

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

**EVALUACIÓN DEL IMPACTO DE UNA PROMOCIÓN DE CANJE DE
PUNTOS EN UN CLUB DE FIDELIZACIÓN DE UNA EMPRESA DE
RETAIL**

MEMORIA PARA OPTAR AL TÍTULO DE INGENIERO CIVIL INDUSTRIAL

MARLENNE DANIELA RIQUELME TORRES

PROFESOR GUÍA:
RICARDO MONTOYA MOREIRA

MIEMBROS DE LA COMISIÓN:
ALEJANDRA PUENTE CHANDÍA
MARÍA CAROLINA OLGUÍN QUIJADA

SANTIAGO DE CHILE
OCTUBRE 2009

**RESUMEN DE LA MEMORIA
PARA OPTAR AL TÍTULO DE
INGENIERO CIVIL INDUSTRIAL
POR: MARLENNE RIQUELME TORRES
FECHA: 26/10/2009
PROF. GUÍA: SR. RICARDO MONTOYA M.**

EVALUACIÓN DEL IMPACTO DE UNA PROMOCIÓN DE CANJE DE PUNTOS EN UN CLUB DE FIDELIZACIÓN DE UNA EMPRESA DE RETAIL

Debido a la gran competitividad en la industria del retail, las empresas se han visto en la necesidad de desarrollar estrategias de marketing que beneficien a sus clientes y generen lealtad. Una de esas estrategias son los clubes de fidelización, los que a través de incentivos buscan concentrar las compras de los clientes en una sola empresa.

El objetivo de esta memoria es evaluar el impacto de una promoción de canje de productos realizada por el club de fidelización de una tienda por departamento, la cual se desarrolló en dos etapas.

Para estimar el efecto de la promoción se usó el modelo de series de tiempo $Sarimax(1,0,2)(1,0,0)$, incorporando variables como los puntos entregados en períodos anteriores y variables dummies de fin de mes, navidad, entre otros. El modelo mostró en el primer período un incremento en el número de canjes, observándose posterior a la ésta un ajuste en los canjes de acuerdo a lo estimado. Durante el segundo período, no se observó el mismo efecto, los canjes se ajustaron a los de una semana normal, lo que se asocia a la incorporación permanente de la promoción, disminuyendo la urgencia de los clientes por canjear, comportamiento esperado en el largo plazo.

Posteriormente, se utilizó el modelo proporcional de Hazard para pronosticar compra. Se diseñó la variable distancia, la cual determinó que a medida que el cliente se encuentra más cerca de un nivel de canje, mayor es la probabilidad de comprar, lo que es un claro indicio de aceleración de compra debido a la estructura del club. A través de una variable de promoción, se determinó que la ésta disminuyó la probabilidad de compra en los cuatro niveles más bajos (que acumulan más del 97% de los canjes).

Usando un modelo logit se encontró que a mayor cercanía al nivel de canje y menor frecuencia, mayor es la probabilidad de canjear, probabilidad que luego se usó para identificar el porcentaje de clientes que usó la promoción, pero que no lo hubiese realizado en circunstancias de canje regular, determinándose que menos del 50% de los clientes hubiesen canjeado en el período.

Para concluir, se realizó un análisis comparativo de los clientes que canjearon durante período promocional, donde se determinó que los clientes que usaron la promoción compran en menor frecuencia y montos más bajos que aquellos que no la usaron. Además, analizando el recency promedio de compra se observó que antes de la promoción, los clientes que canjearon en promoción tenían un recency mayor que el de los clientes que no la usaron, mientras que en las semanas posteriores a la ésta su recency se mantuvo más bajo, lo que puede ser un claro efecto de la promoción, que incentivó a los clientes a comprar más seguido.

Luego, si se considera que con el monto cancelado por el cliente la empresa es capaz de cubrir parte o el total del costo del producto, y que se observó un efecto post-promoción positivo en la compra en los clientes que la usaron, se sugiere a la empresa que continúe realizándola, pues es una forma de recompensar a sus clientes y reducir costos.

Índice

1.	Introducción	1
1.1	Antecedentes Generales	1
1.1.1	CMR Puntos	1
1.1.2	Círculo más	2
1.1.3	RipleyPuntos	4
1.2	CRM.....	5
2.	Justificación y Descripción del Proyecto	6
3.	Objetivos	7
3.1	Objetivo general.....	7
3.2	Objetivos específicos	7
4.	Alcances	8
5.	Resultados esperados	8
6.	Metodología	8
6.1	Estudio de la situación actual	8
6.2	Selección y procesamiento de datos	9
6.3	Transformación de los datos.....	9
6.4	Estimación de número de canjes	9
6.5	Modelo Proporcional de Hazard.....	9
6.6	Modelos de predicción de canjes.....	9
6.7	Análisis descriptivo de los clientes que usaron la promoción	9
7.	Marco conceptual.....	10
7.1	Modelos de series de tiempo	10
7.1.1	Definiciones básicas.....	10
7.1.2	Modelos Arima	12
7.1.3	Modelos Sarimax.....	12
7.1.4	Medidas de error	13
7.2	Modelo Proporcional de Hazard.....	14
7.2.1	Línea base.....	16
7.2.2	Estimación de los parámetros	16
7.2.3	Evaluación de la calidad predictiva	17
7.3	Hipótesis del gradiente – objetivo	18
8.	Desarrollo metodológico	19
8.1	Estudio de la situación actual	19
8.1.1	Programa de fidelización de la empresa	19

8.1.2	Promoción a evaluar	20
8.2	Selección y procesamiento de los datos	21
8.3	Transformación de los datos	22
8.3.1	Variabes RFM	23
8.3.2	Distancia psicológica al objetivo	23
8.3.3	Variabes Dummies	24
8.4	Estimación de número de canje	25
8.4.1	Estimación número de canjes globales	25
8.4.2	Estimación número de canjes para el nivel 1	40
8.4.3	Conclusiones de las estimaciones del número de canjes por niveles	44
8.5	Modelo proporcional de Hazard:	45
8.5.1	Estimación de los parámetros del modelo para el nivel 1	45
8.5.2	Ajuste de modelo para el nivel 1	47
8.5.3	Pronóstico del modelo proporcional de Hazard para el nivel 1	48
8.5.4	Conclusiones sobre el modelo en los niveles	49
8.6	Modelos de predicción de canje	49
8.6.1	Estimación de los parámetros para el nivel 1	50
8.6.2	Ajuste del modelo para el nivel 1	52
8.6.3	Pronóstico del modelo de predicción de canjes	53
8.7	Análisis descriptivo de los clientes que usaron la promoción	55
8.7.1	Sexo	56
8.7.2	Edad	57
8.7.3	Recency de compra	59
8.7.4	Frequency de compra	60
8.7.5	Monto de compra	61
8.7.6	Recency de canje	62
8.7.7	Frequency de canje	63
8.7.8	Monto de canje	64
9.	Conclusiones	65
9.1	Conclusiones sobre la estimación de demanda	65
9.2	Conclusiones sobre el modelo proporcional de Hazard	66
9.3	Conclusiones sobre el modelo de predicción de canje	67
9.4	Conclusiones generales	67
9.5	Limitaciones de los modelos	68
9.6	Trabajos futuros	68
10.	Bibliografía	69

11. Anexos.....	71
11.1 Anexo 1: Cálculo del tamaño muestral	71
11.2 Anexo 2: Número de canjes semanales por mes	72
11.3 Anexo 3: Estimación de canje por niveles	75
11.4 Anexo 4: Modelo proporcional de Hazard.....	79
11.4.1 Nivel 2	79
11.4.2 Nivel 3	81
11.4.3 Nivel 4	83
11.4.4 Nivel 5	85
11.4.5 Nivel 6	87
11.4.6 Nivel 7	88
11.4.7 Nivel 8	90
11.5 Anexo 5: Modelo de predicción de canjes	91
11.6 Anexo 6: Porcentaje de canjes y no canjes predichos para clientes que usaron la promoción en período 1	93

Índice de Gráficos y Tablas

Gráfico 1: Participación de canjes por nivel	20
Gráfico 2: Número de canjes totales semanales.....	25
Gráfico 3: Autocorrelaciones de la serie Número de Canjes Totales.....	27
Gráfico 4: Autocorrelaciones Parciales de la serie Número de Canjes Totales	27
Gráfico 5: Periodograma Serie número de canjes totales por frecuencia	28
Gráfico 6: Número de canjes semanales, mes de enero	29
Gráfico 7: Número de canjes semanales, mes de marzo.....	29
Gráfico 8: Número de canjes semanales, mes de septiembre.....	30
Gráfico 9: Número de canjes semanales, mes de diciembre	30
Gráfico 10: Correlaciones cruzadas Numero de canjes y Puntos acumulados.....	31
Gráfico 11: Correlaciones cruzadas Numero de canjes y Puntos entregados	31
Gráfico 12: Correlaciones cruzadas Numero de canjes y Puntos vencidos	32
Gráfico 13: Ajuste del Modelo 1 en el período de calibración	36
Gráfico 14: Ajuste del Modelo 2 en el período de calibración	37
Gráfico 15: Ajuste del modelo 1 en el período de pronóstico.....	38
Gráfico 16: Ajuste del modelo 2 en el período de pronóstico.....	39
Gráfico 17: Ajuste del Modelo 1 para el nivel 1	42
Gráfico 18: Ajuste del Modelo 2 para el nivel 1	42
Gráfico 19: Ajuste de los Modelos 1 y 2 para el nivel 1	43
Gráfico 20: Proporción de clientes que usaron la promoción en el período 1 por rango de edad.....	57
Gráfico 21: Proporción de clientes que no usaron la promoción en el período 1 por rango de edad.....	57
Gráfico 22: Proporción de clientes que usaron la promoción en el período 2 por rango de edad.....	58
Gráfico 23: Proporción de clientes que no usaron la promoción en el período 1 por rango de edad.....	58
Gráfico 24: Recency de compra promedio de clientes canjeadores del período 1	59
Gráfico 25: Recency de compra promedio de clientes canjeadores del período 2	59
Gráfico 26: Frecuencia de compra promedio de clientes canjeadores del período 1.....	60
Gráfico 27: Frecuencia de compra promedio de clientes canjeadores del período 2.....	60
Gráfico 28: Monto de compra promedio de clientes canjeadores del período 1	61
Gráfico 29: Monto de compra promedio de clientes canjeadores del período 2	61
Gráfico 30: Recency de canje promedio de clientes canjeadores del período 1	62
Gráfico 31: Recency de canje promedio de clientes canjeadores del período 2.....	62
Gráfico 32: Frecuencia de canje promedio de clientes canjeadores del período 1	63
Gráfico 33: Frecuencia de canje promedio de clientes canjeadores del período 2	63
Gráfico 34: Monto de canje promedio de clientes canjeadores del período 1	64
Gráfico 35: Monto de canje promedio de clientes canjeadores del período 1	64
Gráfico 36: Número de canjes semanales, mes de febrero	72
Gráfico 37: Número de canjes semanales, mes de abril.....	72
Gráfico 38: Número de canjes semanales, mes de mayo.....	73
Gráfico 39: Número de canjes semanales, mes de junio	73
Gráfico 40: Número de canjes semanales, mes de julio	73
Gráfico 41: Número de canjes semanales, mes de agosto.....	74
Gráfico 42: Número de canjes semanales, mes de octubre.....	74

Gráfico 43: Número de canjes semanales, mes de noviembre.....	74
Tabla 1: Modalidades de canje Círculo Más	3
Tabla 2: Estructura de acumulación de puntos Círculo Más	3
Tabla 3: Estructura de acumulación de RipleyPuntos.....	4
Tabla 4: Líneas bases paramétricas	16
Tabla 5: Matriz de confusión	17
Tabla 6: Estructura de acumulación de puntos del programa.....	19
Tabla 7: Estructura de Niveles de la empresa	19
Tabla 8: Estructura de Niveles promoción Puntos con Pesos.....	21
Tabla 9: Test Dickey Fuller para la serie número de canjes globales	26
Tabla 10: Parámetros del Modelo 1 de estimación del número de canjes	33
Tabla 11: Parámetros del Modelo 2 de estimación del número de canjes	35
Tabla 12: Ajuste del Modelo 1 en el período de calibración.....	36
Tabla 13: Ajuste del Modelo 2 en el período de calibración.....	37
Tabla 14: Medidas de error Modelo 1 en el período de pronóstico	38
Tabla 15: Medidas de error Modelo 2 en el período de pronóstico	39
Tabla 16: Parámetros del Modelo 1 y 2 para el nivel 1	41
Tabla 17: Ajuste de los Modelos 1 y 2 durante el período de calibración	41
Tabla 18: Medidas de error Estimación número de canje Nivel 1	43
Tabla 19: Parámetros del modelo proporcional de Hazard para el nivel 1	46
Tabla 20: Líneas base del nivel 1	46
Tabla 21: Ajuste del modelo proporcional de Hazard para el nivel 1	47
Tabla 22: Indicadores matriz de confusión (base de calibración)	48
Tabla 23: Indicadores matriz de confusión (base de pronóstico)	48
Tabla 24: Parámetros del modelo de predicción de canjes.....	51
Tabla 25: Ajuste del modelo predicción de canje para el nivel 1.....	52
Tabla 26: Indicadores matriz de confusión (base de calibración)	53
Tabla 27: N° de canjes y no canjes de la base	53
Tabla 28: Indicadores matriz de confusión en (base de pronóstico)	53
Tabla 29: Porcentaje de canjes y no canjes predichos para clientes que usaron la promoción en período 1	54
Tabla 30: Porcentaje de canjes y no canjes predichos para clientes que usaron la promoción en período 1 sin incorporar variables promocionales.	54
Tabla 31: Canjes promoción Puntos con pesos por nivel, primer período	55
Tabla 32: Canjes promoción Puntos con pesos por nivel, segundo período.....	56
Tabla 33: Porcentaje de hombres y mujeres que canjearon en período promocional	56
Tabla 34: Tamaño muestral por niveles	71
Tabla 35: Tamaño de las muestras utilizadas por niveles.....	72
Tabla 36: Resumen del efecto de la primera Promoción por nivel.....	75
Tabla 37: Parámetros por nivel del Modelo 1.....	76
Tabla 38: Parámetros por nivel del Modelo 2.....	77
Tabla 39: Ajuste por nivel de los Modelos 1 y 2.....	78
Tabla 40: Parámetros del modelo proporcional de Hazard para el nivel 2.....	79
Tabla 41: Ajuste del modelo proporcional de Hazard para el nivel 2	79
Tabla 42: Indicadores matriz de confusión (base de calibración)	80
Tabla 43: Indicadores matriz de confusión (base de pronóstico)	80
Tabla 44: Parámetros del modelo proporcional de Hazard para el nivel 3.....	81

