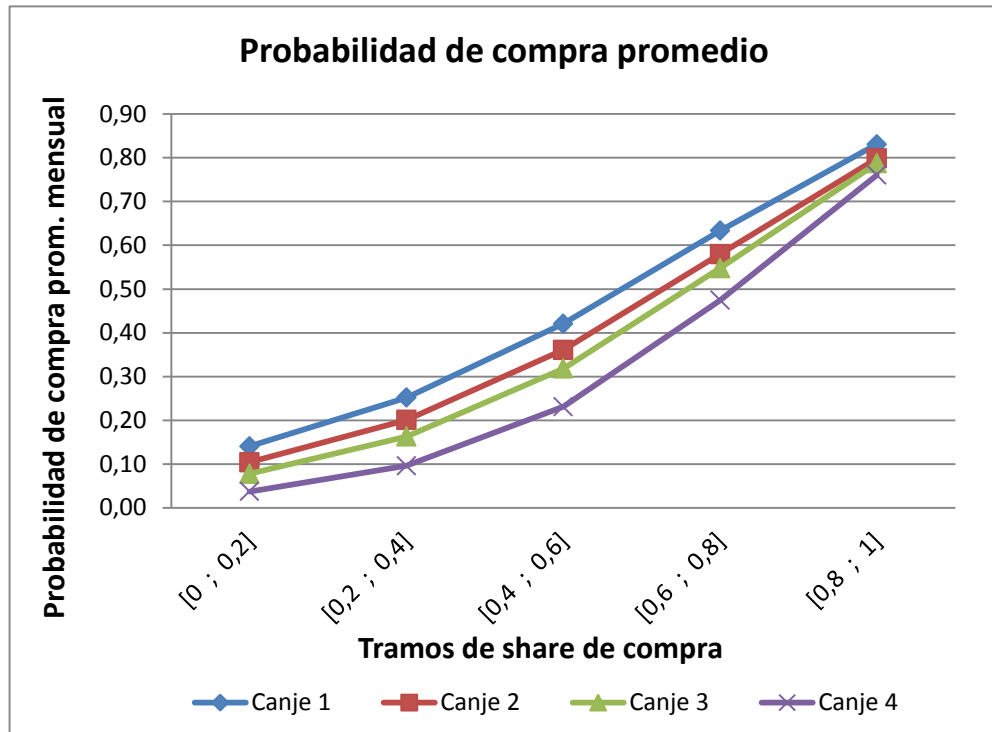


Figura nº 19: Probabilidad de compra promedio en función del tramo del share de compra
Fuente: Elaboración propia.

Como era de esperarse, a medida que el share crece, la probabilidad de compra aumenta, pues el que el share de compra sea más grande implica que la cantidad de compras concretadas en promedio es mayor. De aquí se observa también que el share captura buena parte del comportamiento de compra del cliente, pues variaciones grandes en este producen importantes cambios en la probabilidad. Si, por ejemplo, el share es muy bajo (pocos están comprando), la probabilidad es muy baja, independiente de la distancia al canje.

De forma similar, en la figura nº20, se muestra la probabilidad de compra promedio por tramos, en función de la distancia hasta el próximo canje, para cada uno de los grupos de interés.

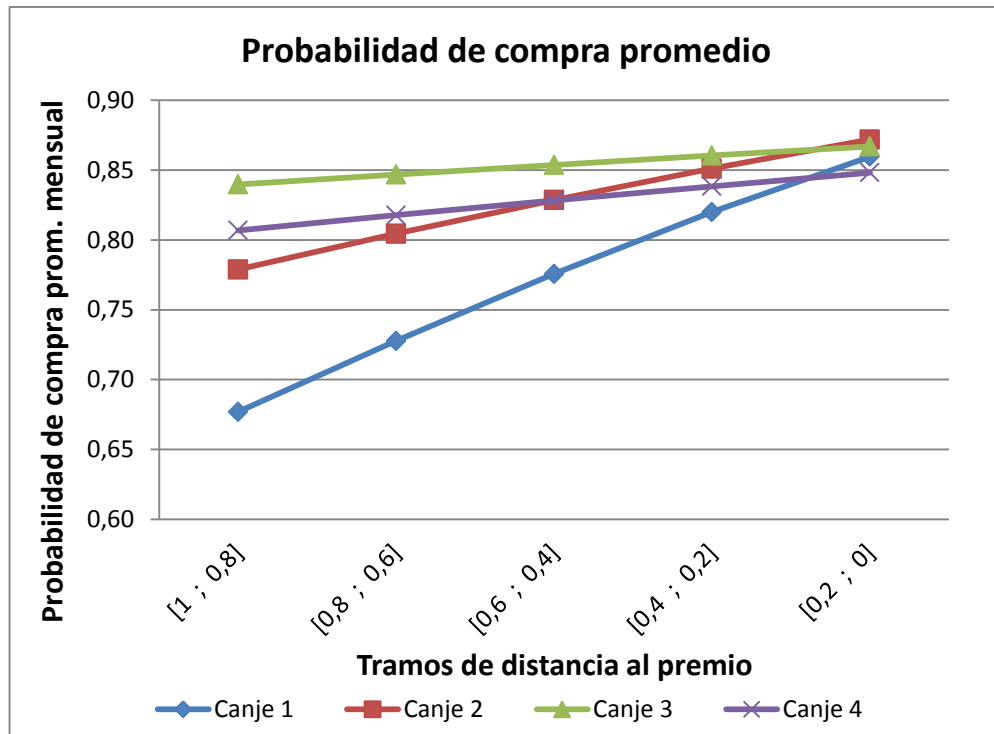


Figura nº 20: Probabilidad de compra promedio en función del tramo de la distancia al canje
Fuente: Elaboración propia.

Se puede observar que, consistente con H1, la probabilidad de compra de los clientes crece a medida que la distancia al premio disminuye, para los 4 canjes. Sin embargo, se observa que los clientes no regresan a un estado análogo al del canje anterior después de un canje, sino más bien su probabilidad de compra es mayor al estado anterior, lo que se observa en los 3 primeros canjes. El cuarto canje, en cambio, registra una disminución en la probabilidad con respecto al tercero. Esto indicaría que para el cuarto canje el programa de lealtad ya no es tan efectivo como en los anteriores.

A continuación, se calcularon los intervalos de confianza de los parámetros β_1^{28} , a partir de los cuales se obtuvieron los intervalos de confianza para la probabilidad promedio de compra al inicio y al final de cada canje, es decir, al estar lo más lejos y más cerca del premio, respectivamente. Estos intervalos se muestran en las tablas nº50 y nº51.

²⁸ Ver anexo G para los intervalos de confianza de β_1

	Probabilidades	
Canje 1	0,651	0,703
Canje 2	0,744	0,811
Canje 3	0,813	0,865
Canje 4	0,777	0,835

Tabla nº 38: Intervalos de confianza para la probabilidad de compra promedio, al inicio del canje
Fuente: Elaboración propia.

	Probabilidades	
Canje 1	0,858	0,862
Canje 2	0,869	0,875
Canje 3	0,864	0,870
Canje 4	0,845	0,851

Tabla nº 39: Intervalos de confianza para la probabilidad de compra promedio, al final del canje
Fuente: Elaboración propia.

Se observa que los intervalos de confianza al inicio de los canjes no se intersectan entre sí, a excepción del canje 4. Esto implica que existe diferencia significativa en las probabilidades al comienzo de las primeras 3 carreras, partiendo en cada canje desde un punto con mayor probabilidad que para el anterior. En el caso del canje 4, no hay diferencia significativa en la probabilidad al comienzo de la carrera con la probabilidad al comienzo del tercer canje.

Por otro lado, comparando el intervalos de confianza al final de un canje con el intervalo de confianza al inicio del canje siguiente, se observa que en todos los casos estos intervalos no se intersectan, por lo que existe una diferencia significativa entre las probabilidades de compra al final de un canje y al inicio del canje siguiente, lo que demuestra la existencia de un reinicio post-premio.

8.4.5 Discusión

De lo anterior, se puede concluir que existe un efecto de reinicio post premio parcial ya que, después de un canje, los clientes vuelven a un estado donde su probabilidad de compra inicial es menor que al final del canje anterior, pero al mismo tiempo mayor que al comienzo de del canje anterior. Este efecto se puede observar al menos para los primeros 3 canjes; en el caso del cuarto canje existe un reinicio post-premio total, ya no se parte con una probabilidad de compra mayor que en el canje anterior. Se debe notar que un reinicio post canje parcial es mejor que un reinicio completo, porque la probabilidad de compra no disminuye hasta el punto inicial del canje anterior sino a un punto superior, lo que indica que a medida que el cliente canjea productos, su probabilidad mínima de compra aumenta.

Por otro lado, el hecho de que la probabilidad al comienzo de un canje sea menor que al final del canje anterior es un resultado sumamente importante, pues confirma que es el premio el que motiva al cliente a acelerar su comportamiento de compra y que

esta aceleración no se debe simplemente a un efecto de inercia o aprendizaje del cliente. Dicho de otra forma, la caída en la probabilidad de compra al iniciar una nueva carrera hacia un canje valida que el cliente está acelerando según su distancia a un premio (hipótesis 1) y no simplemente porque una vez que comienza a comprar se acostumbra a hacerlo cada vez más seguido.

8.4.6 Heterogeneidad

El igual que con la hipótesis 1, se debe considerar el sesgo por heterogeneidad. Para incorporar el efecto, se hará un análisis de clases latentes.

8.4.6.1 Modelo

Para efectos de este análisis, se considerará que todos los parámetros del modelo pueden variar a través de los individuos. De esta forma, al considerar las clases latentes, la función de probabilidad del hazard proporcional en tiempo discreto queda dada por la expresión que se muestra a continuación:

$$P_{ig}(t, X_t) = 1 - \exp\left(-e^{(\beta_{1g}d_t + \beta_{2g}SHARE_t + \beta_{3g}VENCIDOS_t + \beta_{4g}POR_VENCER_t)} \int_{t-1}^t h_g(s) ds\right) \quad (50)$$

En esta expresión, cada uno de los parámetros a estimar varía con respecto a un índice “g”, que representa los grupos o clases a considerar. Luego, la función de verosimilitud queda dada por

$$L = \prod_{i=1}^N \left(\sum_{g=1}^G \left(\prod_{t=1}^{T_i} P_{ig}(t, X_{it})^{\delta_{it}} * (1 - P_{ig}(t, X_{it}))^{1-\delta_{it}} * \pi_g \right) \right) \quad (51)$$

Es importante recordar que la probabilidad π_g es también un parámetro a estimar en este modelo.