Tabla 45: Ajuste del modelo proporcional de Hazard para el nivel 3	81
Tabla 46: Indicadores matriz de confusión (base de calibración)	82
Tabla 47: Indicadores matriz de confusión (base de pronóstico)	82
Tabla 48: Parámetros del modelo proporcional de Hazard para el nivel 4.....	83
Tabla 49: Ajuste del modelo proporcional de Hazard para el nivel 4	83
Tabla 50: Indicadores matriz de confusión (base de calibración)	84
Tabla 51: Indicadores matriz de confusión (base de pronóstico)	84
Tabla 52: Parámetros del modelo proporcional de Hazard para el nivel 5.....	85
Tabla 53: Ajuste del modelo proporcional de Hazard para el nivel 5	85
Tabla 54: Indicadores matriz de confusión (base de calibración)	86
Tabla 55: Indicadores matriz de confusión (base de pronóstico)	86
Tabla 56: Parámetros del modelo proporcional de Hazard para el nivel 6.....	87
Tabla 57: Ajuste del modelo proporcional de Hazard para el nivel 6	87
Tabla 58: Indicadores matriz de confusión (base de calibración)	87
Tabla 59: Indicadores matriz de confusión (base de pronóstico)	88
Tabla 60: Parámetros del modelo proporcional de Hazard para el nivel 7.....	88
Tabla 61: Ajuste del modelo proporcional de Hazard para el nivel 7	88
Tabla 62: Indicadores matriz de confusión (base de calibración)	89
Tabla 63: Indicadores matriz de confusión (base de pronóstico)	89
Tabla 64: Parámetros del modelo proporcional de Hazard para el nivel 8.....	90
Tabla 65: Ajuste del modelo proporcional de Hazard para el nivel 8	90
Tabla 66: Indicadores Indicadores matriz de confusión (base de calibración)	90
Tabla 67: Indicadores matriz de confusión (base de pronóstico)	91
Tabla 68: Parámetros del modelo de predicción de canje por nivel.....	91
Tabla 69: Ajuste del modelo de predicción de canje por nivel	92
Tabla 70: Indicadores matriz de confusión (base de calibración)	92
Tabla 71: Indicadores matriz de confusión (base de pronóstico)	92
Tabla 72: Porcentaje de canjes y no canjes predichos para clientes que usaron la promoción en período 2.	93
Tabla 73: Porcentaje de canjes y no canjes predichos para clientes que canjearon usando la promoción en período 2 sin incorporar variables promocionales.	93

1. Introducción

1.1 Antecedentes Generales

La estabilidad económica que ha presentado el país durante las últimas décadas y la gran competitividad de la industria del retail han permitido que las ventas del sector durante los últimos años hayan tenido un crecimiento promedio anual del 12%, un crecimiento mayor al experimentado por la economía en conjunto¹.

Dentro del sector, uno de los canales que ha experimentado un incremento importante en las ventas han sido las tiendas por departamento, las que surgen de la necesidad de los clientes de cubrir gran parte de sus necesidades en un mismo lugar. Actualmente en Chile, cuatro empresas han asumido un rol protagónico en las tiendas por departamento: Falabella, con una participación de mercado del 37,3%; Ripley, con un 25,4 %; Paris, con el 24,9% y La Polar, con un 12,4%².

La gran competitividad, las altas rotaciones y los bajos márgenes asociados a la industria han llevado a estos actores a una constante búsqueda por diferenciarse de sus competidores y generar lealtad en sus clientes, razón por la cual han diseñado distintos programas de fidelización, cuyo objetivo es premiar a los clientes más leales, entregando descuentos y/o puntos intercambiables por premios, incrementando en los clientes el costo de cambiarse a otro competidor.

Actualmente, los programas de fidelización se han transformado en una de las herramientas de marketing usadas más frecuentemente para estimular el uso de productos y/o servicios ofrecidos por el retail y para retener a los clientes, pues la evidencia indica que adquirir nuevos clientes resulta más costoso para las empresas que retener a los actuales [8].

Los programas de fidelización, de acuerdo a la industria donde se encuentren insertos, pueden presentar similitudes y diferencias, a continuación se presenta una breve descripción de los clubes desarrollados por las tiendas por departamento más importantes del país:

1.1.1 CMR Puntos

En el caso de Falabella, existe un programa llamado CMR Puntos, que entró en vigencia el 1 de marzo de 2003, donde los clientes socios del club pueden acumular puntos que le permitirán canjear productos en las tiendas Falabella y sus alianzas.

La clave de este programa es la tarjeta CMR, la que permite a sus clientes realizar compras en cuotas, lo que facilita el acceso a productos que para algunos grupos

¹ Fuente: Reporte de Mercado, Colliers International

² Fuente: Presentación SACI Falabella, Abril 2007

socioeconómicos son difíciles de adquirir al contado y además acumular puntos a través de compras.

El sistema de acumulación es diferente dependiendo del tipo de transacción realizado por el cliente, así, por cada \$120 de compra de productos o servicios en que el medio de pago utilizado sea la tarjeta CMR se acumulará 1 punto. También se acumula 1 punto por cada \$950 de Avance y Súper Avance en efectivo contra el crédito CMR del socio participante. Por último, si el socio solicita la emisión de facturas como respaldo de la compra del producto o servicio, se acumulará 1 punto cada \$360 de dichas compras.

Los puntos se acumulan en la cuenta del cliente y son intransferibles. La vigencia de los puntos es de 12 meses, es decir, los puntos no canjeados al finalizar dicho período son descontados de la cuenta.

Para canjear, los clientes requieren acumular una cantidad de puntos mínima, la cual depende del nivel en el cual deseen canjear. Los premios a los cuales pueden acceder los socios participantes se detallan en un catálogo que la empresa edita semestralmente, el que se encuentra disponible en las tiendas Falabella, Homecenter Sodimac y en el sitio Web de CMR Falabella.

1.1.2 Círculo más

El Programa Círculo Más es organizado por Cencosud S.A. a través del cual cualquier persona mayor de 18 años puede acumular puntos en una cuenta única por todas las compras realizadas a través de cualquier medio de pago (tarjetas más, efectivo, cheque, Redbanc y tarjetas bancarias) en Jumbo, París, Easy, Santa Isabel, Viajes Paris, Entel PCS y comercios asociados, y con ellos canjear los premios que se encuentran disponibles en el catálogo Círculo Más.

Cada compra efectuada por un cliente en alguna de las empresas nombradas anteriormente otorga a éste una cantidad de puntos que ingresarán a la cuenta. Si la compra es realizada por un tercero, éste tendrá la posibilidad de acumular puntos para el titular de la cuenta mencionando el Rut de éste.

Si la compra es realizada con factura en las empresas asociadas al programa, la cantidad de puntos asociados estará en directa relación a alguno o algunos de los siguientes factores: producto adquirido, monto de la venta, medio de pago utilizado, comercio donde se efectúa la transacción y fecha de la transacción; factores que serán determinados libremente por la empresa.

Los puntos que obtienen los clientes por concepto de las compras realizadas son canjeables por bienes y servicios que son agrupados de acuerdo a niveles de puntuación. Estos pueden ser retirados en los locales habilitados para tales efectos o solicitados mediante sistemas de comunicación remota. En particular, los puntos pueden ser canjeados a través de las siguientes modalidades:

Tabla 1: Modalidades de canje Círculo Más

Modalidad	Tipo de canje
Catálogo Círculo Más	Productos insertos en el catálogo
Viajes París	Viajes o paquetes turísticos
Lan Pass	Kilómetros Lan Pass
Entel PCS	Kits de prepago o recargas de minutos

El sistema de acumulación, por tratarse de un club de fidelización al que pertenecen empresas de distintos rubros, varía según las compras realizadas en una u otra empresa. La Tabla 2 explica la situación:

Tabla 2: Estructura de acumulación de puntos Círculo Más

Por cada \$200 de compras	Puntos pagando sin Tarjeta Más	Puntos pagando con Tarjeta Más
París	1	2
Jumbo	1	3
Easy	1	2
Santa Isabel	1	2
Entel	2	3

Los puntos utilizados por el cliente para canjear un bien o servicio ofrecido por la empresa son descontados de su cuenta personal. El cliente tiene la posibilidad de transferir en forma irrevocable sus puntos a otra cuenta, siempre y cuando cumpla con los procedimientos establecidos.

En caso que Círculo Más acepte la devolución de un bien o servicio canjeado con cargo a la totalidad o parte de sus puntos acumulados, éstos son reintegrados. Si la devolución es de una compra que generó puntos, estos son descontados del total acumulado.

Los puntos obtenidos en cada compra tienen una vigencia de 12 meses, contando desde la fecha de la compra que los origina. Los puntos no canjeados al finalizar dicho periodo caducan y son descontados de la cuenta.

1.1.3 RipleyPuntos

RipleyPuntos es un programa exclusivo para clientes de la tarjeta Ripley y se encuentra vigente desde el 1 de enero de 2008. El objetivo de este programa es premiar a los clientes que realizan compras en Ripley, comercios asociados y seguros o Súper Avance, acumulando en su cuenta RipleyPuntos que pueden ser canjeados por productos de toda la tienda.

El sistema de acumulación es el siguiente:

Tabla 3: Estructura de acumulación de RipleyPuntos

Montro de compra	RipleyPuntos acumulados
\$100 de compras en la tienda, seguros y comercios asociados.	1
\$200 de avances en efectivo.	1

El valor sobre el cual se calculan los RipleyPuntos es el precio contado sin intereses e IVA incluido. Las compras realizadas bajo la modalidad cuota también acumulan puntos, sin embargo, si el pago es un tercio y dos cuotas, sólo se acumulará el monto cargado en la tarjeta. Las compras realizadas por las tarjetas adicionales también acumulan puntos en la tarjeta del titular, no así los avances en efectivo.

La vigencia de los puntos es de 12 meses desde la fecha de transacción para clientes normales. Los clientes Premier RipleyPuntos pueden aumentar la vigencia de sus puntos a 24 meses, realizando compras con su tarjeta durante el año.

Los canjes pueden ser realizados por titulares y adicionales de la tarjeta, que se encuentren con cuentas al día y tengan un mínimo de 3.000 puntos. Al canjear, cada RipleyPunto corresponde a \$1 (de moneda de curso legal), pudiendo utilizarse el monto total o parcial del puntaje acumulado para un producto de menor valor, o combinarse con pago a través de la tarjeta Ripley para productos con mayor valor.

Al devolver productos canjeados con RipleyPuntos, estos se abonan nuevamente a la cuenta, sin embargo, si lo que se devuelve es un producto que acumuló RipleyPuntos, estos serán descontados. De esta forma, si se realizó un canje previo a la devolución, los RipleyPuntos acumulados pueden llegar a ser negativos, por lo que serán descontados de la acumulación en compras posteriores.

Al comparar los tres programas de fidelización asociados a las principales tiendas por departamento, es posible observar que ellos presentan diferencias tanto en la forma de acumular puntos como en la de realizar el canje de premios, por lo que las diferencias de diseño entre cada uno de estos programas resulta una importante herramienta de diferenciación entre una empresa y otra. Al existir un entorno competitivo en esta materia, los directos favorecidos son los clientes, pues el diseño de

los programas está enfocado en entregarles los más atractivos beneficios con el objetivo de conseguir lealtad por parte de ellos.

1.2 CRM

La administración de la relación con los clientes o CRM, forma parte de una estrategia de negocio centrada en el cliente que busca aumentar el valor del cliente para una empresa.

Recopilar la mayor cantidad de información posible sobre los clientes y analizar las interacciones que tienen éstos con la empresa, son partes fundamentales del CRM, pues a través de ellas pueden comprender el comportamiento del cliente y desarrollar estrategias enfocadas a sus necesidades, lo que incrementa la lealtad de los clientes y aumenta los costos de cambiarse a otra empresa.

Actualmente, con el uso de las tarjetas de crédito y la inscripción en los programas de fidelización, las empresas cuentan con importante información transaccional y demográfica de los clientes, lo que les da la posibilidad de diseñar estrategias de CRM con el objetivo de prolongar el ciclo de vida e incrementar la lealtad de estos [7].

De acuerdo a la etapa del ciclo de vida en que se encuentre el cliente, la estrategia CRM será distinta, distinguiéndose tres etapas [7]:

- **Adquisición:** el objetivo de esta etapa es obtener nuevos clientes. La firma invierte recursos a través de distintos canales para lograr generar una relación con el cliente, la que se concreta luego de la primera interacción del cliente con la empresa.
- **Desarrollo de la relación:** el objetivo de esta etapa es aumentar los ingresos de los clientes. Una vez generada la relación es posible desarrollarla ofreciendo a los clientes productos o servicios que se ajusten a las necesidades de éstos en los distintos períodos de su ciclo de vida y por el canal más apropiado para ellos. Esto permite a la empresa incrementar sus ventas y aumentar la participación de la firma en la cartera de los clientes.
- **Retención del cliente:** el objetivo de esta etapa es prolongar el ciclo de vida y la rentabilidad de los clientes a lo largo de éste. Retener a los clientes implica entender los factores que influyen en la decisión de fuga de los clientes, ser capaces de predecirla, generando estrategias que concentren esfuerzos para prolongar la relación con el cliente.

El diseño de programas de fidelización en base a estrategias CRM, son usados cada vez más por una gran variedad de empresas, las cuales buscan a través de ellos desarrollar relaciones, estimular el uso de productos y servicios, y retener a los clientes, sin embargo, la efectividad de estos programas aún no es clara, pues diversos estudios realizados han mostrado efectos distintos.

El presente trabajo se basa en el estudio de una promoción realizada por un programa de fidelización, el cual se basa en información transaccional y demográfica de los clientes. En el siguiente capítulo se presenta con más detalle el problema a abordar y su respectiva justificación.

2. Justificación y Descripción del Proyecto

La industria del retail en Chile durante las últimas décadas ha experimentado un enorme crecimiento, generándose una gran competencia entre las empresas del rubro. Debido a esto, las empresas se encuentran permanentemente buscando formas de diferenciarse de sus competidores, y en esta búsqueda se han dado cuenta que para lograrlo necesitan entender al cliente, comprendiendo sus necesidades, ocasiones de compra, entre otros.

Actualmente, a través del uso de tarjetas de crédito y la inscripción a clubes de fidelización, las empresas de retail cuentan con información de los clientes y sus historiales de compra, lo que les ha permitido caracterizarlos en base a sus patrones de consumo y características demográficas, y a su vez, cuantificar el efecto de sus distintas acciones de marketing, logrando determinar el impacto de ellas en sus ventas.

Por esta razón, surge como una propuesta para diferenciarse comprender el efecto que tienen en los clientes las distintas acciones de marketing: catálogos, promociones, descuentos en precios, cupones, etc., así, además de ofrecerles productos de acuerdo a sus necesidades, podrán hacerlo a través de la acción de marketing que tenga un mayor impacto para ellos.

De esta manera, el trabajo consistirá en la evaluación de una promoción de canje de productos desarrollada por el programa de fidelización de una empresa de retail. Para ello, se realizará una estimación de demanda que sirva de línea base para determinar el impacto de la promoción en unidades de productos canjeados totales y por nivel, y que además permita determinar si la promoción privilegió el canje de un nivel en particular en desmedro de otros.

Una vez determinado este impacto, se aplicarán modelos econométricos para determinar si la promoción tuvo un efecto de anticipación de compra en los clientes, quienes para acceder al nivel mínimo de puntos requeridos por la promoción debieron adelantar compra en tiempo o en montos gastados. Este modelo además permitirá identificar qué variables influyen en la decisión de compra de los clientes en un determinado período.

A través de modelos de predicción de canje, se espera identificar entre los clientes que canjearon durante el período de promoción a aquellos clientes que:

- Canjearon durante el período usando la promoción, pero que no lo hubiesen hecho en situaciones de canje regulares (sin promoción) en el mismo período.
- Canjearon durante el período usando la promoción y que lo hubiesen hecho en situaciones de canje regulares.

A partir de lo anterior, se identificarán características comunes en aquellos clientes que usaron la promoción, que permitan generar perfiles en base a sus patrones de compra y canje y a sus características demográficas.

3. Objetivos

3.1 Objetivo general

Evaluar el impacto de una promoción de canje de puntos en el club de fidelización de una empresa de retail.

3.2 Objetivos específicos

- Estimar el impacto a nivel agregado de la promoción en el número de productos canjeados globales y por nivel.
- Determinar las variables que influyen en la decisión de canje de los clientes y aplicarlas a través de modelos econométricos para estimar probabilidades de canje discretas.
- Determinar variables que influyen en la decisión de compra de los clientes.
- Analizar si la promoción generó un efecto acelerador de compra.
- Identificar características comunes entre los clientes y generar perfiles que permitan predecir, de acuerdo a patrones de compra y canje, su comportamiento ante la promoción.
- Determinar el efecto de la promoción a corto y largo plazo.

4. Alcances

- Se trabajará sólo con información transaccional y demográfica disponible en la empresa, no se contempla realizar ningún tipo de estudio de mercado.
- Para el desarrollo metodológico se utilizará una muestra de clientes.
- Los modelos econométricos de predicción de compra y canje se estimarán para cada uno de los niveles de canje del programa de fidelización de la empresa.

5. Resultados esperados

Se espera entregar:

- Un modelo de estimación de demanda que permita determinar si la promoción generó un incremento en el número de canjes globales y por niveles.
- Una estimación de probabilidad de compra a través del cual se pueda identificar las variables que influyen en la decisión de compra de los clientes y si la promoción generó un efecto acelerador de compra en sus clientes.
- Un modelo de probabilidad de canje que permita identificar si la promoción generó en los clientes un efecto de aceleración de compra durante el período de la promoción.
- Perfiles de clientes en base a información transaccional y variables demográficas .
- Análisis a corto y largo plazo del impacto de la promoción.

6. Metodología

A continuación se expone la metodología a desarrollar para cumplir los objetivos:

6.1 Estudio de la situación actual

En esta etapa se describirán las características del club de fidelización de la empresa, se detallarán las características de la promoción y se revisarán los datos disponibles para llevar a cabo la evaluación.

6.2 Selección y procesamiento de datos

En esta fase se seleccionarán los datos con los cuales se trabajará para el desarrollo de los modelos, limpiando de las bases valores perdidos e incoherencias que puedan presentarse.

6.3 Transformación de los datos

En este paso se realizarán las transformaciones necesarias a los datos para obtener las variables que serán utilizadas en los modelos.

6.4 Estimación de número de canjes

A través de modelos de series de tiempo se espera estimar la demanda de canjes por niveles y global durante el período de la promoción, que servirá de línea base para tener una idea de cómo hubiese sido el comportamiento de los clientes durante el mismo período de no haberse realizado la promoción y permitirá conocer el impacto en unidades de productos canjeados de ésta.

6.5 Modelo Proporcional de Hazard

Usando modelos econométricos de aceleración de compra, se espera identificar si los clientes que usaron la promoción debieron adelantar o incrementar niveles de compra para poder acceder a la promoción.

6.6 Modelos de predicción de canjes

En esta etapa se espera determinar el efecto de la promoción y de otras variables que pueden influir en la decisión de canje de los clientes, lo que permitirá identificar dentro de los clientes que utilizaron la promoción a aquellos que no hubiesen canjeado durante el mismo período de los que si lo hubiesen hecho de no haberse realizado la promoción.

6.7 Análisis descriptivo de los clientes que usaron la promoción

En esta etapa se espera describir a los clientes que usaron la promoción, identificando características comunes que permitan desarrollar perfiles.

7. Marco conceptual

7.1 Modelos de series de tiempo

En 1970, Box y Jenkins desarrollaron una metodología para modelar el comportamiento de series temporales, los modelos Arima. Estos modelos utilizan los mismos datos de la serie para identificar si las relaciones inter-temporales entre las variables se mantienen en el tiempo generando patrones que puedan ser capaces de explicar el comportamiento de la serie en el futuro [13].

Para comprender el modelo Arima, es necesario definir algunos conceptos:

7.1.1 Definiciones básicas

7.1.1.1 *Proceso estocástico*

Se define como una sucesión de variables aleatorias Y_t ordenadas y equidistantes cronológicamente. Un ejemplo de proceso estocástico es el siguiente: $Y_{-5}, Y_{-4}, Y_{-3}, \dots, Y_2, Y_3, Y_4$.

Donde cada una de las variables Y_t tiene su función de distribución propia, y cada grupo de variables su función de distribución conjunta, que permite caracterizar el proceso estocástico, sin embargo, estas funciones pueden resultar difíciles de conocer, por lo que para caracterizar el proceso basta con especificar la media y la varianza para cada Y_t y la covarianza para variables referidas a distintos valores de t :

$$E[Y_t] = \mu_t \quad (1)$$

$$\sigma_t^2 = \text{Var}(y_t) = E[(y_t - \mu_t)^2] \quad (2)$$

$$Y_{t,s} = \text{Cov}(Y_t, Y_s) = E[(y_t - \mu_t)(y_s - \mu_s)] \quad (3)$$

7.1.1.2 *Series de tiempo*

Se define como una serie de tiempo a una sucesión de n valores observados de una variable referidos a periodos de tiempo generalmente regulares, estas se representan de la siguiente forma $\{Y_1, \dots, Y_n\}$, donde Y_t corresponde a la observación de la serie en el período t .

7.1.1.3 *Estacionariedad*

Un proceso estocástico es estacionario en sentido estricto o fuerte si sus funciones de distribución conjuntas son constantes invariantes con respecto a un desplazamiento en el tiempo, es decir:

$$F(Y_t, Y_{t+1}, \dots, Y_{t+k}) = F(Y_{t+m}, Y_{t+1+m}, \dots, Y_{t+k+m}) \quad (4)$$

Como se mencionó anteriormente, conocer las funciones de distribución puede resultar difícil, por lo que es posible relajar la estacionariedad estricta utilizando la denominada estacionariedad en sentido amplio o débil:

$$E[Y_t] = E[Y_{t+m}] \quad \forall m \quad (5)$$

$$\text{Var}[Y_t] = \text{Var}[Y_{t+m}] \neq \infty \quad \forall m \quad (6)$$

$$\text{Cov}(Y_t, Y_s) = \text{Cov}(Y_{t+m}, Y_{s+m}) \quad \forall m \quad (7)$$

7.1.1.4 Modelo Autorregresivo

Un modelo se denomina autorregresivo si la variable aleatoria Y puede expresarse como una combinación lineal de sus valores en períodos anteriores, más un término de error. Estos modelos se abrevian con la palabra $AR(p)$, donde p indica el orden, que corresponde al número de observaciones retrasadas que intervienen en la ecuación.

Un modelo $AR(p)$ queda representado por la siguiente ecuación:

$$Y_t = \varphi_0 + \varphi_1 * Y_{t-1} + \varphi_2 * Y_{t-2} + \dots + \varphi_p * Y_{t-p} + a_t \quad (8)$$

Donde a_t corresponde a un término de error, el que se denomina ruido blanco si cumple con las siguientes condiciones:

- Tiene media nula $\rightarrow E(a_t) = 0$
- Varianza constante $\rightarrow \text{Var}(a_t) = \sigma^2$
- Covarianza nula entre errores de observaciones distintas $\rightarrow \text{Cov}(a_i, a_j) = 0, \forall i \neq j$

La ecuación (1), puede describirse de la siguiente forma:

$$\Phi_p(L) * Y_t = \varphi_0 + a_t \quad (9)$$

Donde $\Phi_p(L)$ corresponde al operador polinomial de retardos

$$\Phi_p(L) * Y_t = (1 - \varphi_1 * L - \varphi_2 * L^2 - \dots - \varphi_p * L^p) * Y_t \quad (10)$$

y L se conoce como operador retardo:

$$L * Y_t = Y_{t-1} \quad (11)$$

7.1.1.5 Modelo de medias móviles

Un modelo de medias móviles explica el valor de una variable en un período t en función de una sucesión de errores correspondientes a períodos anteriores y un término independiente. Estos modelos se abrevian con la letra $MA(q)$, donde q indica el orden.

Un modelo $AR(p)$ queda representado por la siguiente ecuación:

$$Y_t = \mu + a_t + \theta_1 * a_{t-1} + \theta_2 * a_{t-2} + \dots + \theta_q * a_{t-q} \quad (12)$$

Esta ecuación puede describirse de la siguiente forma:

$$Y_t = \Theta_q(L) * a_t + \mu \quad (13)$$

Donde μ es constante y $\Theta_q(L)$ corresponde al polinomio de retardos, que tiene la siguiente estructura:

$$\Theta_q(L) * a_t = (1 - \theta_1 * L - \theta_2 * L^2 - \dots - \theta_q * L^q) * a_t \quad (14)$$

7.1.2 Modelos Arima

Como su nombre lo indica, los modelos Arima combinan componentes autorregresivos (AR) y de medias móviles (MA) para modelar series.