8.4.6.2 Resultados

Para estimar los parámetros, se maximiza la verosimilitud indicada anteriormente, usando la parametrización Weibull. Como primer paso, se estima el modelo con una sola clase latente, con lo que se recuperan los resultados del caso homogéneo. Después, se evalúan los modelos incrementando el número de clases latentes, de uno en uno. Finalmente, se decide cuál es el número de clases latentes que mejor describe a la muestra, usando el criterio de información bayesiana.

Canje 1

En la tabla nº36, se encuentra la evaluación de los modelos con distinta cantidad de grupos, mostrando para cada uno de ellos el logaritmo de la verosimilitud, así como el estadístico del criterio de información bayesiana (BIC) y la diferencia entre el BIC para un modelo y el anterior.

Grupos	LL	BIC	Diferencia
1	-5.629	11.325	-
2	-5.204	10.543	782,5
3	-5.126	10.453	89,6
4	-5.101	10.470	-16,8

Tabla nº 40: Evaluación del modelo con distinto número de clases latentes
Fuente: Elaboración propia

El modelo con 3 clases latentes es el que presenta un mejor BIC. Los parámetros estimados para este modelo se muestran en la tabla nº 37.

	β_1 (Dt)	β_2 (SHAREt)	β_3 (VENCIDOST)	β_4 (X_VENCERT)	γ	α	π
Grupo 1	-1,299 (0,501)	5,9848 (0,587)	0,076* (0,102)	-0,162* (0,163)	0,006 (0,631)	0,883 (0,019)	0,145 (0,1)
Grupo 2	-0,732 (0,141)	4,333 (0,116)	-0,115* (0,338)	-0,195* (0,432)	0,017 (0,493)	0,759 (0,064)	0,426 (0,163)
Grupo 3	-4,371 (1,383)	6,2993 (0,338)	0,4737 (0,236)	-0,416* (0,911)	0,654 (0,549)	0,559 (0,033)	0,428 (-)**

* No son estadísticamente significativos, según test t de dos colas, $P > 0.1$

** Este es el nivel base de la probabilidad, por lo que la desviación estándar es nula.

Tabla nº 41: Parámetros estimados para el canje 1, con 3 clases latentes

Fuente: Elaboración propia

Por último, en la tabla nº38 se presenta la probabilidad de compra de los clientes de cada grupo, evaluado en la distancia al canje inicial y final (1 y 0 respectivamente)²⁹. Con ello, se puede calcular el incremento porcentual en la probabilidad de compra, lo que permite observar a qué grupos afecta en mayor o menor medida el programa de lealtad.

	$P(d_t=1)$	$P(d_t=0)$	Incremento %
Grupo 1	0,346	0,789	128,07%
Grupo 2	0,527	0,790	49,70%
Grupo 3	0,840	1,000	19,02%

Tabla nº 42: Probabilidad de compra para cada grupo, al comienzo y al final de la carrera

Fuente: Elaboración propia

²⁹ Para el resto de los parámetros se usan valores promedio

Análisis de resultados canje 1

Se puede notar que la distancia al canje normalizada es significativa para todos los grupos, lo que confirma la presencia de la aceleración de compra para el primer canje, a medida que los clientes se acercan al premio, consistente con H1.

De los tres grupos, el primero es el más pequeño, el cuál se ve impactado en mayor medida por el programa de lealtad, aumentando su probabilidad de compra en un 128,07%. Los otros dos grupos son relativamente similares en tamaño, siendo el último el que menos cambia su comportamiento, debido principalmente a que son clientes que ya compraban muy seguido, incluso sin el programa de lealtad.

Los puntos por vencer y puntos vencidos no juegan un rol importante en ninguno de los grupos en el primer canje. El único caso donde los puntos vencidos son significativos es en el tercer grupo, y aun así este es el grupo donde el programa no tiene un gran efecto, por lo que no parece ser un aspecto relevante..

Por último, es importante destacar que el parámetro asociado al share de compra es significativo para todos los grupos, lo que muestra que parte de la tendencia de compra por estacionalidad se captura mediante este parámetro.

Canje 2

En la tabla nº39, se encuentra la evaluación de los modelos con distinta cantidad de grupos, mostrando para cada uno de el logaritmo de la verosimilitud, así como el BIC y la diferencia entre el BIC para un modelo y el anterior.

Grupos	LL	BIC	Diferencia
1	-2.169	4.402	
2	-2.022	4.172	230,2
3	-2.010	4.211	-39,8

Tabla nº 43: Evaluación del modelo con distinto número de clases latentes
Fuente: Elaboración propia

El modelo con 2 clases latentes es el que presenta un mejor BIC. Los parámetros estimados para este modelo se muestran en la tabla nº 40.

	β_1 (Dt)	β_2 (SHAREt)	β_3 (VENCIDOST)	β_4 (X_VENCERT)	γ	α	π
Grupo 1	-1,313 (0,756)	12,576 (1,239)	-0,025* (0,054)	-0,122* (0,097)	2,796 (0,236)	0,855 (0,036)	0,181 (0,087)
Grupo 2	-0,314 (0,131)	3,761 (0,983)	0,026* (0,33)	0,033* (0,364)	0,002 (0,116)	0,411 (0,049)	0,818 (-)**

* No son estadísticamente significativos, según test t de dos colas, $P > 0.1$

** Este es el nivel base de la probabilidad, por lo que la desviación estándar es nula.

Tabla nº 44: Parámetros estimados para el canje 2, con 2 clases latentes

Fuente: Elaboración propia

Por último, en la tabla nº41 se presenta la probabilidad de compra de los clientes de cada grupo, evaluado en la distancia al canje inicial y final, y el incremento porcentual de dicha probabilidad.

	$P(d_t=1)$	$P(d_t=0)$	Incremento %
Grupo 1	0,382	0,832	118,17%
Grupo 2	0,848	0,924	19,01%

Tabla nº 45: Probabilidad de compra para cada grupo, al comienzo y al final de la carrera

Fuente: Elaboración propia

Análisis de resultados canje 2

La distancia al canje normalizada es, nuevamente, significativa para todos los grupos, lo que confirma la presencia de la aceleración de compra para el canje 2 a medida que los clientes se acercan al premio.

En este canje existen a penas dos grupos, uno menos que en el primero. Observando detenidamente se puede concluir que los últimos dos grupos del canje anterior se fusionan en este canje en uno solo (grupo 2). Esto se deduce del tamaño de los grupos y sus características en cuanto al aumento en las probabilidades. Por otro lado, era esperable que existiera un grupo grande de gente que compra muy seguido debido a la forma en que se seleccionó la muestra, condicionando que los clientes elegidos tuvieran al menos cuatro canjes en el nivel 1 (lo que se traduce en muchas compras).

Es importante destacar además que, analizando el grupo 1 en el primer y segundo canje, se cumple el hecho de que existe un reinicio parcial de la probabilidad, consistente con la hipótesis 2.

Por otro lado, los puntos por vencer y puntos vencidos no presentan valores significativos en ninguno de los grupos, mientras que el parámetro asociado al share de compra es significativo para todos los grupos, lo que muestra nuevamente que parte de la tendencia de compra por estacionalidad se captura mediante este parámetro.

Canje 3

En la tabla n°42, se encuentra la evaluación de los modelos con distinta cantidad de grupos, mostrando para cada uno de ellos el logaritmo de la verosimilitud, así como el BIC y la diferencia entre el BIC para un modelo y el anterior.

Grupos	LL	BIC	Diferencia
1	-1.223	2.508	
2	-1.120	2.364	143,8
3	-1.099	2.385	-20,2

Tabla n° 46: Evaluación del modelo con distinto número de clases latentes
Fuente: Elaboración propia

El modelo con 2 clases latentes es el que presenta un mejor BIC. Los parámetros estimados para este modelo se muestran en la tabla n° 43.

	β_1 (Dt)	β_2 (SHAREt)	β_3 (VENCIDOST)	β_4 (X_VENCERT)	γ	α	π
Grupo 1	-1,021 (0,169)	12,168 (0,579)	-0,001* (0,019)	0,041* (0,076)	4,576 (0,493)	0,7452 (0,019)	0,2416 (0,321)
Grupo 2	-0,152 (0,069)	1,031 (0,496)	0,049* (0,097)	0,105* (0,33)	0,205 (0,397)	0,0971 (0,066)	0,7583 (-)**

* No son estadísticamente significativos, según test t de dos colas, $P > 0.1$

** Este es el nivel base de la probabilidad, por lo que la desviación estándar es nula.

Tabla n° 47: Parámetros estimados para el canje 3, con 2 clases latentes
Fuente: Elaboración propia

Por último, en la tabla n°44 se presenta la probabilidad de compra de los clientes de cada grupo, evaluado en la distancia al canje inicial y final, y el incremento porcentual de dicha probabilidad.

	$P(d_i=1)$	$P(d_i=0)$	Incremento %
Grupo 1	0,528	0,875	65,87%
Grupo 2	0,867	0,904	4,33%

Tabla n° 48: Probabilidad de compra para cada grupo, al comienzo y al final de la carrera
Fuente: Elaboración propia

Análisis de resultados canje 3

Nuevamente, la distancia al canje normalizada es significativa para todos los grupos, lo que confirma la presencia de la aceleración de compra para el canje 3 a medida que los clientes se acercan al premio, consistente con la hipótesis 1.

Para este canje, nuevamente existen sólo dos grupos, análogos a los del canje anterior. El grupo 2 es el grupo grande de gente que compra muy seguido y se ve poco afectada por el programa, mientras que el grupo 1 es el de los clientes que sufren un

mayor impacto en su probabilidad de compra. El incremento en probabilidad del grupo 1 es menor que en el canje anterior, al igual que en el caso homogéneo, y se observa que se sigue cumpliendo el reseteo post canje en ambos grupos.

Por último, los parámetros de los puntos vencidos y por vencer no son significativos, mientras que el parámetro asociado al share de compra es significativo para todos los grupos, lo que muestra nuevamente que parte de la tendencia de compra por estacionalidad se captura mediante este parámetro.

Canje 4

En la tabla nº45, se encuentra la evaluación de los modelos con distinta cantidad de grupos, mostrando para cada uno de ellos el logaritmo de la verosimilitud, así como el BIC y la diferencia entre el BIC para un modelo y el anterior.

Grupos	LL	BIC	Diferencia
1	-1.362	2.786	
2	-1.276	2.677	109,4
3	-1.266	2.717	-40,5

Tabla nº 49: Evaluación del modelo con distinto número de clases latentes
Fuente: Elaboración propia

El modelo con 2 clases latentes es el que presenta un mejor BIC. Los parámetros estimados para este modelo se muestran en la tabla nº 46.

	β_1 (Dt)	β_2 (SHAREt)	β_3 (VENCIDOST)	β_4 (X_VENCERT)	γ	α	π
Grupo 1	-0,567 (0,213)	11,906 (0,496)	0,170* (0,137)	-0,021* (0,041)	5,209 (0,512)	0,730 (0,026)	0,167 (0,101)
Grupo 2	-0,064 (0,023)	4,423 (0,643)	-0,181* (0,156)	0,158* (0,171)	0,001 (0,364)	0,420 (0,041)	0,832 (-)**

* No son estadísticamente significativos, según test t de dos colas, $P > 0.1$

** Este es el nivel base de la probabilidad, por lo que la desviación estándar es nula.

Tabla nº 50: Parámetros estimados para el canje 4, con 2 clases latentes
Fuente: Elaboración propia

Por último, en la tabla nº47 se presenta la probabilidad de compra promedio de los clientes de cada grupo, evaluado en la distancia al canje inicial y final, y el incremento porcentual de dicha probabilidad.

	$P(d_t=1)$	$P(d_t=0)$	Incremento %
Grupo 1	0,655	0,847	29,33%
Grupo 2	0,815	0,835	2,40%

Tabla nº 51: Probabilidad de compra para cada grupo, al comienzo y al final de la carrera
Fuente: Elaboración propia

Análisis de resultados nivel 4

Nuevamente, la distancia al canje normalizada es significativa para todos los grupos, lo que confirma la presencia de la aceleración de compra para el canje 4 a medida que los clientes se acercan al premio.

Al igual que en el segundo y tercer canje, se observan dos grupos donde uno de ellos domina en tamaño y corresponde a clientes que compran muy seguido y no son afectados en mayor medida por el programa. Se puede notar que el efecto de reseteo post canje sigue presente con respecto al canje anterior. En particular, el grupo 1 presenta el efecto deseado, donde la probabilidad al iniciar la carrera al cuarto canje es mayor que al iniciarla hacia el tercero, mientras que en el grupo 2 ocurre lo visto en el caso homogéneo, donde la probabilidad en el cuarto canje es menor. Esto es bueno, pues indica que el grupo de clientes que está siendo influenciado en mayor medida sufre del efecto del reseteo post premio parcial en todos los canjes, incluyendo el cuarto, lo que en el caso homogéneo no se podía observar.

Los puntos vencidos y por vencer no juegan ningún rol: Ambos son no significativos para los 2 grupos. Por último, es importante destacar que el parámetro asociado al share de compra es significativo para todos los grupos, lo que muestra nuevamente que parte de la tendencia de compra por estacionalidad se captura mediante este parámetro.

8.5 Hipótesis 4 – Preferencias por tipo de premio

H₄: La preferencia por premios hedónicos aumenta al aumentar el esfuerzo necesario para el canje de un premio.

En el contexto de los programas de lealtad, estudios muestran que la preferencia por el tipo de premios (hedónico v/s utilitario) depende de la cantidad de esfuerzo necesario para conseguirlos. En efecto, Kivetz y Simonson (2002) mostraron que, al aumentar los requerimientos de un programa de lealtad, los clientes prefieren recompensas hedonistas en vez de utilitaristas³⁰. De estos estudios surge la motivación de analizar si este comportamiento se repite en el programa de lealtad estudiado en esta investigación.

Dada la estructura de niveles de canje del programa, donde para canjear un premio se debe alcanzar una cierta cantidad de puntos fija y cada nivel va exigiendo cada vez más puntos, se espera que en niveles superiores la preferencia por premios hedónicos sea mayor que en niveles inferiores (debido a que para acumular más puntos se debe hacer un gasto mayor y, por ende, el esfuerzo necesario aumenta).

³⁰ Ver capítulo 4.2 para descripción detallada de este fenómeno

8.5.1 Datos

Se utilizaron los datos de todos los clientes que realizaron canjes en alguno de los primeros 4 niveles del programa. En total, corresponden a 20.697 clientes, los cuales realizaron 55.793 canjes (1.362 artículos diferentes) durante los 62 meses del período de estudio. En la tabla n°52, se muestra el total de clientes y canjes en cada nivel, así como la cantidad de artículos distintos que se canjearon en cada uno.

	Cientes	Canjes	SKU distintos
Nivel 1	17.827	35.114	608
Nivel 2	7.011	10.605	332
Nivel 3	5.170	8.746	224
Nivel 4	879	1.328	178

Tabla n° 52: Cantidad de clientes, canjes y artículos distintos que se canjearon, por nivel
Fuente: Elaboración propia.

Se puede notar que en los niveles inferiores hay una mayor cantidad de canjes y, por ende, se seleccionan más artículos distintos que en otros niveles. Esto se debe principalmente a que para canjear en niveles superiores, se necesita una mayor cantidad de puntos y, por ende, se debe gastar una mayor cantidad de dinero. Bajo este esquema, es esperable que haya menos gente que logre acumular puntos suficientes para canjear premios en los niveles superiores. Por ejemplo, para canjear un premio del nivel 1, se deben acumular 5.000 puntos, lo que equivale a un gasto de \$600.000, mientras que para canjear un premio del nivel 4, se necesitan 24.000 puntos, equivalente a \$2.880.000. En la figura n°21, se muestra la cantidad de dinero que se debe invertir para acumular puntos suficientes para cada nivel de canje del programa.

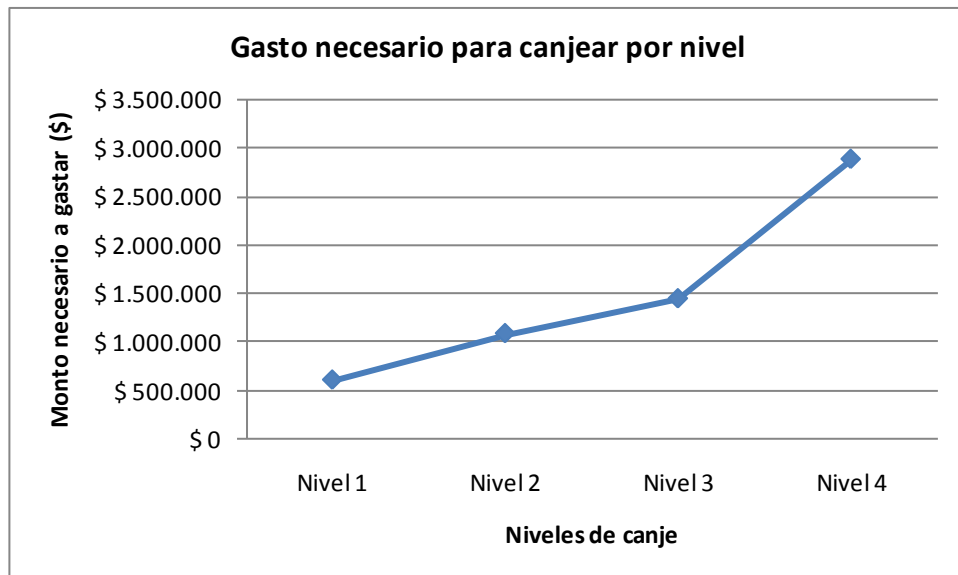


Figura n° 21: Gasto necesario para canjear un premio, por nivel
Fuente: Elaboración propia.

Como se comentaba anteriormente, se puede apreciar que en niveles superiores la cantidad de dinero a gastar es mayor, además se ve un crecimiento importante al pasar del nivel 3 al 4.

8.5.2 Modelo

Para ajustar los datos, se utiliza una regresión logística binaria³¹. El evento a modelar es el tipo premio que elige el cliente al canjear. De aquí, la variable dependiente corresponde a:

$$y_{ikt} = \begin{cases} 1 & \text{Si el premio } k \text{ canjeado por el cliente } i \text{ en el mes } t \text{ es hedónico} \\ 0 & \text{Sino} \end{cases}$$

Se desea saber cómo afectan ciertos factores externos a la probabilidad de canje hedonista. Las variables independientes que representan a estos factores corresponden a:

$$x_{iktn} = \begin{cases} 1 & \text{Si el premio } k \text{ canjeado por cliente } i \text{ en mes } t \text{ es del nivel de canje } n \\ 0 & \text{Sino} \end{cases}$$

$$z_{ikt} = \begin{cases} 1 & \text{Si premio } k \text{ canjeado por cliente } i \text{ en mes anterior a } t \text{ fue hedonista} \\ 0 & \text{Sino} \end{cases}$$

$$m_{ikt} = \begin{cases} 1 & \text{Si premio } k \text{ canjeado por cliente } i \text{ se canjeó el mes } t \\ 0 & \text{Sino} \end{cases}$$

La primera variable independiente indica el nivel al que pertenece el premio canjeado, mientras que la segunda indica si el cliente canjeó o no un premio hedonista en su canje anterior. Esta última variable se agrega para ver si existe una dependencia del tipo de premio elegido con respecto al canje anterior. Es posible que se dé una especie de “lealtad” hacia el tipo de premio, donde una vez que el cliente canjea un premio hedonista o utilitarista, ocurra que la probabilidad de volver a elegir un premio de este tipo en el canje siguiente sea mayor. La última variable indica si un premio se canjeó en cierto mes, y sirve para ver si existe efecto de estacionalidad.

Para clasificar los premios en hedónicos o utilitarios, se utilizaron los conceptos propuestos por Kevin Voss, Eric Spangenberg y Bianca Grohmann (2003). Según esta clasificación, de los 55.793 canjes, 3.255 se consideran hedonistas, lo que corresponde al 5.8% de los canjes.

³¹ Ver capítulo 5.3.3.

8.5.3 Resultados

En la tabla nº53 se muestran los resultados del logit.

	β
Constante	-3,630 (0,078)
Nivel 2	0,150 (0,055)
Nivel 3	1,103 (0,043)
Nivel 4	1,991 (0,072)
Último_canje	0,847 (0,064)
Febrero	-0,110* (0,108)
Marzo	-0,026* (0,102)
Abril	0,706 (0,092)
Mayo	0,766 (0,096)
Junio	0,816 (0,095)
Julio	0,684 (0,096)
Agosto	0,833 (0,095)
Septiembre	0,394 (0,104)
Octubre	0,097 (0,107)
Noviembre	0,307 (0,103)
Diciembre	0,432 (0,093)

Tabla nº 53: Resultados modelo logit binario con las 3 variables independientes
Fuente: Elaboración propia.