Si la serie es no estacionaria, es posible transformarla aplicando el operador diferencial ∇^d , donde d indica el número de veces que debe ser diferenciada la serie para transformarla en estacionaria:

$$\nabla^d Y_t = Y_t - Y_{t-d} \quad (15)$$

Este modelo se abrevia como $ARIMA(p,d,q)$ y se representan por la siguiente ecuación:

$$\Phi_p(L) * (\nabla^d Y_t - \mu) = \Theta_q(L) * a_t \quad (16)$$

7.1.3 Modelos Sarimax

El modelo SARIMAX es una extensión del modelo ARIMA, pues permite incorporar estacionalidad (S) y la influencia de variables externas (X).

Si la serie en estudio presenta un comportamiento estacional S, es posible captar ese comportamiento incorporando al modelo tres nuevos términos:

- sp: término estacional autorregresivo
- sd: término estacional diferencial
- sq: término estacional de medias móviles

Tal como se definió en el modelo anterior, al existir estacionalidad se define el operador diferencial estacional ∇_{sd}^s como sigue:

$$\nabla_{sd}^s Y_t = Y_t - Y_{t-s*sd} \quad (17)$$

Este modelo se abrevia como SARIMA(p,d,q)(sp,sd,sq)X y se representa por la siguiente ecuación:

$$\Phi_p(L) * \Phi_{sp}(L) * [(\nabla^d \nabla_{sd}^s (Y_t - \sum_{i=1}^r C_i * X_i) - \mu) = \Theta_q(L) * \Theta_{sq}(L) * a_t \quad (18)$$

Donde $\Phi_{sp}(L)$ y $\Theta_{sq}(L)$ son los operadores de retardos estacionales autorregresivos y de medias móviles respectivamente, y son de la forma:

$$\Phi_{sp}(L) = 1 - \phi_1 * L - \phi_2 * L^2 - \dots - \phi_{sp} * L^{sp} \quad (19)$$

$$\Theta_{sq}(L) = 1 - \theta_1 * L - \theta_2 * L^2 - \dots - \theta_{sq} * L^{sq} \quad (20)$$

Dentro de los modelos de series de tiempo expuestos anteriormente, es SARIMAX (p,d,q)(sp.sd,sq)X el que resulta ser el más completo para analizar la serie de número de canjes semanales, pues combina tanto las aproximaciones autorregresivas como las de medias móviles, no requiere que la serie sea estacionara e incorpora el efecto de estacionalidad y variables externas que puedan influir en el comportamiento de la serie.

Sin embargo, la decisión final del modelo a usar dependerá de las características de la serie, lo que será detallado en la sección 8.4.1.1.

7.1.4 Medidas de error

Para medir el desempeño de los modelos se utilizarán medidas de error que comparen los valores observados (Y) y los estimados (\hat{Y}). Estas medidas son:

7.1.4.1 Error Porcentual Absoluto Medio (MAPE)

Esta medida entrega el promedio de los valores absolutos de los errores porcentuales.

$$MAPE = \frac{1}{N} \sum_k \left| \frac{Y_k - \hat{Y}_k}{Y_k} \right| \quad (21)$$

Donde N corresponde al número de observaciones de la serie.

7.1.4.2 Error cuadrático medio normalizado

El error cuadrático medio normalizado es un estimador de las desviaciones en general entre los valores observados y pronosticados.

$$NMSE = \frac{\sum_k (Y_k - \hat{Y}_k)^2}{\sum_k (Y_k - \bar{Y}_k)^2} = \frac{1}{\sigma^2 N} \sum_k (Y_k - \hat{Y}_k)^2 \quad (22)$$

Donde σ^2 es la varianza de la serie observada.

7.2 Modelo Proporcional de Hazard

Para poder determinar si la promoción tuvo un efecto acelerador en las compras de los clientes, se usará el Modelo Proporcional de Hazard, un modelo de supervivencia propuesto por Cox en 1972, que plantea la probabilidad de ocurrencia de un determinado evento, en este caso la compra, en base al tiempo que ha transcurrido desde la ocurrencia anterior [14].

Esta probabilidad, llamada función Hazard ($h_i(t, X_t)$), plantea una relación multiplicativa entre una línea base ($h_i(t)$), que caracteriza la distribución de probabilidad entre compras de los individuos, y una función de covariables ($\exp(\beta \cdot X_t)$), que captura la influencia variables de marketing y características de los individuos. En el caso continuo, esta probabilidad se ve representada por la siguiente ecuación:

$$h_i(t, X_t) = h_i(t) * e^{\beta \cdot X_t} \quad (23)$$

Esta relación multiplicativa es la que permite proporcionalidad en el modelo de Hazard, pues asume que cambios de nivel en las variables independientes producen cambios proporcionales en la función de Hazard, independiente del tiempo, lo que puede verse en:

$$\left[\frac{h_i(t, X_t)}{h_i(t)} \right] = e^{\beta \cdot X_t} \quad (24)$$

Si consideramos que $f(t, X_t)$ corresponde a la función de distribución de la función Hazard para un individuo en el tiempo t y $F(t, X_t)$ a la función de distribución acumulada, es posible entonces reescribir la función de Hazard como sigue:

$$h_i(t, X_t) = \frac{f(t, X_t)}{1 - F(t, X_t)} \quad (25)$$

Donde $f(t, X_t)$ representa la probabilidad que el individuo realice compra en el período t , mientras que $(1 - F(t, X_t))$ corresponde a la función de supervivencia, es decir,

la probabilidad que el cliente sobreviva sin comprar hasta el período t, y se representa por $S(t, X_t)$.

Combinando las ecuaciones (23) y (25) es posible escribir:

$$\frac{dF(t, X_t)}{1 - F(t, X_t)} = h(t) * e^{\beta * X_t} * dt \quad (26)$$

Luego, resolviendo la ecuación de primer orden puede obtenerse la función de supervivencia:

$$F(t, X_t) = 1 - e^{-\int_0^t h(u) * e^{\beta * X_u} * du} \Leftrightarrow S(t, X_t) = e^{-\int_0^t h(u) * e^{\beta * X_u} * du} \quad (27)$$

Sustituyendo la ecuación (27) en la (25):

$$f(t, X_t) = h(t) * e^{\beta * X_t} * e^{-\int_0^t h(u) * e^{\beta * X_u} * du} \quad (28)$$

Desde donde es posible obtener la probabilidad de compra condicionada al tiempo que ha transcurrido desde la compra anterior. Sin embargo, haciendo el supuesto que los clientes realizan compras en tiempos discretos, usualmente en semanas, es posible reescribir la función de supervivencia para tiempos discretos:

$$S(t, X_t) = e^{-\sum_{u=1}^t e^{\beta * X_t} \int_{u-1}^u h(w) * dw} \quad (29)$$

A partir de la cual es posible reescribir la función de Hazard a tiempo discreto como sigue:

$$P(t, X_t) = 1 - \frac{S(t, X_t)}{S(t-1, X_{t-1})} = 1 - e^{-e^{\beta * X_t} \int_{t-1}^t h(u) du} \quad (30)$$

Donde $P(t, X_t)$ representa la probabilidad en tiempo discreto que la compra ocurra en el tiempo t con covariantes X_t .

Se escogió el modelo Proporcional de Hazard pues mediante este es posible determinar cómo influyen variables de marketing en el comportamiento intertemporal de compra de los clientes (determinado a través de la línea base del modelo) aumentando o disminuyendo la probabilidad que el cliente realice compra en una determinada semana, lo que permite determinar si las variables de marketing inducen o no a los clientes a acelerar compra.

7.2.1 Línea base

La forma de la función de línea base puede ser estimada de manera no paramétrica, a partir de los datos de compra y no compra de los clientes, o de forma paramétrica, asumiendo que el patrón intertemporal de compra sigue una distribución conocida.

A continuación se presentan tres distribuciones que serán testeadas para determinar la línea base paramétrica:

Tabla 4: Líneas bases paramétricas

Exponencial	$h(t) = \gamma$	Donde $\gamma > 0$, esta distribución tiene forma plana
Weibull	$h(t) = \gamma\alpha(\gamma t)^{\alpha-1}$	Donde $\gamma > 0$. Su forma puede ser plana, monótonica creciente y monótonica decreciente. Es la forma más usada en la literatura
Log-logística	$h(t) = \frac{\gamma\alpha(\gamma t)^{\alpha-1}}{1+(\gamma t)^\alpha}$	Donde $\gamma, \alpha > 0$. La forma de esta distribución puede ser monótonica decreciente o tener forma de U invertida.

La elección de la línea base dependerá entonces del ajuste del modelo a los datos.

7.2.2 Estimación de los parámetros

La estimación de los parámetros del modelo se realizará maximizando la función de verosimilitud, cuya forma es la siguiente:

$$L = \prod_{v=1}^T P(v, X_v)^{\delta_v} * [1 - P(v, X_v)]^{1-\delta_v} \quad (31)$$

Donde δ_v es una variable que toma el valor 1 cuando el cliente realiza compra en la semana v y toma el valor 0 en el caso contrario.

Así, maximizando la función L respecto a los parámetros buscados, será posible obtener aquellos que con mayor verosimilitud generaron los datos.

Luego, para determinar entre un modelo y otro se usará el criterio de información Bayesiano o BIC, el que se calcula como sigue:

$$\text{BIC} = k \cdot \ln(n) - 2 \cdot \log L_{\max} \quad (32)$$

Donde k corresponde al número de parámetros estimados, n al número de observaciones y $\log L_{\max}$ al valor máximo que toma el logaritmo de la función de verosimilitud. El modelo escogido será aquel que minimice el BIC, criterio que penaliza el ajuste del modelo por el número de parámetros usados.

7.2.3 Evaluación de la calidad predictiva

Para evaluar la calidad predictiva del modelo se usará la matriz de confusión [3], la que está representada por:

Tabla 5: Matriz de confusión

		Predicción	
		Compra	No compra
Real	Compra	A	B
	No compra	D	C

Donde:

- A: Número de compras correctamente predichos
- B: Número de compras mal identificados por el modelo
- C: Número de no compras correctamente predichos
- D: Número de no compras mal predichas por el modelo
- A + B: Número de compras reales
- A + D: Número de compras predichas
- C + D: Número de no compras reales
- B + C: Número de no compras predichas

La información aportada por esta matriz, permite calcular ciertos indicadores que ayudarán a medir la capacidad predictiva de los modelos. Estos indicadores son:

- $\text{Predic}_{\text{ok}}$: proporción de predicciones correctas
- C_{ok} : proporción de compras predichas correctamente
- NC_{ok} : proporción de no compras predichas correctamente
- C_{error} : proporción de compras mal identificadas (modelo predijo compra y era no compra)

- NC_{error} : proporción de no compras mal identificadas (modelo predijo no compra y era compra)

Y se calculan como sigue:

$$\text{Predic}_{\text{ok}} = \frac{A + C}{A + B + C + D} \quad (33)$$

$$C_{\text{ok}} = \frac{A}{A + B} \quad (34)$$

$$NC_{\text{ok}} = \frac{C}{C + D} \quad (35)$$

$$C_{\text{error}} = \frac{D}{C + D} \quad (36)$$

$$NC_{\text{error}} = \frac{B}{A + B} \quad (37)$$

7.3 Hipótesis del gradiente – objetivo

Originalmente propuesta por Hull en 1932, la hipótesis del gradiente – objetivo plantea que a medida que la meta se encuentra más cercana se incrementan los esfuerzos realizados para alcanzarla [9].

Esta hipótesis ha sido ampliamente utilizada para estudiar el comportamiento de animales, sin embargo, el estudio en humanos se ha visto limitado mayoritariamente a temas fisiológicos.

Actualmente, con los programas de fidelización y el CRM, se cuenta con una gran cantidad de información de los clientes, por lo que es posible adaptar esta teoría a la psicología humana para comprender como la cercanía a una recompensa puede influir en la decisión de compra.

En particular, para el presente trabajo, se pretende determinar como influye en la decisión de compra de los clientes la cercanía al nivel de canje y una promoción de canje de productos realizado por un club de fidelización de una empresa de retail. Para esto, se diseñará una variable que cuantifique esta distancia y que será incorporada en los modelos a desarrollar.

8. Desarrollo metodológico

8.1 Estudio de la situación actual

8.1.1 Programa de fidelización de la empresa

El programa de fidelización de la empresa en estudio presenta una estructura de recompensa basada en la entrega de puntos intercambiables por productos.

La forma de acumular puntos es la siguiente:

Tabla 6: Estructura de acumulación de puntos del programa

Cada	Puntos acumulados
\$120 en compras con tarjeta	1
\$950 en avances en efectivo	1
\$360 en emisión de facturas	1

Los puntos acumulados por cada cliente son intransferibles y tienen una vigencia de aproximadamente 12 meses, venciendo el último día del mismo mes que se realizó la compra del año siguiente.

Para poder realizar canjes, los clientes deben haber alcanzado un nivel de puntos mínimo requerido por cada nivel pudiendo, en caso de tener una mayor cantidad de puntos, dejarlos guardados en su cuenta puntos para un canje posterior.