Se observa que el parámetro β para los niveles 2, 3 y 4 es positivo, lo que implica que la probabilidad de que un cliente elija un premio hedonista en estos niveles de canje es mayor que si el cliente está en el nivel 1 (nivel base). Además se observa que el parámetro β es creciente en los niveles, es decir, a niveles más altos la probabilidad de preferir un premio hedonista va aumentando. Todo esto es consistente con H3. Por otro lado, el parámetro beta de la variable independiente último_canje es positivo, lo

que indica que efectivamente existe una mayor probabilidad de elegir un premio hedónico si en el canje anterior se eligió también uno hedónico.

Se puede observar que los meses de febrero y marzo no entregan resultados significativos. Para el resto, el parámetro es positivo, lo que implica que en estos meses (de abril a diciembre) la probabilidad de canjear premios hedonistas es mayor que en enero. Además, es a mediados de año cuando la probabilidad de elegir premios hedonistas es mayor. Esto podría explicarse como una forma de los clientes de darse un gusto, pues a esas alturas del año están en sus trabajos y muy lejos de unas posibles vacaciones.

Por último, la constante es negativa, lo que indica que, en general, los clientes prefieren canjear premios utilitaristas.

8.5.4 Discusión

De los resultados anteriormente expuestos, se puede concluir que existe una mayor probabilidad de elegir un premio hedónico al estar en niveles superiores de canje, lo que es consistente con H4. Sin embargo, hasta el momento no se ha dado cuenta de una situación que podría estar influyendo en estos resultados: la cantidad de premios hedonistas para elegir. Podría ocurrir que los clientes tiendan a preferir productos hedónicos en niveles superiores de canje, simplemente porque en esos niveles la cantidad de opciones de premios hedónicos disponibles sea mayor. Asimismo, puede que los clientes en niveles inferiores no elijan premios hedónicos debido a que en ellos la mayoría de los premios sean de carácter utilitarios.

Para descartar este efecto, en la tabla nº54 se muestra el porcentaje de productos hedonistas por nivel, en cada uno de los catálogos utilizados durante el período de estudio.

	Nivel 1	Nivel 2	Nivel 3	Nivel 4
Catálogo 1	19,2%	22,6%	21,6%	16,4%
Catálogo 2	22,2%	25,8%	14,9%	12,6%
Catálogo 3	25,0%	26,9%	16,7%	30,7%
Catálogo 4	16,7%	14,3%	24,6%	20,8%
Catálogo 5	21,1%	27,4%	20,9%	20,0%
Catálogo 6	18,6%	16,4%	23,7%	17,9%
Promedio	20,5%	22,2%	20,4%	19,7%

Tabla nº 54: Porcentaje de premios hedonistas disponibles por nivel, por catálogo
Fuente: Elaboración propia.

Se observa que la cantidad de premios hedónicos es más bien constante en los catálogos a través de los niveles, incluso ocurre que en niveles superiores hay menos premios hedónicos que en niveles inferiores en algunos de ellos. En la figura nº22 se muestra la probabilidad estimada promedio de elegir un premio hedonista en cada nivel

(según los resultados de la regresión logística binaria) y el porcentaje de premios hedonistas promedio por nivel.

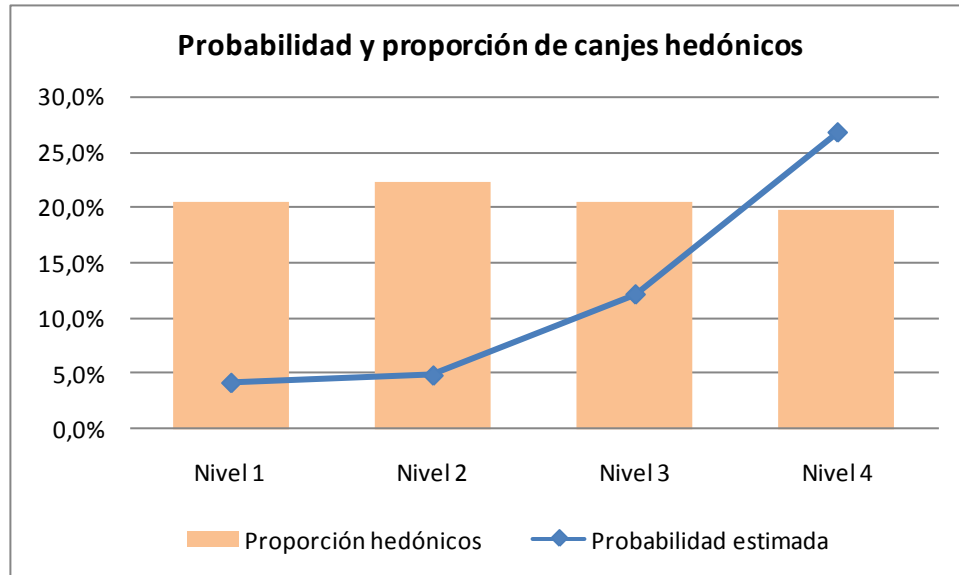


Figura nº 22: Probabilidad estimada y proporción de canjes hedónicos
Fuente: Elaboración propia.

De aquí se observa claramente que la proporción de premios hedonistas es relativamente constante entre los niveles, mientras que la probabilidad de elegir un premio de este tipo aumenta exponencialmente con los niveles. En otras palabras, elegir un premio hedonista es más probable en los niveles de canje superiores, donde la cantidad de esfuerzo necesario es mayor, lo que es consistente con H4.

Se puede concluir entonces que la diferencia entre la probabilidad de elegir un premio hedonista y la proporción de premios de este tipo en los catálogos se debe a la diferencia en los requerimientos de cada uno de los niveles. A mayor requerimiento, mayor es la probabilidad de elegir un premio hedónico. Además, parte de este efecto se explica también por una “lealtad” al tipo de canje, es decir, cuando un cliente canjea un premio de carácter hedonista en un período, es más probable que vuelva a elegir un premio de este tipo en el canje siguiente. Por último, es importante notar que estos posiblemente no son los únicos factores que inciden en el problema, pues podría haber diferencias también de acuerdo al sexo del cliente, estrato socio-económico etc. Sin embargo, no se cuenta con esta información al momento de este estudio, por lo que queda propuesto para estudios futuros.

8.5.5 Heterogeneidad

Un último aspecto a considerar es el sesgo por heterogeneidad, por lo que se realizará un análisis de clases latentes.

8.5.5.1 Modelo

Para efectos de este análisis, se considerará que todos los parámetros del modelo pueden variar a través de los individuos. De esta forma, al considerar las clases latentes, la función de probabilidad de que un cliente canjee un premio hedonista queda dada por la siguiente expresión:

$$P_{ig} = \frac{\exp\left(\sum_{w=1}^W \beta_{wg} X_{iw}\right)}{1 + \exp\left(\sum_{w=1}^W \beta_{wg} X_{iw}\right)} \quad (52)$$

donde el índice W son los atributos o variables de marketing utilizadas en el modelo. En esta expresión, cada uno de los parámetros a estimar varía con respecto a un índice “g”, que representa los grupos o clases a considerar. Luego, la función de verosimilitud queda dada por

$$L = \prod_{i=1}^N \left(\sum_{g=1}^G \left(P_{ig}^{\delta_i} * (1 - P_{ig})^{1-\delta_i} * \pi_g \right) \right) \quad (53)$$

Es importante recordar que la probabilidad π_g es también un parámetro a estimar en este modelo.

8.5.5.2 Resultados

Como primer paso, se estima el modelo con una sola clase latente, con lo que se recuperan los resultados del caso homogéneo. Después, se evalúan los modelos incrementando el número de clases latentes, de uno en uno. Finalmente, se decide cuál es el número de clases latentes que mejor describe a la muestra, usando el criterio de información bayesiana.

En la tabla nº55, se encuentra la evaluación de los modelos con distinta cantidad de grupos, mostrando para cada uno de ellos el logaritmo de la verosimilitud, así como el estadístico del criterio de información bayesiana (BIC) y la diferencia entre el BIC para un modelo y el anterior.

Grupos	LL	BIC	Diferencia
1	-11.614	23.414	
2	-11.500	23.371	42,8
3	-11.495	23.548	-176,2

Tabla nº 55: Evaluación del modelo con distinto número de clases latentes
Fuente: Elaboración propia

El modelo con 2 clases latentes es el que presenta un mejor BIC. Los parámetros estimados para este modelo se muestran en la tabla nº 56.

	Grupo 1	Grupo 2
Constante	-4,194 (0,129)	-4,074 (0,897)
Nivel 2	-8,268 (0,005)	0,957 (0,067)
Nivel 3	-2,534 (0,000)	1,971 (0,220)
Nivel 4	-8,616 (0,001)	2,922 (0,401)
Último_canje	3,175 (0,001)	0,508 (0,104)
Febrero	-0,561 (0,000)	-0,082* (0,073)
Marzo	1,449 (0,009)	-0,184* (0,173)
Abril	3,359 (0,026)	0,175 (0,074)
Mayo	3,606 (0,021)	0,156 (0,055)
Junio	3,864 (0,022)	0,083 (0,050)
Julio	3,269 (0,022)	0,197 (0,061)
Agosto	3,182 (0,017)	0,485 (0,075)
Septiembre	0,642 (0,001)	0,405 (0,080)
Octubre	0,488 (0,001)	0,099 (0,069)
Noviembre	1,688 (0,005)	0,170 (0,068)
Diciembre	0,682 (0,002)	0,455 (0,138)
Prob	0,145 (0,214)	0,854 (-)**

* No son estadísticamente significativos, según test t de dos colas, $P > 0.1$

** Este es el nivel base de la probabilidad, por lo que la desviación estándar es nula.

Tabla nº 56: Parámetros estimados con 2 clases latentes

Fuente: Elaboración propia

8.5.5.3 Análisis de los resultados

Del análisis con clases latentes se obtienen dos grupos de clientes que representan a la población. Es interesante notar que el segundo grupo es el más grande, y es análogo a los resultados obtenidos en el caso homogéneo: el parámetro β para los niveles 2, 3 y 4 es positivo y creciente a través de los niveles, lo que implica que los clientes tienen una mayor probabilidad de elegir premios hedónicos que en el nivel 1 y, además, la probabilidad va creciendo a medida que se va subiendo de nivel. Por otro lado, y también análogo al caso homogéneo, se observa en el grupo 2 que febrero y marzo son negativos pero no significativos, mientras que el resto de los meses son todos positivos. Finalmente, se observa que el tipo de premio del canje anterior es también un factor significativo, indicando que si el canje anterior fue hedónico, el actual es más probable que también lo sea.

Los resultados para el grupo 1, en cambio, presentan características un tanto distintas. En este grupo, es en el nivel 1 donde existe mayor probabilidad de elegir un premio hedónico, siendo los parámetros asociados a los niveles 2, 3 y 4 todos negativos, los que tampoco presentan ningún tipo de tendencia. Una explicación para esto es que es posible que en este grupo haya clientes que compran muy seguido y a quienes el programa de lealtad no los afecta en mayor medida, por lo que les da igual el tipo de premio a canjear. En este grupo se observa que el tipo de premio del último canje es aún más importante y, además, todos los meses tienen un valor positivo y significativo. Finalmente, en ambos grupos la constante es negativa, lo que muestra que los consumidores prefieren en general premios utilitarios.

De este análisis se puede concluir que el grupo importante es el grupo 2, de mayor tamaño, el que ya estaba bien representado por el estudio del caso homogéneo. El estudio de la heterogeneidad ayudó a identificar que existe un grupo relativamente pequeño de clientes que no se comportan de la misma forma, y que parecieran no dejarse influenciar tanto por el programa.

8.6 Hipótesis 5 – Aceleración de compra hedonista versus utilitarista

H₅: Los clientes que canjean premios hedonistas aceleran más que aquellos que canjean premios utilitarios

De acuerdo con la hipótesis de aceleración de compra H1, la motivación para esforzarse aumenta con el progreso que se ha realizado hacia el canje de un premio. Por otro lado, la hipótesis de preferencias por el tipo de premio H4 indica que al aumentar los requerimientos del programa de lealtad, la preferencia por premios hedónicos aumenta. De lo anterior, un análisis interesante consiste en estudiar si los clientes se esfuerzan más a medida que se acercan al canje de un premio (es decir, aceleran más) si canjean un premio hedónico con respecto a si canjean un premio utilitario, dado un nivel de canje particular. Por ejemplo, la intuición indica que si se toma el nivel de canje 1, los clientes que desean canjear un premio de tipo hedónico

debieran acelerar más que otro que espera canjear un premio utilitario en el mismo nivel, de acuerdo a lo discutido en las hipótesis anteriores.

8.6.1 Datos

Se utilizaron los datos de todos los clientes que realizaron canjes en alguno de los primeros 4 niveles del programa. En total, corresponden a 20.697 clientes, los cuales realizaron 55.793 canjes durante los 62 meses del período de estudio. Cada una de las 4 bases de datos, correspondientes a los 4 niveles de canje, fue dividida a su vez en 2 partes: datos de canjes hedonistas y datos de canjes utilitaristas. Con esto, se obtuvo un total de 8 bases de datos. En las tablas nº57 y nº58 se muestra el número de clientes, canjes y canjes promedio por cliente en los 4 niveles para canjes de tipo hedonistas y utilitaristas respectivamente.

	Cientes	Canjes	Canjes promedio
Nivel 1	1.273	1.449	1,14
Nivel 2	452	468	1,04
Nivel 3	916	1.010	1,10
Nivel 4	270	318	1,18

Tabla nº 57: Clientes, canjes y canjes promedio por nivel, canjes hedonistas
Fuente: Elaboración propia

	Cientes	Canjes	Canjes promedio
Nivel 1	17.508	33.665	1,92
Nivel 2	6.809	10.137	1,49
Nivel 3	4.813	7.736	1,61
Nivel 4	693	1.010	1,46

Tabla nº 58: Clientes, canjes y canjes promedio por nivel, canjes utilitarios
Fuente: Elaboración propia

8.6.2 Modelo

Para estudiar el fenómeno se utiliza un modelo hazard proporcional en tiempo discreto. De acuerdo a los resultados de H1, se usa el modelo 4 que fue el de mejor ajuste y BIC. Las covariables que se utilizan son entonces la distancia al canje suavizada por el número de canjes D_t , el share de compra y las binarias de puntos vencidos y por vencer. La función de probabilidad queda como se muestra a continuación:

$$\Pr(t, X_t) = 1 - \exp \left(- e^{(\beta_1 D_t + \beta_2 SHARE_t + \beta_3 VENCIDOS_t + \beta_4 POR_VENCER)} \int_{t-1}^t h(s) ds \right) \quad (54)$$

8.6.3 Resultados

Para estimar los parámetros, se utilizó la parametrización Weibull del modelo hazard proporcional en tiempo discreto, pues fue el que produjo un mejor ajuste en la hipótesis 1. El ajuste se midió en base al logaritmo natural de la verosimilitud (-2LL) y el criterio de información bayesiano (BIC).

En las tablas nº59 y nº60 se muestran los resultados de los 4 niveles para los canjes hedonistas y utilitaristas respectivamente.

	β_1 (Dt)	β_2 (SHAREt)	β_3 (VENCIDost)	β_4 (X_VENCERT)	γ	α	LL	BIC
Nivel 1	-0,838 (0,002)	2,929 (0,107)	0,060* (0,051)	-0,036* (0,032)	0,101 (0,003)	0,524 (0,000)	-5.547	11.161
Nivel 2	-0,837 (0,006)	2,930 (0,951)	0,016* (0,064)	-0,095* (0,064)	0,114 (0,049)	0,548 (0,001)	-1.444	2.946
Nivel 3	-0,785 (0,004)	2,464 (0,210)	-0,063* (0,058)	-0,001* (0,022)	0,199 (0,036)	0,447 (0,000)	-2.108	4.280
Nivel 4	-0,724 (0,027)	2,864 (1,344)	0,029* (0,029)	-0,089* (0,077)	0,085 (0,075)	0,455 (0,001)	-596	1.246

* No son estadísticamente distintos de cero, según test t, $P > 0.1$

Tabla nº 59: Parámetros estimados modelo Hazard, parametrización weibull, canjes hedónicos
Fuente: Elaboración propia

	β_1 (Dt)	β_2 (SHAREt)	β_3 (VENCIDost)	β_4 (X_VENCERT)	γ	α	LL	BIC
Nivel 1	-0,811 (0,000)	1,969 (0,001)	0,039* (0,033)	0,055* (0,050)	0,428 (0,001)	0,499 (0,000)	-168.874	337.839
Nivel 2	-0,901 (0,000)	1,670 (0,007)	0,043* (0,040)	0,066* (0,073)	0,893 (0,019)	0,458 (0,000)	-31.898	63.876
Nivel 3	-0,965 (0,000)	1,517 (0,012)	0,021* (0,020)	0,066* (0,061)	1,459 (0,110)	0,430 (0,000)	-18.812	37.703
Nivel 4	-1,064 (0,005)	2,227 (0,247)	-0,113* (0,104)	0,010* (0,013)	0,509 (0,344)	0,403 (0,000)	-699	1.462

* No son estadísticamente distintos de cero, según test t, $P > 0.1$

Tabla nº 60: Parámetros estimados modelo Hazard, parametrización weibull, canjes utilitarios
Fuente: Elaboración propia

De las tablas anteriores, se puede observar que el parámetro estimado asociado a la distancia al canje suavizada es menor que cero para todos los niveles lo que es consistente con lo estudiado en la hipótesis 1. Se puede observar también que los puntos vencidos y por vencer no son significativos en ningún caso.

8.6.4 Discusión

Para determinar si los consumidores que canjean premios hedonistas presentan un comportamiento distinto a los que canjean premios utilitarios, se calculan las probabilidades de compra al comienzo y al final de la carrera (distancia igual a 1 y 0 respectivamente), para todos los niveles y ambos tipos de canje. Los resultados se presentan en la tabla nº61 y 62.

	$P(D_t=1)$	$P(D_t=0)$	Incremento %
Nivel 1	0,755	0,961	27,3%
Nivel 2	0,785	0,971	23,7%
Nivel 3	0,800	0,971	21,3%
Nivel 4	0,853	0,981	15,0%

Tabla nº 61: Probabilidad de compra, al comienzo y al final de la carrera, canjes hedonistas
Fuente: Elaboración propia

	$P(D_t=1)$	$P(D_t=0)$	Incremento %
Nivel 1	0,777	0,966	24,3%
Nivel 2	0,820	0,985	20,2%
Nivel 3	0,830	0,990	19,3%
Nivel 4	0,881	0,998	13,3%

Tabla nº 62: Probabilidad de compra, al comienzo y al final de la carrera, canjes utilitaristas
Fuente: Elaboración propia

De las tablas anteriores se puede observar que, consistente con lo discutido en capítulos anteriores, el incremento en la probabilidad va disminuyendo a través de los niveles. Esto se observa para ambos tipos de canje. Además, consistente con la hipótesis 5, se observa que en todos los niveles, los canjes hedónicos tienen un mayor incremento relativo en la probabilidad de compra, comparado con aquellos clientes que canjean premios utilitarios³².

IX. CONCLUSIONES Y TRABAJOS FUTUROS

Los programas de lealtad son, hoy en día, una herramienta ampliamente utilizada por las empresas de retail en el país. Su objetivo principal es tratar de retener a los clientes en la empresa, ofreciéndoles un servicio adicional que les entrega valor, al premiarlos de acuerdo a su frecuencia de compra. Sin embargo, no es fácil determinar

³² Esta diferencia es significativa, para demostrarlo se calcularon los intervalos de confianza de los parámetros asociados a la distancia, con lo cuál se estimaron los intervalos de confianza de las probabilidades al comienzo y al final del canje (similar a lo expuesto en la hipótesis 3), y se calculó el incremento porcentual según estas probabilidades, las que no se intersectaron en ningún caso.

si los programas de lealtad son realmente rentables para la empresa, es decir, si se justifica su implementación y el gasto adicional que estos significan, pues el impacto que tienen sobre el comportamiento de compra de sus clientes es difícil de cuantificar.