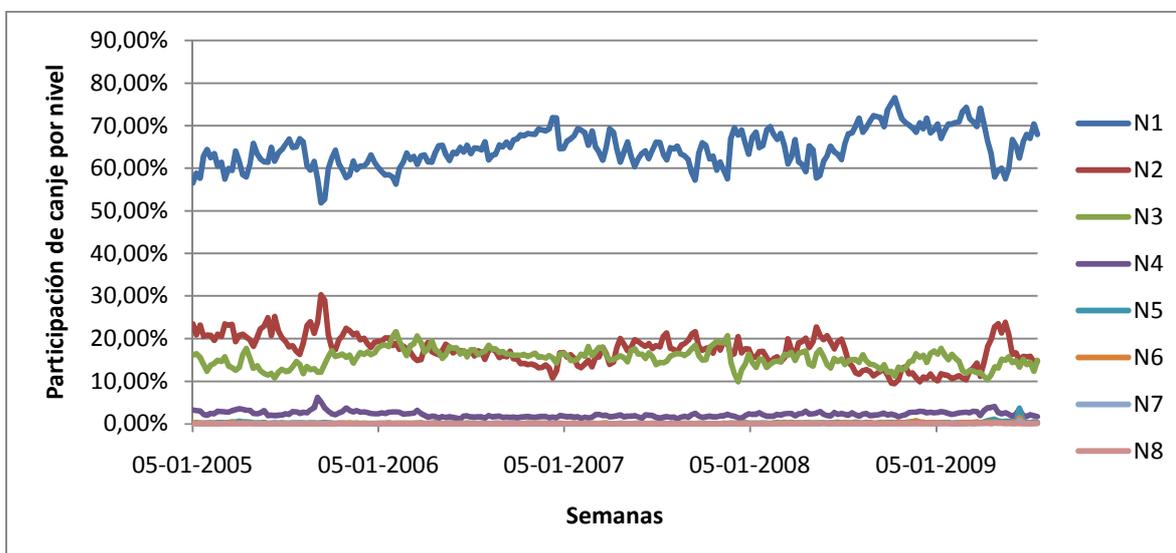
El programa cuenta con 8 niveles de canjes, los que son descritos a continuación:

Tabla 7: Estructura de Niveles de la empresa

Nivel	Puntos acumulados requeridos
1	5.000
2	9.000
3	12.000
4	24.000
5	48.000
6	60.000
7	90.000
8	120.000

La participación de cada uno de los niveles en los canjes semanales puede verse reflejada en el siguiente gráfico:

Gráfico 1: Participación de canjes por nivel



Donde se observa que el nivel 1 en promedio tiene una participación del 65% de los canjes, el nivel 2 de un 17,2% y el nivel 3 de un 15,4%, concentrando estos niveles aproximadamente el 97% de los canjes de una semana.

8.1.2 Promoción a evaluar

La acción de marketing a evaluar es la promoción denominada “Puntos con pesos”, que fue realizada por la empresa durante el período del 2 al 19 de octubre del 2008.

Para acceder a la promoción, los clientes debían contar con un puntaje mínimo (inferior al requerido por nivel en períodos sin promoción) y cancelar un monto en pesos, lo que les otorgaba la posibilidad de canjear en niveles superiores a los que les correspondía de acuerdo a la cantidad de puntos acumulados.

A continuación se presenta la tabla donde se especifica el puntaje mínimo y el monto a cancelar para cada nivel de canje:

Tabla 8: Estructura de Niveles promoción Puntos con Pesos

Nivel	Puntos acumulados requeridos	Monto a pagar
1	3.000	\$ 3.990
2	5.000	\$ 6.990
3	9.000	\$ 9.990
4	12.000	\$ 24.990
5	24.000	\$ 59.990
6	48.000	\$ 99.900
7	60.000	\$ 199.900
8	90.000	\$399.000

Posteriormente, a contar del 17 de abril del 2009, esta modalidad de canje fue incorporada de forma permanente en el programa de puntos.

Este segundo período promocional será utilizada para validar los modelos y determinar el efecto de la promoción en un largo plazo.

8.2 Selección y procesamiento de los datos

Para el desarrollo de de la memoria se cuenta con información transaccional y demográfica de 3. 918.692 clientes que pertenecen al programa de fidelización.

La información demográfica de los clientes corresponde a su edad, sexo y antigüedad de la cuenta. La información transaccional consiste en:

- Puntos acumulados por los clientes que pertenecen al programa de fidelización de la empresa al 31 de diciembre del 2004.
- Puntos entregados, canjeados y vencidos diarios desde el 1° de enero del 2005 hasta el 31 de julio del 2009.
- Detalle de los canjes diarios desde el 1° de enero del 2005 hasta el 31 de julio del 2009.
- Edad, sexo, antigüedad de la cuenta y actividad de los clientes.

Del total de los clientes, aproximadamente un 31% (1.214.617 clientes) realizó al menos un canje durante el período de estudio.

Debido a que la promoción a analizar en su primera etapa se extendió por dieciocho días, se determinó tratar los datos de los clientes de manera semanal, por lo

que fue necesario transformar los datos. Luego la estructura conseguida por cliente fue la siguiente:

- $PE_i(t)$: Puntos entregados la semana t al cliente i .
- $PC_i(t)$: Puntos canjeados por el cliente i en la semana t .
- $PV_i(t)$: Puntos vencidos la semana t al cliente i .

Una vez transformados los datos a semanales, se procedió a limpiar de las bases los errores, que principalmente consistían en puntos entregados o canjeados negativos, debido a devoluciones realizadas por los clientes, los que fueron corregidos en las bases.

Posteriormente se calculó el puntaje acumulado por cada cliente al final de cada semana y la consistencia entre puntos acumulados, entregados, canjeados y vencidos semanalmente usando las fórmulas:

$$PA_i(t) = PA_i(t-1) + PE_i(t) - PC_i(t) - PV_i(t) \quad (38)$$

$$PA_i(t-1) + PE_i(t) \geq PC_i(t) + PV_i(t) \quad (39)$$

La restricción (39) permite identificar a aquellos clientes “inconsistentes” que canjean sin contar con la cantidad de puntos acumulados necesarios para hacerlo, por ejemplo, aquellos clientes que luego de canjear un producto realizaron la devolución de alguna compra con la que acumularon puntos para el canje, quedando con una acumulación negativa de puntos. Del total de los clientes canjeadores, un 37% presenta alguna inconsistencia, lo que es un porcentaje bastante alto, sin embargo es importante considerar que, en su mayoría, estas inconsistencias se producen una vez dentro del período.

La selección de la muestra a utilizar para el desarrollo de los modelos, se realizó con aquellos que resultaron ser consistentes. Se excluyó del análisis a los inconsistentes, pues pueden generar errores en los modelos.

En el Anexo 1: Cálculo del tamaño muestral, se detalla el procedimiento usado para estimar tamaño de la muestra para cada uno de los niveles de canje.

8.3 Transformación de los datos

Para obtener los datos a utilizar en el desarrollo de los distintos modelos, se procedió a transformar los datos seleccionados en la etapa anterior, estas transformaciones se presentan a continuación:

8.3.1 Variables RFM

A continuación se expone la metodología usada para obtener las variables RFM de compra y canje.

8.3.1.1 Recency

Corresponde al número de semanas que ha pasado desde la última ocurrencia del evento, para el caso en estudio la compra y el canje. Dado que se cuenta con información semanal a contar de enero del 2005, se asume que el evento ocurrió la última semana del 2004, por lo que los clientes inician el período en estudio con recency 1.

8.3.1.2 Frequency

Se define la variable frecuencia como:

$$F_{i,t} = \frac{\sum_{k=\alpha}^t \Delta_{i,k}}{t - \alpha + 1} \quad (40)$$

Donde α corresponde a la semana en que ocurre la compra o el canje por primera vez y Δ_{ik} toma valor 1 si el evento ocurre en la semana k y 0 si no.

8.3.1.3 Monetary

Se define la variable monto como el promedio acumulado de puntos canjeados o entregados (compra) semanalmente hasta la semana t . Su fórmula es la siguiente:

$$M_{i,t} = \frac{\sum_{k=\alpha}^t PX_{i,k}}{t - \alpha + 1} \quad (41)$$

Donde α corresponde a la semana en que ocurre el evento por primera vez y PX representa los puntos entregados (PE_i) o puntos canjeados (PC_i) para el monto de compra y canje respectivamente.

8.3.2 Distancia psicológica al objetivo

La distancia psicológica corresponde a la variable diseñada para cuantificar la cercanía o lejanía al nivel de canjes. A través de esta se espera testear la hipótesis del gradiente objetivo. Su forma es la siguiente:

$$\text{Dist-Canje} = \frac{DN_i}{1 + \ln(1 + \text{Num_canje_acum})} \quad (42)$$

Donde Num_canje_acum corresponde al número de canjes acumulados durante el último año y DN_i a la distancia relativa, en puntos acumulados, al nivel de canje i y su forma es la siguiente:

$$DN_i = \begin{cases} \frac{P_Nivel_j - PA_t}{P_Nivel_j} & \text{Si } PA_t \leq P_Nivel_j \\ 0 & \text{Si no} \end{cases} \quad (43)$$

Esta variable toma valores entre 0 y 1, donde valores cercanos a 1 indican que el cliente tiene una baja acumulación de puntos para alcanzar el nivel y valores cercanos a 0 indican que el cliente está cercano a canjear.

El objetivo de incorporar el número de canjes acumulados durante el último año es testear si el efecto acelerador que se desea investigar a través de la distancia relativa se ve influenciado por el número de canjes acumulados por el cliente durante el último período. El uso del logaritmo natural se debe a que se desea evitar la linealidad de este efecto.

8.3.3 Variables Dummies

Las variables consideradas son:

Fin de mes:	$\begin{cases} 1 \\ 0 \end{cases}$	$\begin{cases} \text{Si la semana es fin de mes} \\ \text{Si no} \end{cases}$
Fiestas patrias	$\begin{cases} 1 \\ 0 \end{cases}$	$\begin{cases} \text{Si la semana es la semana de fiestas patrias} \\ \text{Si no} \end{cases}$
Navidad	$\begin{cases} 1 \\ 0 \end{cases}$	$\begin{cases} \text{Si la semana es es la semana de navidad} \\ \text{Si no} \end{cases}$
Semana santa	$\begin{cases} 1 \\ 0 \end{cases}$	$\begin{cases} \text{Si la semana es semana santa} \\ \text{Si no} \end{cases}$
Promoción	$\begin{cases} 1 \\ 0 \end{cases}$	$\begin{cases} \text{Si la semana corresponde a una semana de promoción} \\ \text{Si no} \end{cases}$
Pre-promoción	$\begin{cases} 1 \\ 0 \end{cases}$	$\begin{cases} \text{Si la semana es una de las 3 semanas previas a la promoción} \\ \text{Si no} \end{cases}$

Post-promoción $\begin{cases} 1 & \text{Si la semana es una de las 3 semanas posteriores a la promoción} \\ 0 & \text{Si no} \end{cases}$

8.4 Estimación de número de canje

8.4.1 Estimación número de canjes globales

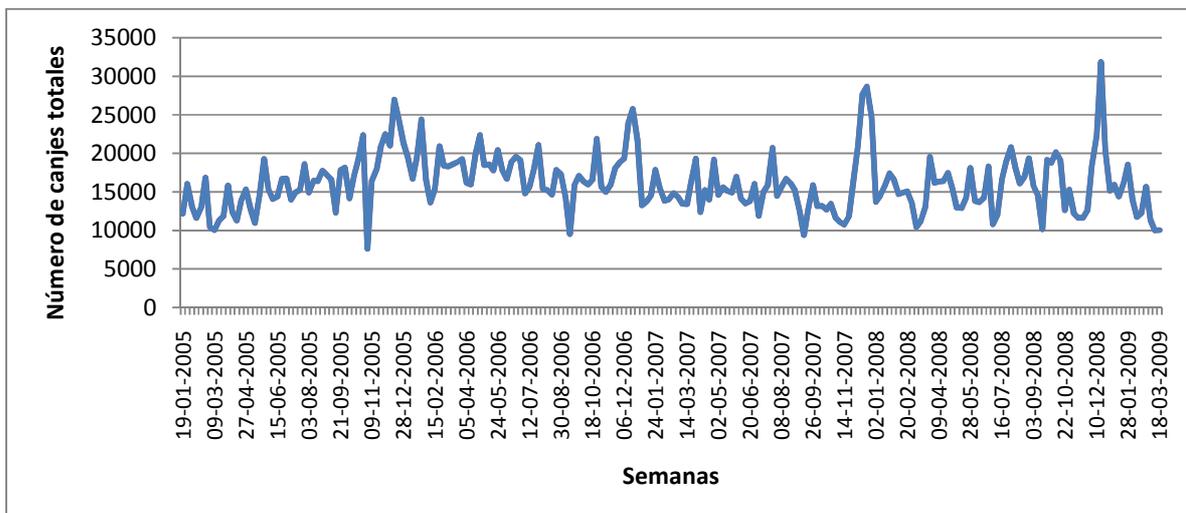
Para determinar el impacto de la promoción en el número de productos canjeados semanalmente se describirá la serie, analizando estacionariedad, el orden de los modelos y las variables externas que puedan influir en el comportamiento de la serie. Esto permitirá contar con la información necesaria para calibrar el modelo, el que luego será validado y evaluado de acuerdo a la calidad de su ajuste.

Como un primer análisis, se calibrará el modelo sin incluir ninguno de los períodos promocionales (Modelo 1), esto con el objetivo de determinar cuál hubiese sido el comportamiento de la variable canje de no haberse realizado la promoción durante esas semanas. Una segunda parte de este análisis consistirá en incluir el primer período promocional a la calibración para luego pronosticar el número de canjes durante la segunda promoción (Modelo 2).