En la presente investigación, se han estudiado distintos efectos que afectan el comportamiento de compra y canje de los clientes pertenecientes a un programa de lealtad. En particular, se ha demostrado que:

- Los clientes del programa de lealtad aceleran sus compras a medida que se acercan al canje de un premio. Esta aceleración es de distinta magnitud para diferentes grupos de clientes en cada nivel de canje, pero el efecto en sí siempre se observa. Además, se encontró que la aceleración va decreciendo a medida que se aumenta el nivel de canje. Esto indica que, en general, los clientes que canjean premios en niveles superiores tienden a verse menos motivados por el programa de lealtad, probablemente porque para ellos se ha hecho habitual el comprar más seguido y el programa no genera un gran impacto. Al segmentar los clientes, se observó que los efectos, en general, se mantienen, salvo algunos grupos de clientes que compran muy seguido y que no son afectados en mayor medida por el programa. De todas formas, estos grupos son relativamente pequeños.
- Los clientes del programa de lealtad aumentan el esfuerzo invertido en cada compra, medido en dinero gastado, a medida que se acercan al canje de un premio. Los puntos vencidos y por vencer mostraron ser un parámetro relevante para la magnitud del esfuerzo, donde aquellos clientes que tienen puntos por vencer gastan más (para acumular más puntos, canjear y no perderlos) y los que tienen puntos vencidos gastan menos (probablemente por un sentimiento de rechazo a esta pérdida de puntos).
- Los clientes del programa, después de canjear, experimentan un efecto de reinicio post premio parcial, volviendo a un estado donde su probabilidad de compra inicial es menor que al final del canje anterior, pero al mismo tiempo mayor que al comienzo de del canje anterior. Esto es bueno para el programa, pues indica que los clientes se van motivando a medida que van canjeando y cada vez su probabilidad base al comienzo de la carrera hacia un premio es mayor. Este efecto se observó para los primeros 3 canjes. Al segmentar los clientes se encontró que aparece un grupo pequeño de consumidores que no se ven afectados en mayor medida por el programa, pues compran muy seguido de todas formas, y que el otro segmento (que incluye a los clientes que sí cambian su comportamiento) tiene un reinicio parcial en los 4 niveles estudiados..
- Los requisitos del programa tienen un impacto en el tipo de premio seleccionado por el cliente. Se observó que existe una mayor probabilidad de elegir un premio hedónico cuando el cliente canjea en niveles de canje más altos. Además, se descartó que este efecto se debiera simplemente a una mayor cantidad de

premios hedónicos disponibles, pues en efecto la proporción de premios hedónicos era similar para todos los niveles. También se demostró que existe una tendencia a elegir premios hedónicos cuando ya se ha canjeado un premio de este tipo en el canje anterior, lo que indica la existencia de una cierta lealtad hacia el tipo de premio. Al segmentar a los clientes se observó que existe un grupo pequeño de clientes que no se ven afectados en mayor medida por el programa, y el resto cae en un grupo con características similares a las ya explicadas.

- La magnitud de aceleración en los tiempos entre compras sucesivas de los clientes que canjean premios hedónicos, es mayor que para aquellos que canjean premios utilitarios, para todos los niveles de canje.

De los resultados anteriores, se puede concluir que el programa de lealtad efectivamente está modificando favorablemente el comportamiento de compra y canje de sus clientes. La disminución en los tiempos entre compras sucesivas es quizás el resultado más importante, pues este efecto indica que los clientes están aumentando su frecuencia de compra gracias al programa de lealtad. Se demostró que los clientes que canjean en niveles superiores aceleran menos que los de niveles inferiores, pero de todas formas se logra un impacto en el comportamiento, por lo que es buena la existencia de estos niveles. Además, es importante destacar que la aceleración no es sólo en la frecuencia de compra, sino también en la magnitud del esfuerzo invertido en cada una de ellas. Dicho de otro modo, los clientes están yendo a comprar más seguido y, además, están gastando más dinero en cada una de estas compras, a medida que se acercan al nivel de canje deseado, lo que indica que el programa de lealtad es sumamente efectivo.

Se demostró además la importancia de los premios hedonistas y utilitarios. En particular, se observó que aquellos clientes que canjean premios hedónicos aceleran más que los que prefieren premios utilitarios. Esto indica que es necesaria la existencia de premios hedónicos en el mix de premios, para que aquellos clientes que los prefieren los puedan elegir, pues ellos aceleran en mayor medida que el resto. Además, se probó que los clientes que se esfuerzan más, tienden a preferir premios hedónicos. Sin embargo, hay que tener cuidado con este resultado pues, como se observó, la elección de un premio hedónico podría significar simplemente una lealtad a este tipo de premios; el nivel de canje es importante, y explica parte de la elección del tipo de premio, pero no hay que dejar de lado el efecto de la lealtad.

Una posibilidad de mejora al programa de lealtad estudiado consiste en una mejor elección de los niveles y la calidad de los premios entregados. Actualmente, la mayoría de los canjes ocurren en los primeros niveles de canje, y en los niveles superiores (omitidos en este estudio) casi no existen canjes, debido a la alta cantidad de dinero necesaria. La idea principal de una estructura de canje por niveles es motivar al cliente a que pase de un nivel al siguiente, para que compre con mayor frecuencia e invierta más dinero en la tienda. Sin embargo, la estructura actual no está logrando totalmente este objetivo.

Esta investigación ha resultado ser útil para entender de mejor forma las respuestas de los consumidores ante un programa de lealtad. De todas formas, existen muchos estudios adicionales que se podrían realizar. Algunos trabajos futuros propuestos podrían ir en la siguiente dirección:

- Estudiar el efecto de la venta de puntos en el comportamiento de los consumidores. Existen en la actualidad muchos programas de lealtad que ofrecen comprar los puntos restantes para el canje de un premio. A primera vista, esto permitiría al cliente no perder los puntos cuando estos están por expirar. Sin embargo, ¿Es esto rentable para la empresa? ¿Cuál es el precio al que deberían vender los puntos? ¿Es mejor venderlos más caro o más barato que los puntos que se acumulan al realizar una compra de un producto de la tienda? Estas son algunas de las problemáticas que se podrían abordar.
- Realizar un estudio que permita determinar cuál es la mejor estructura de canjes para este programa de lealtad particular. Esto podría dar luces de qué estructura es mejor para las tiendas de retail en general. Como se comentó con anterioridad, la estructura de niveles logra impactar el comportamiento de compra y canje de los clientes, sin embargo pareciera que el esfuerzo necesario no se condice con los premios entregados en los distintos niveles. Esto se observa del hecho que los clientes concentran sus canjes en los primeros niveles, pero pareciera que no se ven mayormente estimulados a intentar alcanzar el nivel de canje siguiente. A la fecha, el programa de lealtad ha evolucionado, agregando nuevos niveles de canje y cambiando los premios entregados, por lo que se podrían usar estos datos para realizar los estudios pertinentes y comparar la situación actual al año 2012 con la situación estudiada en esta investigación.
- Realizar experimentos que permitan estudiar otros de los efectos planteados en el capítulo IV, como por ejemplo el efecto de ilusión al progreso, o el efecto del esfuerzo como determinante en la elección de premios con riesgo. Para esto, se necesitaría una alta colaboración por parte de la empresa para poder realizar los experimentos pertinentes.

REFERENCIAS

- Bender, Augustin y Blettner (2005): “**Generating Survival Times to Simulate Cox Proportional Hazards Models**”. *Statistics in Medicine*, 24(11), 1713-1723.
- Bettman, Luce y Payne (1998): “**Constructive Consumer Choice Processes**”. *Journal of Consumer Research*, 25 (3), 187–217.
- Blattberg, R., Kim, B. y Neslin, S. (2008). “**Database Marketing: Analyzing and Managing Customers**”. *International Series in Quantitative Marketing*. Volumen 18.
- Crowley, A., Spangenberg, E. y Voss, K. (1997). “**Measuring the Hedonic and Utilitarian Dimensions of Attitude: A Generally Applicable Scale**”. *Advances in Consumer Research*, 24 (1), 235-241.
- Cox, D. R. (1972). “**Regression Models and Life Tables**”. *Journal of the Royal Statistical Society Series B*, 34, 187-220.
- Durán, G. y Kremerman, M. (2008). “**Caracterización del Sector Retail: Una mirada general**”. *Cuaderno de investigación Fundación Sol*, nº7, 19p..
- Greene, William (2003). “**Econometric Analysis**”. Prentice Hall.
- Humby, Hunt y Phillips (2007). “**Scoring Points: How Tesco Continues to Win Customer Loyalty**”. Segunda edición. Kogan Page Publishers.
- Kivetz, Ran, Oleg Urminsky y Yuhuang Zheng (2006). “**The Goal-Gradient Hypothesis Resurrected: Purchase Acceleration, Illusionary Goal Progress, and Customer Retention**”. *Journal of Marketing Research*, 43 (1), 39-58.
- Kivetz, Ran y Simonson, Itamar (2002). “**Earning the Right to Indulge: Effort as a Determinant of Customer Preferences Toward Frequency Program Rewards**”. *Journal of Marketing Research*, 39 (2), 155-170.
- Kivetz, Ran (2003). “**The effect of effort and intrinsic motivation on risky choice**”. *Marketing Science*, 22 (4), 477-502.
- Kivetz, Ran y Simonson, Itamar (2003). “**The Idiosyncratic Fit Heuristic: Effort Advantage as a Determinant of Consumer Response to Loyalty Programs**”. *Journal of Marketing Research*, 40 (4), 454-467.
- Kivetz, Ran (2005). “**Promotion Reactance: The Role of Effort-Reward Congruity**”. *Journal of Consumer Research*, 31 (Marzo). 725-736.

- Kopalle, P.K. y Neslin, S.A. (2003). **“The Economic Viability of Frequency Reward Programs in a Strategic Competitive Environment”**. Review of Marketing Science, 1 (1).
- Lancaster, T. (1979). **“Econometric methods for the duration of unemployment”**. Econometrica, 47, 939–956.
- Monroy, Vidal y Saade (2009). **“Análisis de clases latentes: una técnica para detectar heterogeneidad en poblaciones”**. CENEVAL (México). 65p.
- Neculmán, Gonzalo (2010). **“Retail en Chile Evolución, Crecimiento y Riesgos”**. Reporte especial, clasificadora de riesgo Humphreys. Agosto 2010. 30p. .
- Rubinson y Baldinger (1996). **“Brand Loyalty, the link between attitude and behaviour”**. Journal of Advertising Research, 36 (6), 22-34.
- Sarkis, Sebastián (2009). **“Medición del efecto de un programa de lealtad en el comportamiento de compra y canje de sus clientes”**. Memoria para optar al título de Ingeniero Civil Industrial. Universidad de Chile, Departamento de Ingeniería Civil Industrial. 77p.
- Seetharaman, P. y Chintagunta, P. (2003). **“The Proportional Hazard Model for Purchase Timing: A comparison of Alternative Specifications”**. Journal of Business and Economic Statistics, 21 (3), 368-382.
- Sigelman, Lee y Langche Zeng (1999). **“Analyzing Censored and Sample-Selected Data with Tobit and Heckit Models”**. Political Analysis, 8 (2). 167–182.
- Tobin, James (1958), **“Liquidity Preference as Behavior Towards Risk”** The Review of Economic Studies, 25 (1), 65–86.
- Voss, Kevin, Eric Spangenberg y Bianca Grohmann (2003). **“Measuring the Hedonic and Utilitarian Dimensions of Consumer Attitude”**. Journal of Marketing Research, 40 (3), 310-320.
- Walke, Rainer (2010). **“Example for a Piecewise Constant Hazard Data Simulation in R”**. Max Planck Institute for Demographic Research.
- Winer, Russell (1986). **“A Reference Price Model of Brand Choice for Frequently Purchased Products”**. Journal of Consumer Research, 13 (Septiembre), 250–256.

- Wooldridge, Jeffrey (2002). **“Econometric Analysis of Cross Section and Panel Data”**. MIT Press.
- Yi, Youjae y Jeon, Hoseong (2003). **“Effects of Loyalty Programs on Value Perception, Program Loyalty and Brand Loyalty”**. Journal of the Academy of Marketing Science, 31 (3), 229-240.

ANEXOS

Anexo A: Estimación de parámetros del modelo 1 según otras parametrizaciones del modelo hazard, por nivel de canje:

	β_1 (dt)	β_2 (LN_Canjes)	γ	LL	BIC
Nivel 1	-1,122 (0,039)	0,438 (0,331)	2,783 (0,086)	-204.888	409.815
Nivel 2	-1,080 (0,184)	0,432 (0,077)	3,189 (0,764)	-40.212	80.460
Nivel 3	-1,070 (0,000)	0,479 (0,000)	3,290 (0,002)	-25.871	51.776
Nivel 4	-1,041 (0,016)	0,430 (0,037)	3,843 (0,055)	-2.544	5.115

Tabla nº 63: Parámetros estimados según distribución exponencial
Fuente: Elaboración propia

	β_1 (dt)	β_2 (LN_Canjes)	γ	α	LL	BIC
Nivel 1	-0,802 (0,000)	0,328 (0,000)	10,086 (0,009)	0,493 (0,000)	-177.028	354.108
Nivel 2	-0,757 (0,000)	0,317 (0,000)	15,796 (0,219)	0,455 (0,000)	-33.726	67.498
Nivel 3	-0,722 (0,000)	0,339 (0,000)	20,413 (0,196)	0,427 (0,000)	-21.087	42.218
Nivel 4	-0,732 (0,001)	0,317 (0,004)	31,200 (0,089)	0,414 (0,000)	-2.039	4.115

Tabla nº 64: Parámetros estimados según distribución weibull
Fuente: Elaboración propia

	β_1 (dt)	β_2 (LN_Canjes)	γ	α	LL	BIC
Nivel 1	-0,766 (0,046)	0,327 (0,037)	3,975 (0,986)	2,186 (0,137)	-178.241	356.534
Nivel 2	-0,731 (0,000)	0,316 (0,000)	4,702 (0,009)	2,214 (0,001)	-33.905	67.856
Nivel 3	-0,700 (0,000)	0,337 (0,000)	5,317 (0,026)	2,125 (0,002)	-21.184	42.413
Nivel 4	-0,721 (0,008)	0,317 (0,013)	5,650 (0,084)	2,370 (0,025)	-2.041	4.119

Tabla nº 65: Parámetros estimados según distribución log-logística
Fuente: Elaboración propia

	β_1 (dt)	β_2 (LN_Canjes)	LL	BIC
Nivel 1	-1,035 (0,053)	0,305 (0,097)	-180.050	360.803
Nivel 2	-1,036 (0,015)	0,286 (0,089)	-34.494	69.533
Nivel 3	-1,063 (0,007)	0,306 (0,041)	-21.802	44.125
Nivel 4	-0,997 (0,064)	0,292 (0,076)	-2.096	4.656

Tabla nº 66: Parámetros estimados según modelo no paramétrico
Fuente: Elaboración propia

Anexo B: Estimación de parámetros del modelo 2 según otras parametrizaciones del modelo hazard, por nivel de canje:

	β_1 (Dt)	γ	LL	BIC
Nivel 1	-1,340 (0,001)	3,234 (0,002)	-204.536	409.098
Nivel 2	-1,419 (0,002)	4,090 (0,012)	-39.973	79.969
Nivel 3	-1,533 (0,084)	4,700 (0,008)	-25.775	51.572
Nivel 4	-1,497 (0,002)	5,468 (0,028)	-2.524	5.066

Tabla nº 67: Parámetros estimados según distribución exponencial
Fuente: Elaboración propia

	β_1 (Dt)	γ	α	LL	BIC
Nivel 1	-0,961 (0,000)	12,562 (0,016)	0,494 (0,000)	-177.090	354.219
Nivel 2	-1,005 (0,000)	23,117 (0,207)	0,458 (0,000)	-33.695	67.424
Nivel 3	-1,058 (0,000)	37,069 (0,226)	0,429 (0,000)	-21.104	42.241
Nivel 4	-1,105 (0,003)	60,319 (0,058)	0,417 (0,000)	-2.029	4.086

Tabla nº 68: Parámetros estimados según distribución weibull
Fuente: Elaboración propia

	β_1 (Dt)	γ	α	LL	BIC
Nivel 1	-0,925 (0,001)	4,051 (0,010)	2,428 (0,003)	-178.392	356.823
Nivel 2	-0,980 (0,002)	4,769 (0,096)	2,655 (0,016)	-33.895	67.825
Nivel 3	-1,037 (0,000)	5,453 (0,024)	2,737 (0,003)	-21.215	42.463
Nivel 4	-1,100 (0,003)	5,629 (0,288)	3,184 (0,041)	-2.042	4.112

Tabla nº 69: Parámetros estimados según distribución log-logística
Fuente: Elaboración propia

	β_1 (Dt)	LL	BIC
Nivel 1	-1,117 (0,001)	-178.821	358.332
Nivel 2	-1,125 (0,003)	-34.026	68.586
Nivel 3	-1,266 (0,001)	-21.571	43.650
Nivel 4	-1,317 (0,002)	-2.075	4.604

Tabla nº 70: Parámetros estimados según modelo no paramétrico
Fuente: Elaboración propia

Anexo C: Estimación de parámetros del modelo 3 según otras parametrizaciones del modelo hazard, por nivel de canje:

	β_1 (dt)	β_2 (LN_Canjes)	β_3 (SHAREt)	β_4 (VENCIDot)	β_5 (X_VENCeT)	γ	LL	BIC
Nivel 1	-1,000 (0,043)	0,366 (0,040)	1,707 (0,959)	0,092 (0,025)	0,079 (0,023)	0,643 (0,129)	-202.037	404.152
Nivel 2	-1,007 (0,019)	0,418 (0,010)	0,976 (0,051)	0,096* (0,095)	0,086 (0,023)	1,284 (0,118)	-39.947	79.964
Nivel 3	-1,016 (0,000)	0,507 (0,000)	0,191 (0,012)	0,113 (0,000)	0,089* (0,080)	2,540 (0,014)	-25.756	51.580
Nivel 4	-1,037 (0,013)	0,404 (0,007)	0,634 (0,025)	-0,038* (0,061)	0,008 (0,034)	2,198 (0,068)	-2.542	5.140

* No son estadísticamente distintos de cero, según test t, $P > 0.1$

Tabla nº 71: Parámetros estimados según distribución exponencial

Fuente: Elaboración propia

	β_1 (dt)	β_2 (LN_Canjes)	β_3 (SHAREt)	β_4 (VENCIDot)	β_5 (X_VENCeT)	γ	α	LL	BIC
Nivel 1	-0,692 (0,002)	0,243 (0,000)	1,736 (0,001)	-0,057* (0,050)	0,065* (0,061)	0,522 (0,000)	0,498 (0,000)	- 174.763	355.783
Nivel 2	-0,689 (0,010)	0,280 (0,000)	1,261 (0,009)	-0,066* (0,049)	0,073* (0,103)	1,325 (0,025)	0,457 (0,000)	-33.490	67.961
Nivel 3	-0,670 (0,000)	0,323 (0,009)	0,841 (0,014)	0,055* (0,058)	0,072* (0,060)	3,260 (0,073)	0,428 (0,000)	-20.998	42.552
Nivel 4	-0,716 (0,001)	0,249 (0,000)	1,608 (0,026)	0,040* (0,033)	-0,022* (0,037)	0,513 (0,004)	0,444 (0,000)	-2.031	4.877

* No son estadísticamente distintos de cero, según test t, $P > 0.1$

Tabla nº 72: Parámetros estimados según distribución weibull

Fuente: Elaboración propia

	β_1 (dt)	β_2 (LN_Canjes)	β_3 (SHAREt)	β_4 (VENCIDot)	β_5 (X_VENCeT)	β_1 (dt)	β_2 (LN_Canjes)	β_3 (SHAREt)	β_4 (VENCIDot)
Nivel 1	-0,692 (0,049)	0,243 (0,045)	1,736 (0,899)	-0,057* (0,049)	0,065* (0,056)	0,522 (0,364)	0,498 (0,031)	- 174.998	350.087
Nivel 2	-0,689 (0,000)	0,280 (0,000)	1,261 (0,006)	-0,066* (0,060)	0,073 (0,001)	1,325 (0,015)	0,457 (0,000)	-33.545	67.171
Nivel 3	-0,640 (0,000)	0,293 (0,000)	1,428 (0,010)	-0,039* (0,056)	0,067 (0,000)	1,945 (0,040)	0,801 (0,001)	-21.043	42.166
Nivel 4	-0,707 (0,002)	0,241 (0,001)	1,838 (0,009)	-0,048 (0,003)	-0,024* (0,032)	1,396 (0,039)	0,705 (0,022)	-2.036	4.136