8.4.1.1 Análisis de la serie número de canjes

Para poder estimar el número de canjes, es necesario primero analizar la serie. A continuación se presenta el gráfico de la serie número de canjes semanales durante el período del 5 de enero del 2005 al 18 de marzo del 2009:

Gráfico 2: Número de canjes totales semanales



Observando el gráfico anterior, es posible ver que el número de canjes se mantiene relativamente constante en el tiempo, sin embargo, esto no asegura la estacionariedad. Para testear este supuesto, se aplicará el test de raíces unitarias de Dickey Fuller.

Tabla 9: Test Dickey Fuller para la serie número de canjes globales

ADF Test Statistic	-7.682990	1% Critical Value*	-3.4597
		5% Critical Value	-2.8739
		10% Critical Value	-2.5733

*MacKinnon critical values for rejection of hypothesis of a unit root.

Augmented Dickey-Fuller Test Equation

Dependent Variable: D(N_TOTAL)

Method: Least Squares

Date: 09/14/09 Time: 02:07

Sample(adjusted): 1/26/2005 7/22/2009

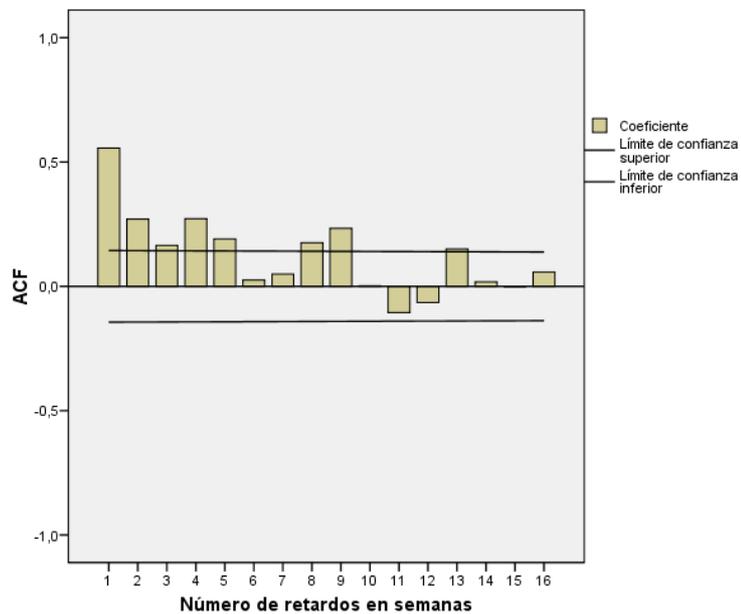
Included observations: 235 after adjusting endpoints

Variable	Coefficient	Std. Error	t-Statistic	Prob.
N_TOTAL(-1)	-0.400111	0.052077	-7.682990	0.0000
C	6332.553	846.3789	7.481937	0.0000
R-squared	0.202132	Mean dependent var	10.49362	
Adjusted R-squared	0.198708	S.D. dependent var	3392.849	
S.E. of regression	3037.106	Akaike info criterion	18.88367	
Sum squared resid	2.15E+09	Schwarz criterion	18.91312	
Log likelihood	-2216.831	F-statistic	59.02833	
Durbin-Watson stat	1.927733	Prob(F-statistic)	0.000000	

Dado que el estadístico ADF es menor que los valores críticos para cada uno de los intervalos de confianza, y que el estadístico Durbin-Watson es cercano a dos, es posible rechazar la hipótesis nula que indica la existencia de raíces unitarias. Luego es posible decir que la serie es estacionaria alrededor de una constante.

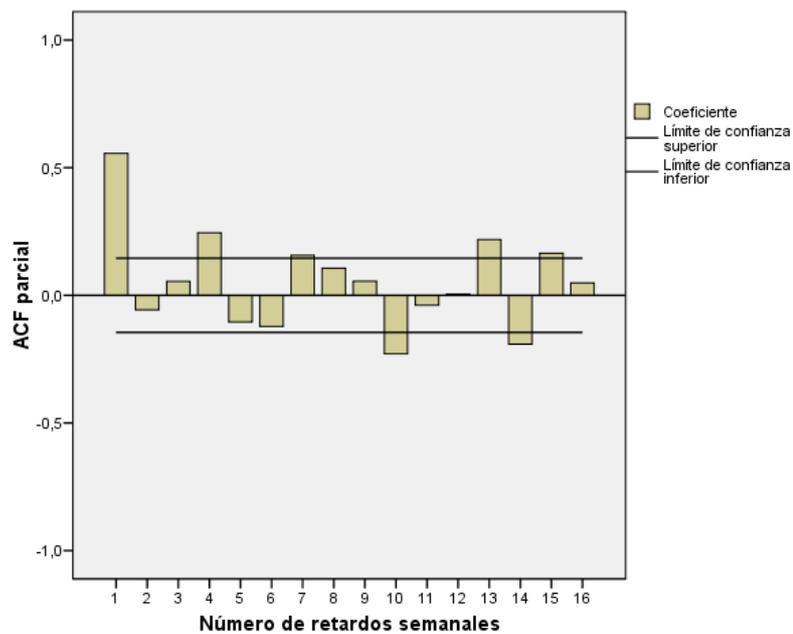
Para determinar el orden del modelo, es necesario analizar la función de autocorrelación y la de correlación cruzadas de la serie.

Gráfico 3: Autocorrelaciones de la serie Número de Canjes Totales



Es posible observar que el número de canjes de la semana t se correlaciona con los retardos uno y dos de la serie, es decir, con el número de canjes de la semana $t-1$ y $t-2$. Esto sugiere la presencia de un proceso de medias móviles de orden 2 ($q=2$).

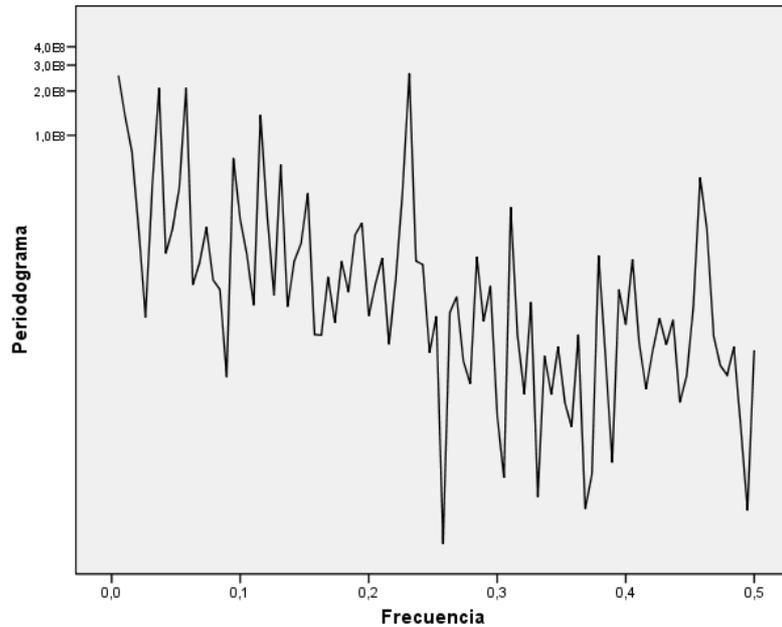
Gráfico 4: Autocorrelaciones Parciales de la serie Número de Canjes Totales



La presencia de un peak en la autocorrelación parcial de la serie número de canjes indica la presencia de un proceso autorregresivo de orden uno ($p=1$).

Ahora se debe entonces analizar la presencia de estacionalidad en la serie, para lo cual se analizará el periodograma de ésta:

Gráfico 5: Periodograma Serie número de canjes totales por frecuencia



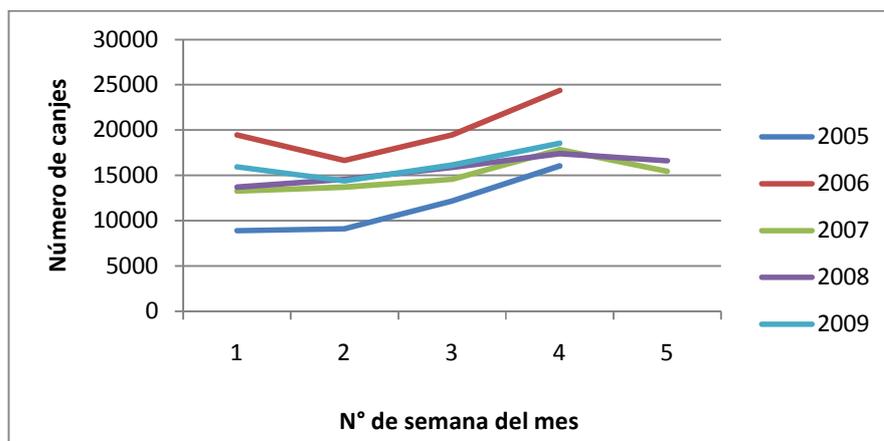
El gráfico muestra el peak de frecuencia más baja cercano a la frecuencia 0,0. Dado que el modelo es semanal, es posible entonces sospechar la presencia de periodicidad anual. Luego, dado que frecuencia y período son recíprocos, es posible estimar la frecuencia resolviendo $1/52$, que es igual a 0,0192 y que es aproximadamente la frecuencia donde se produce el primer peak.

Luego, el modelo a testear es el $ARIMA(1,0,2)(1,0,0)$.

8.4.1.2 Variables externas

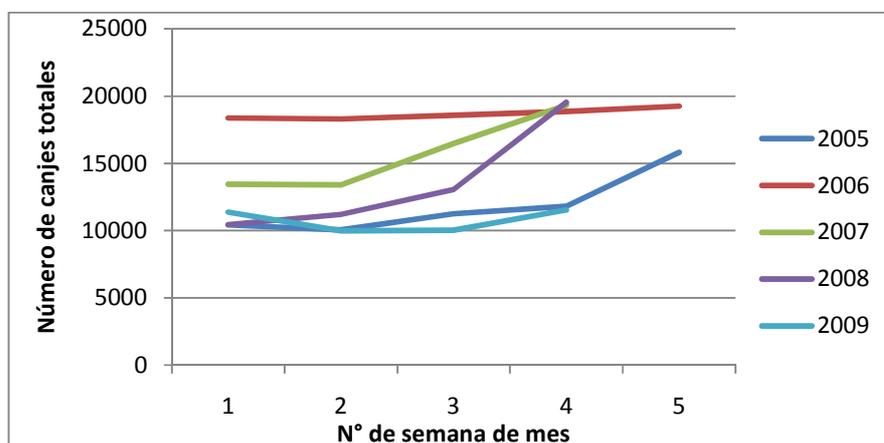
Para determinar qué variables externas pueden influir en el comportamiento de la serie número de canjes, se analizará el comportamiento de la serie mes a mes. El análisis de los meses más relevantes se presenta a continuación, el resto de los meses se encuentra en el Anexo 2: Número de canjes semanales por mes

Gráfico 6: Número de canjes semanales, mes de enero



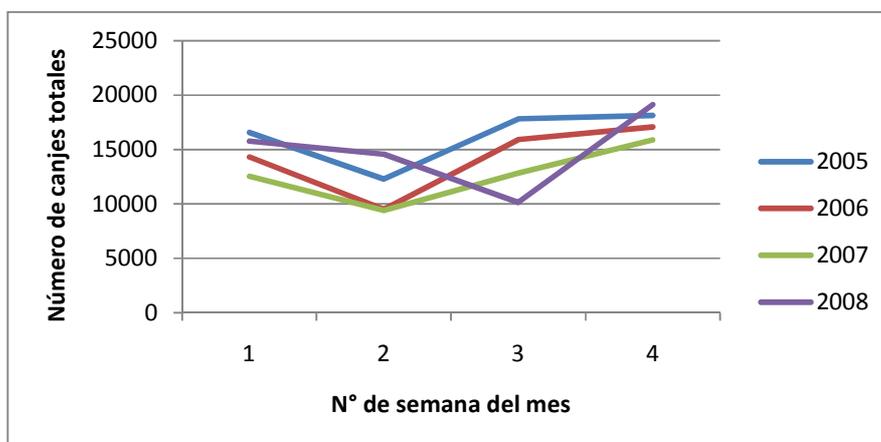
Observando el gráfico, es posible determinar que el comportamiento del mes es similar a lo largo de los años, presentando una disminución del número de canjes en la segunda semana, seguida por un aumento a contar de la semana 3 y 4. También es posible observar que durante el año 2006 se realizó un mayor número de canjes que en los años posteriores, lo que puede deberse al cambio en la tasa de conversión de los puntos entregados, la que fue modificada en el año 2006 y que se ve reflejada al año siguiente debido a la vigencia de los puntos

Gráfico 7: Número de canjes semanales, mes de marzo



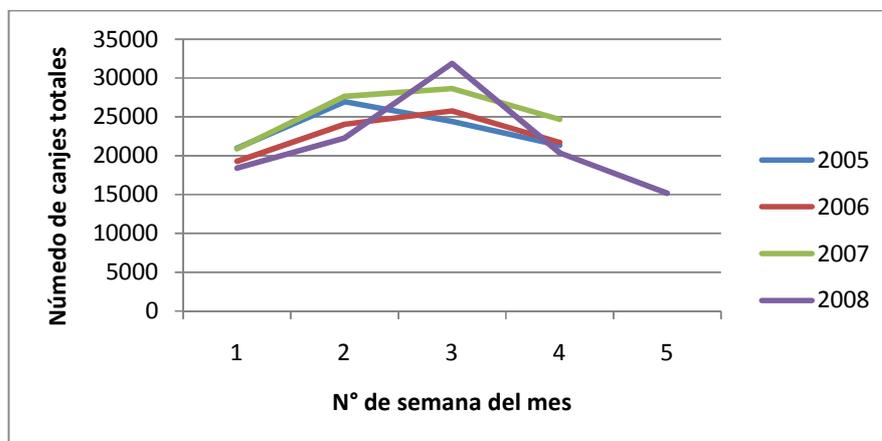
Es posible observar un aumento considerable del número de canjes durante la última semana del mes, la que puede deberse al incremento de puntos producidos por las compras escolares del año actual y del pasado, pues los puntos entregados por compras escolares realizadas a comienzos de marzo de un año vencerán a fines de marzo del año siguiente.