* No son estadísticamente distintos de cero, según test t, $P > 0.1$

Tabla nº 73: Parámetros estimados según distribución log-logística

Fuente: Elaboración propia

	β_1 (dt)	β_2 (LN_Canjes)	β_3 (SHAREt)	β_4 (VENCIDot)	β_5 (X_VENCet)	LL	BIC
Nivel 1	-0,909 (0,003)	0,346 (0,002)	0,162 (0,064)	0,021* (0,023)	0,028 (0,013)	-179.494	359.730
Nivel 2	-0,827 (0,010)	0,368 (0,016)	0,875 (0,104)	-0,050* (0,041)	0,030* (0,038)	-34.037	68.654
Nivel 3	-0,841 (0,007)	0,405 (0,009)	1,491 (0,085)	0,019 (0,034)	0,005 (0,029)	-21.369	43.291
Nivel 4	-0,822 (0,011)	0,365 (0,007)	1,397 (0,136)	0,009* (0,019)	0,007* (0,016)	-2.062	4.615

* No son estadísticamente distintos de cero, según test t, $P > 0.1$

Tabla nº 74: Parámetros estimados según modelo no paramétrico

Fuente: Elaboración propia

Anexo D: Estimación de parámetros del modelo 4 según otras parametrizaciones del modelo hazard, por nivel de canje:

	β_1 (Dt)	β_2 (SHAREt)	β_3 (VENCIDost)	β_4 (X_VENCERT)	γ	LL	BIC
Nivel 1	-1,186 (0,051)	2,055 (0,184)	0,064 (0,055)	0,058 (0,023)	0,562 (0,191)	-201.157	402.379
Nivel 2	-1,318 (0,000)	1,701 (0,006)	0,058 (0,046)	0,063 (0,000)	0,893 (0,004)	-39.602	79.262
Nivel 3	-1,450 (0,000)	1,366 (0,010)	0,052 (0,000)	0,063 (0,014)	1,339 (0,014)	-25.608	51.273
Nivel 4	-1,412 (0,003)	1,399 (0,129)	0,074 (0,002)	0,003 (0,002)	1,492 (0,250)	-2.843	5.732

Tabla nº 75: Parámetros estimados según distribución exponencial
Fuente: Elaboración propia

	β_1 (Dt)	β_2 (SHAREt)	β_3 (VENCIDost)	β_4 (X_VENCERT)	γ	α	LL	BIC
Nivel 1	-0,814 (0,060)	1,952 (0,183)	-0,039* (0,032)	0,053* (0,045)	0,445 (0,033)	0,500 (0,004)	-174.448	348.974
Nivel 2	-0,898 (0,000)	1,729 (0,007)	-0,038* (0,031)	0,061* (0,056)	0,800 (0,016)	0,461 (0,000)	-33.376	66.770
Nivel 3	-0,954 (0,000)	1,572 (0,011)	0,013* (0,011)	0,059* (0,056)	1,304 (0,084)	0,430 (0,000)	-20.949	41.965
Nivel 4	-0,979 (0,004)	1,938 (0,158)	-0,055* (0,044)	0,017* (0,015)	0,796 (0,026)	0,417 (0,000)	-2.016	4.087

* No son estadísticamente distintos de cero, según test t, P>0.1

Tabla nº 76: Parámetros estimados según distribución weibull
Fuente: Elaboración propia

	β_1 (Dt)	β_2 (SHAREt)	β_3 (VENCIDost)	β_4 (X_VENCERT)	γ	α	LL	BIC
Nivel 1	-0,774 (0,056)	2,234 (0,176)	0,034 (0,029)	0,052 (0,027)	0,592 (0,072)	0,726 (0,039)	-174.683	349.444
Nivel 2	-0,859 (0,000)	2,113 (0,005)	0,031 (0,015)	0,059 (0,019)	0,854 (0,011)	0,711 (0,000)	-33.424	66.918
Nivel 3	-0,909 (0,000)	1,990 (0,008)	0,005 (0,009)	0,057 (0,022)	1,210 (0,036)	0,709 (0,001)	-20.983	42.033
Nivel 4	-0,965 (0,003)	2,103 (0,101)	0,059 (0,026)	0,019 (0,002)	1,257 (0,034)	0,698 (0,010)	-2.034	4.123

Tabla nº 77: Parámetros estimados según distribución log-logística
Fuente: Elaboración propia

	β_1 (Dt)	β_2 (SHAREt)	β_3 (VENCIDost)	β_4 (X_VENCERT)	LL	BIC
Nivel 1	-1,137 (0,086)	-0,383 (0,254)	0,095 (0,017)	0,066 (0,025)	-179.616	359.961
Nivel 2	-1,112 (0,008)	-0,537 (0,014)	0,090 (0,001)	0,066 (0,001)	-33.954	68.475
Nivel 3	-1,200 (0,010)	-0,839 (0,004)	0,065 (0,009)	0,072 (0,013)	-21.328	43.199
Nivel 4	-1,244 (0,059)	-1,088 (0,026)	0,0171 (0,013)	0,0034 (0,003)	-2.058	4.597

Tabla nº 78: Parámetros estimados según modelo no paramétrico
Fuente: Elaboración propia

Anexo E: Impacto de la distancia en la probabilidad de compra, según los distintos modelos usados en la hipótesis 1

	$P(D_t=1)$	$P(D_t=0)$	Incremento
Nivel 1	0,78	0,96	24,1%
Nivel 2	0,85	0,98	15,9%
Nivel 3	0,88	0,99	11,9%
Nivel 4	0,92	0,99	8,1%

Tabla nº 79: Influencia de la distancia sobre la probabilidad, modelo 1

Fuente: Elaboración propia

	$P(D_t=1)$	$P(D_t=0)$	Incremento
Nivel 1	0,74	0,97	31,6%
Nivel 2	0,79	0,99	25,3%
Nivel 3	0,80	0,99	23,1%
Nivel 4	0,84	1,00	18,6%

Tabla nº 80: Influencia de la distancia sobre la probabilidad, modelo 2

Fuente: Elaboración propia

	$P(D_t=1)$	$P(D_t=0)$	Incremento
Nivel 1	0,80	0,96	20,2%
Nivel 2	0,86	0,98	14,3%
Nivel 3	0,89	0,99	11,0%
Nivel 4	0,92	0,99	7,9%

Tabla nº 81: Influencia de la distancia sobre la probabilidad, modelo 3

Fuente: Elaboración propia

Anexo F: Estimación de parámetros del modelo de H3 según otras parametrizaciones del modelo hazard, por nivel de canje:

	β_1 (dt)	β_2 (SHAREt)	β_3 (VENCIDOST)	β_4 (X_VENCERT)	γ	LL	BIC
Canje 1	-0,935 (0,039)	3,554 (0,265)	-0,0109 (0,033)	-0,073 (0,030)	0,1583 (0,034)	-6.145	12.318
Canje 2	-0,471 (0,04)	4,172 (0,063)	-0,1228 (0,028)	0,087 (0,031)	0,0672 (0,214)	-2.395	4.818
Canje 3	-0,154 (0,059)	7,416 (0,054)	-0,019* (0,077)	0,091 (0,050)	0,0028 (0,000)	-1.331	2.689
Canje 4	-0,238 (0,067)	5,927 (0,076)	0,0463* (0,081)	-0,003* (0,055)	0,0122 (0,028)	-1.565	3.157

* No son estadísticamente distintos de cero, según test t, $P > 0.1$

Tabla nº 82: Parámetros estimados según distribución exponencial

Fuente: Elaboración propia

	β_1 (dt)	β_2 (SHAREt)	β_3 (VENCIDOST)	β_4 (X_VENCERT)	γ	α	LL	BIC
Canje 1	-0,759 (0,041)	3,438 (0,827)	0,001* (0,042)	-0,031* (0,037)	1,0929 (0,126)	0,6596 (0,067)	-5.629	11.297
Canje 2	-0,450 (0,058)	3,691 (0,178)	0,0688* (0,051)	-0,084 (0,045)	10,0243 (0,01)	0,5613 (0,012)	-2.171	4.378
Canje 3	-0,132 (0,061)	3,983 (0,041)	0,009* (0,080)	0,062* (0,051)	1,0087 (0,003)	0,5277 (0,028)	-1.227	2.490
Canje 4	-0,223 (0,061)	4,172 (0,087)	0,039* (0,076)	0,094 (0,050)	1,0046 (0,001)	0,4898 (0,022)	-1.379	2.793

* No son estadísticamente distintos de cero, según test t, $P > 0.1$

Tabla nº 83: Parámetros estimados según distribución log-logística

Fuente: Elaboración propia

	β_1 (dt)	β_2 (SHAREt)	β_3 (VENCIDOST)	β_4 (X_VENCERT)	LL	BIC
Canje 1	-0,887 (0,051)	0,521 (0,608)	0,011* (0,047)	-0,022 (0,042)	-5.741	11.665
Canje 2	-0,416 (0,071)	4,570 (0,577)	0,068* (0,056)	0,088* (0,057)	-2.191	4.555
Canje 3	-0,153 (0,069)	0,368 (0,102)	0,043* (0,087)	0,063* (0,055)	-1.241	2.606
Canje 4	-0,247 (0,064)	2,216 (0,083)	0,015* (0,086)	0,067* (0,063)	-1.453	3.030

* No son estadísticamente distintos de cero, según test t, $P > 0.1$

Tabla nº 84: Parámetros estimados según modelo no paramétrico

Fuente: Elaboración propia

Anexo G: Intervalos de confianza para parámetro beta1

Los intervalos de confianza para el estimador beta1 se calculan aproximándolo con una distribución normal en forma estándar. Los intervalos se calculan para un nivel de confianza $1 - \alpha = 95\%$, como se muestra a continuación:

$$P(a \leq \beta \leq b) = P(-b \leq -\beta \leq -a) = P\left(\frac{\hat{\beta} - b}{\hat{\sigma}} \leq \frac{\hat{\beta} - \beta}{\hat{\sigma}} \leq \frac{\hat{\beta} - a}{\hat{\sigma}}\right) = 1 - \alpha = 95\%$$

Considerando que $\frac{\hat{\beta} - \beta}{\hat{\sigma}} \rightarrow N(0,1)$ entonces $a = \hat{\beta} - 1,96\hat{\sigma}$ y $b = \hat{\beta} + 1,96\hat{\sigma}$. Dado lo anterior, en la tabla n°X se muestran los intervalos de confianza para los parámetros beta1 del modelo.

	Intervalos de conf. Parámetro beta1
Canje 1	[-0,85 ; -0,68]
Canje 2	[-0,58 ; -0,32]
Canje 3	[-0,26 ; -0,02]
Canje 4	[-0,33 ; -0,08]

Tabla n° 85: Intervalos de confianza para el parámetro beta1

Fuente: Elaboración propia.