Gráfico 8: Número de canjes semanales, mes de septiembre



Se observa una disminución del número de canjes entre la segunda y tercera semana del mes, el que se asocia a las Fiestas Patrias.

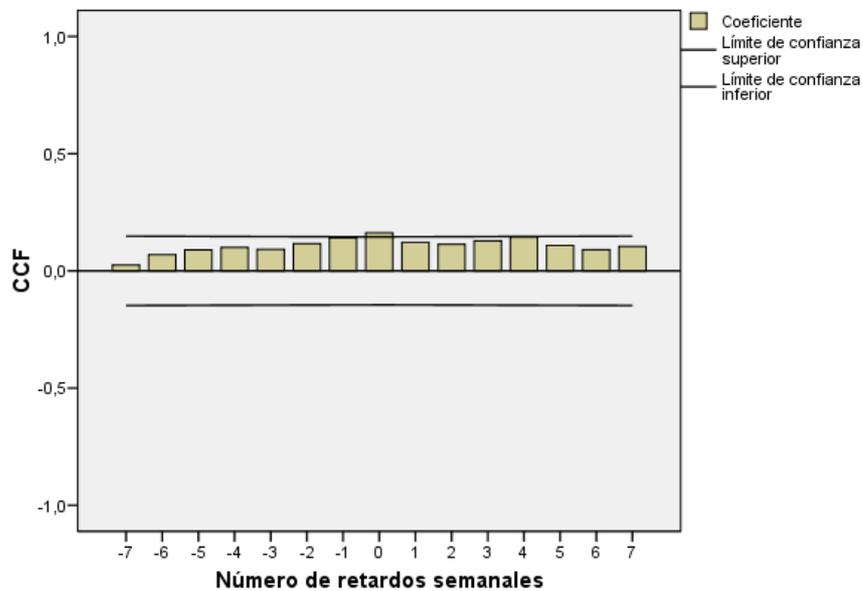
Gráfico 9: Número de canjes semanales, mes de diciembre



A partir del gráfico es posible identificar un incremento del número de canjes durante la tercera semana del mes, el que se asocia a canjes realizados durante el período navideño.

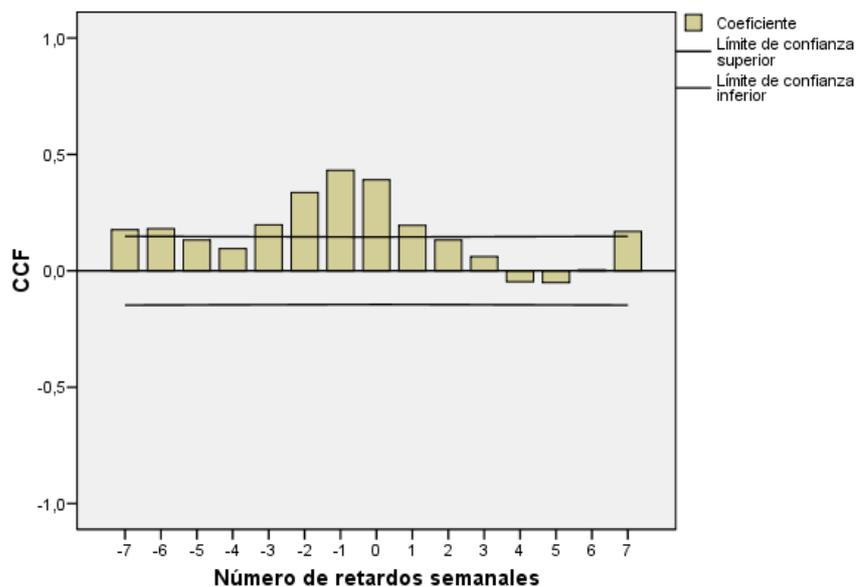
Otras variables que puede influir en el número de canjes son los puntos acumulados, entregados y vencidos, el que puede determinarse analizando los gráficos de correlaciones cruzadas.

Gráfico 10: Correlaciones cruzadas Numero de canjes y Puntos acumulados



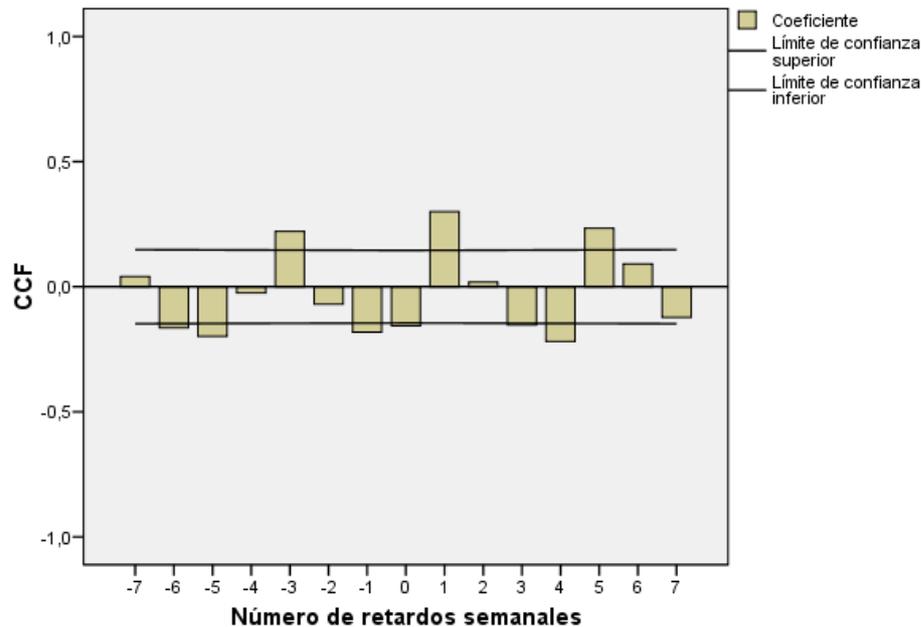
Del gráfico se desprende, que la única correlación cruzada que supera el límite de confianza es con el retardo cero, lo que indica que el número de canjes de una semana se correlaciona con los puntos acumulados de la misma semana, y dado que lo que se desea es pronosticar el número de canjes en base a información disponible de períodos anteriores, esta variable no será considerada en el modelo.

Gráfico 11: Correlaciones cruzadas Numero de canjes y Puntos entregados



Es posible observar una correlación cruzada significativa con los retardos 0, -1 y -2, lo que es indicio que el número de canjes de la semana t se relaciona con los puntos entregados durante la semana t y con los puntos entregados durante las dos semanas anteriores.

Gráfico 12: Correlaciones cruzadas Numero de canjes y Puntos vencidos



Es posible identificar del gráfico una correlación significativa del número de canjes con los puntos vencidos 3 semanas antes.

Luego, a partir del análisis, las variables a incorporar al modelo son:

- Fin de mes enero, febrero, marzo, julio, septiembre, noviembre
- Semana santa
- Fiestas Patrias
- Navidad
- Puntos entregados durante las dos semanas anteriores
- Puntos vencidos en tres períodos anteriores.

8.4.1.3 Estimación de los parámetros

Una vez conocidos los órdenes del modelo y las variables externas a incorporar, se procede a estimar los parámetros.

Como se mencionó en la sección 8.4.1, se estimarán dos modelos, con distintos períodos de calibración, los que se presentan a continuación.

8.4.1.3.1 Modelo 1

El período de calibración de este modelo va desde el 5 de enero del 2005 al 3 de septiembre del 2008.

Las variables que resultaron significativas son las siguientes:

Tabla 10: Parámetros del Modelo 1 de estimación del número de canjes

		Estimaciones	t	Sig. aprox.
Retardos no estacionales	AR1	0,90	17,84	-
	MA1	0,27	2,88	0,00
	MA2	0,23	2,56	0,01
Retardos estacionales	AR1 Estacional	0,50	7,03	-
Coefficientes de regresión	PE_1	1,35E-05	2,79	0,01
	PATRIAS	- 4.585,66	- 3,01	0,00
	NAVIDAD	3.474,26	3,31	0,00
	SEMSANTA	- 2.413,11	- 3,05	0,00
	fdm_1	1.923,17	1,96	0,05
	fdm_2	1.822,55	1,85	0,07
	fdm_3	3.095,72	2,18	0,03
	fdm_4	1.987,77	2,02	0,04
	fdm_5	1.861,01	1,89	0,06
	fdm_6	3.302,11	2,37	0,02
	fdm_7	3.788,05	3,85	0,00
	fdm_10	3.142,86	3,06	0,00
	fdm_12	3.230,79	2,10	0,04
Constante		12.826,43	8,41	-

Analizando el valor de los parámetros obtenidos para las variables externas, es posible identificar que estos toman valores de acuerdo a lo esperado según lo analizado en la sección 8.4.1.2.

Es importante mencionar, que si bien el parámetro asociado a los puntos entregados en períodos anteriores (PE_1 y PE_2) son pequeños, éstas series son del orden de 10^8 , por lo que representan un porcentaje considerable de los canjes realizados en el período.

8.4.1.3.2 Modelo 2

El período de calibración de este modelo es del 5 de enero del 2005 al 18 de marzo del 2009.

Para este modelo se introducen tres nuevas variables dummies: pre-promoción, promoción, y post-promoción, las que fueron explicadas en la sección 8.3.3.

A continuación se presenta la tabla donde se exponen aquellas variables que resultaron significativas para el modelo 2:

Tabla 11: Parámetros del Modelo 2 de estimación del número de canjes

		Estimaciones	t	Sig. aprox.
Retardos no estacionales	AR1	0,92	22,29	-
	MA1	0,33	4,05	0,00
	MA2	0,25	3,12	0,00
Retardos estacionales	AR1 estacional	0,53	8,46	-
Coefficientes de regresión	PE_1	9,40E-06	2,18	0,03
	PROMO	4.171,59	2,69	0,01
	PATRIAS	- 4.930,64	- 4,67	0,00
	NAVIDAD	4.504,08	4,31	0,00
	SEMSANTA	- 2.060,74	- 2,57	0,01
	fdm_1	2.200,34	2,19	0,03
	fdm_2	2.100,47	2,10	0,04
	fdm_3	3.212,36	2,09	0,04
	fdm_4	1.980,07	1,91	0,06
	fdm_6	3.157,26	2,08	0,04
	fdm_7	3.815,51	3,67	0,00
	fdm_9	2.992,30	1,95	0,05
	fdm_10	3.267,98	3,14	0,00
fdm_12	2.397,43	2,31	0,02	
Constante		13.280,22	8,94	-

Donde es posible observar que, además del hecho que los parámetros toman valores de acuerdo a lo esperado, resulta significativa la variable promoción y su efecto es positivo, por lo que estar en una semana de promoción incrementa el número de canjes.

8.4.1.4 Ajuste de los modelos

En esta sección se analiza el ajuste de los modelos durante el período de calibración.

8.4.1.4.1 Modelo 1

La siguiente tabla presenta la calidad del ajuste del modelo 1:

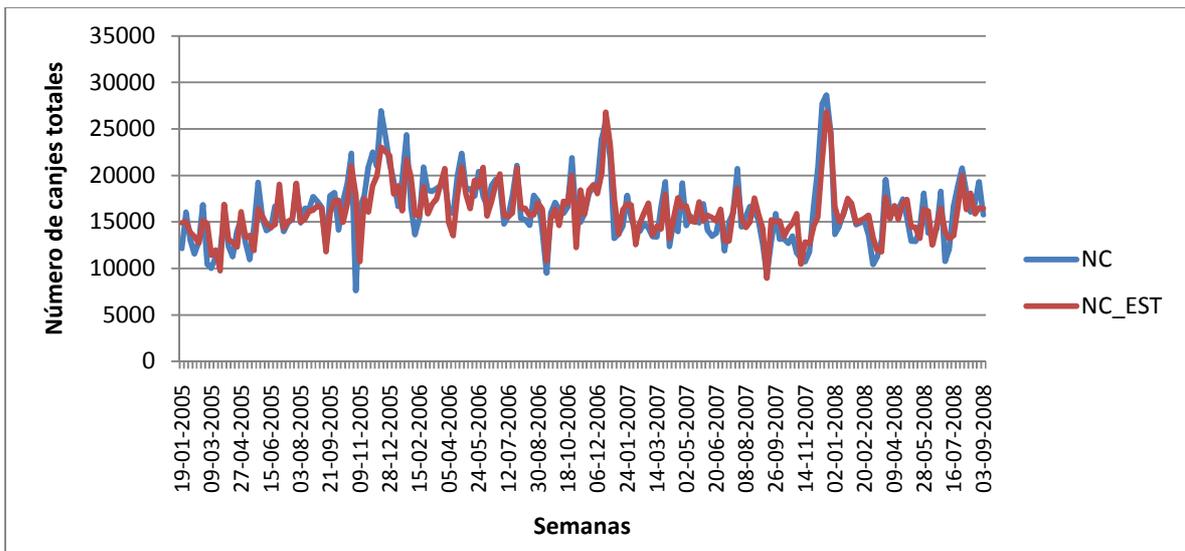
Tabla 12: Ajuste del Modelo 1 en el período de calibración

R²	66,23%
R	81,38%
MAPE	10,20%

Se observa de la tabla anterior que aproximadamente un 66,23% de la varianza de los datos puede ser explicada por el Modelo 1. El MAPE del modelo durante este período indica que, en promedio, la diferencia porcentual entre el número de canjes reales y el estimado es de un 10,2%, que es un porcentaje de error bajo.

El ajuste del modelo puede verse también reflejado gráficamente:

Gráfico 13: Ajuste del Modelo 1 en el período de calibración



8.4.1.4.2 Modelo 2

La siguiente tabla presenta la calidad del ajuste del Modelo 2:

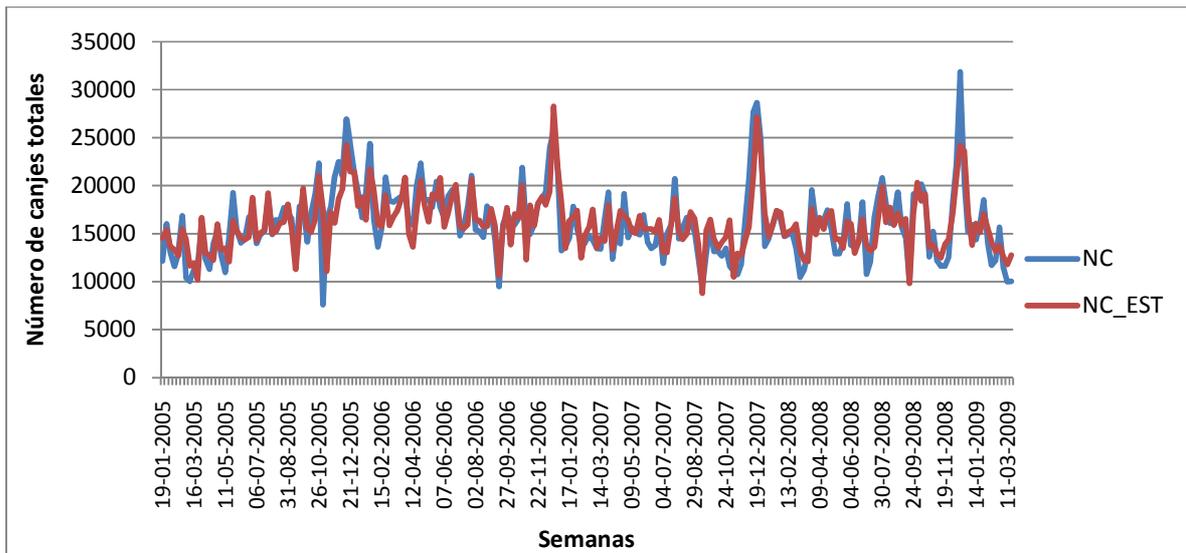
Tabla 13: Ajuste del Modelo 2 en el período de calibración

R ²	68,95%
R	83,04%
MAPE	10,23%

Se observa de la tabla anterior que aproximadamente un 68,95% de la varianza de los datos puede ser explicada por el modelo. El MAPE del modelo durante este período, al igual que para el Modelo 1, fue de un 10,2%.

Gráficamente, el ajuste de Modelo 2 puede verse a continuación:

Gráfico 14: Ajuste del Modelo 2 en el período de calibración



8.4.1.5 **Pronóstico de los modelos**

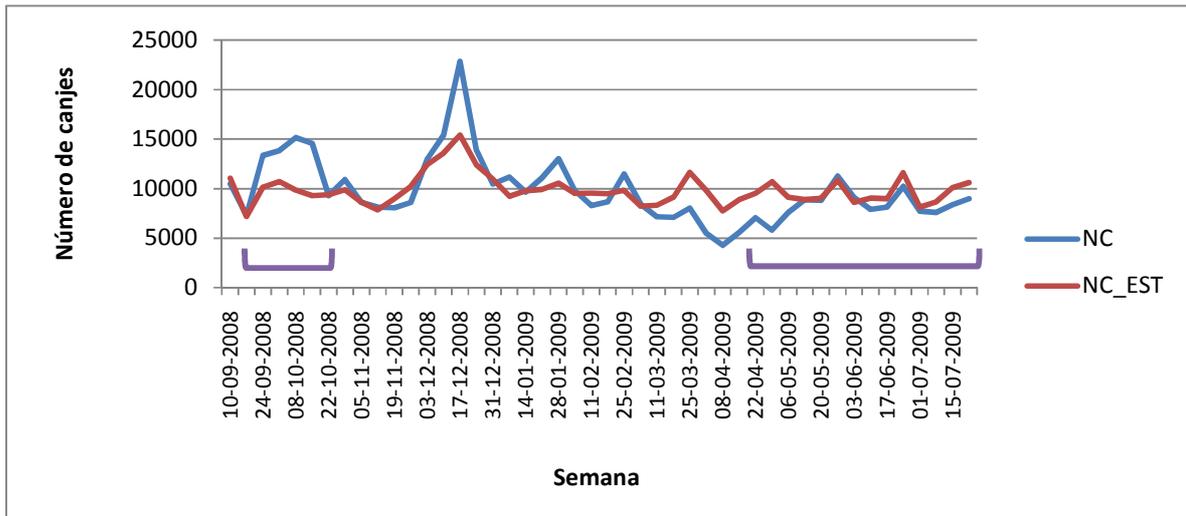
Para determinar la validez de los modelos en cuanto a capacidad predictiva, se realizará el pronóstico del número de canjes de las semanas posteriores a la calibración.

8.4.1.5.1 Modelo 1

El período de pronóstico del modelo se extiende del 10 de septiembre del 2008 al 22 de julio del 2009.

El ajuste de la estimación puede verse en el siguiente gráfico:

Gráfico 15: Ajuste del modelo 1 en el período de pronóstico



Del gráfico se desprende que la estimación subestima el número de canjes durante el primer período promocional, para luego volver a ajustarse, manteniéndose de acuerdo a lo estimado en las semanas posteriores y durante el segundo período promocional.

A continuación se presentan las medidas de error para el modelo 1 propuestas en la sección 7.1.4.

Tabla 14: Medidas de error Modelo 1 en el período de pronóstico

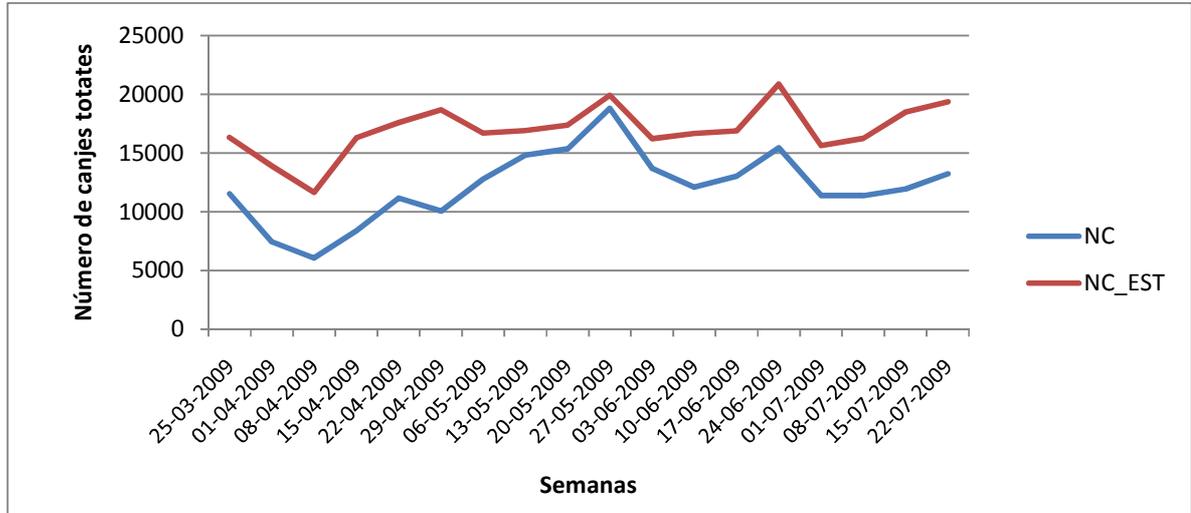
MAPE	21,59%
NMSE	55,89%

Se observa un incremento en el error porcentual absoluto promedio, el cual se duplica durante el período de pronóstico.

8.4.1.5.2 Modelo 2

El período de pronóstico de este modelo se extiende desde el 25 de marzo del 2009 al 22 de julio del 2009.

Gráfico 16: Ajuste del modelo 2 en el período de pronóstico



El gráfico permite identificar que el modelo 2 sobreestima el número de canjes del período, lo que puede verse reflejado en el incremento de las medidas de error presentadas en la

Tabla 15:

Tabla 15: Medidas de error Modelo 2 en el período de pronóstico

MAPE	45,92%
NMSE	312,66%

A partir de las medidas de error y de los gráficos presentados en la sección anterior, es posible concluir que el modelo 1 se ajusta mejor al comportamiento de la serie número de canjes totales en los períodos de calibración y validación.

A través de ambos modelos fue posible captar el incremento en el número de canjes durante el período promocional, en el modelo 2 a través de la variable promoción y en el modelo 1 comparando el número de canjes reales y el estimado por el modelo, que presentaron una diferencia de más de 7.300 canjes durante las semanas de promoción y post-promoción.

En el pronóstico del segundo período promocional los modelos presentaron grandes diferencias. El modelo 2, que incorpora en la estimación el efecto de la promoción, tendió a sobreestimar el número de canjes, mientras que el modelo 1 presentó un mejor ajuste en la predicción, lo que indica que el nivel de canjes durante este período se mantuvo de acuerdo a lo esperado para una semana normal. Este efecto se debe a que a partir del segundo período promocional la modalidad de canjes que combina puntos con pesos fue incorporada de manera permanente como forma de

canje por el programa, restando urgencia a utilizar la promoción y haciendo que el número de canjes totales en el largo plazo se mantuviese de acuerdo a lo esperado en una semana regular.

Sin embargo, es importante considerar que éste fue el efecto en el número de canjes agregado por niveles, y que puede ser no representativo para todos los niveles, por lo que resulta importante estimar los canjes por niveles para determinar el impacto de la promoción en cada uno de ellos.

A continuación se presenta la estimación del número de canjes para el primer nivel de canjes.

8.4.2 Estimación número de canjes para el nivel 1

El procedimiento para estimar el orden de los modelos y las variables externas es similar al desarrollado en las secciones 8.4.1.1 y 8.4.1.2. Por esta razón, en esta sección se presentará un resumen de los resultados obtenidos.

Al igual que para la estimación de canjes totales se desarrollaron 2 modelos:

- Modelo 1: Calibración: 05-01-2005 al 03-09-2008
 Validación: 10-09-2008 al 22-07-2009
- Modelo 2: Calibración: 05-01-2005 al 18-03-2009
 Validación: 25-03-2009 al 22-07-2009

8.4.2.1 Estimación de los parámetros

Los parámetros obtenidos para este nivel canje son los siguientes:

Tabla 16: Parámetros del Modelo 1 y 2 para el nivel 1

	Modelo 1	Modelo 2
AR1	0,73	0,69
AR1 estacional	0,42	0,38
PE_1	8,32E-06	5,60E-06
PROMO		3.740,16
PATRIAS	- 3.060,28	- 3.092,30
NAVIDAD	2.413,49	3.624,04
SEMSANTA	- 1.542,30	- 1.375,41
fdm_1	1.304,37	1.631,61
fdm_2	1.390,27	1.717,82
fdm_3	2.037,29	2.064,70
fdm_4	1.448,82	1.454,35
fdm_5	1.464,91	1.485,64
fdm_6	2.250,17	2.171,74
fdm_7	2.283,27	2.313,12
fdm_9		2.107,51
fdm_10	2.122,87	2.346,35
fdm_12	1.560,08	1.218,64
constante	8.297,62	8.840,42

Es posible observar que en el modelo 2, la variable que indica si una semana es o no de promoción resulta ser significativa y tiene un efecto positivo en el número de canjes.

8.4.2.2 Ajuste de los modelos

La siguiente tabla refleja el ajuste de los modelos durante el período de calibración:

Tabla 17: Ajuste de los Modelos 1 y 2 durante el período de calibración

	Modelo 1	Modelo 2
R ²	67,49%	68,90%
R	82,15%	83,01%
MAPE	11,19%	11,13%

A partir de la información anterior es posible decir que ambos modelos son capaces de explicar un porcentaje de varianza similar de los datos y que su error promedio es similar.

Gráficamente el ajuste de los modelos puede verse a continuación:

Gráfico 17: Ajuste del Modelo 1 para el nivel 1

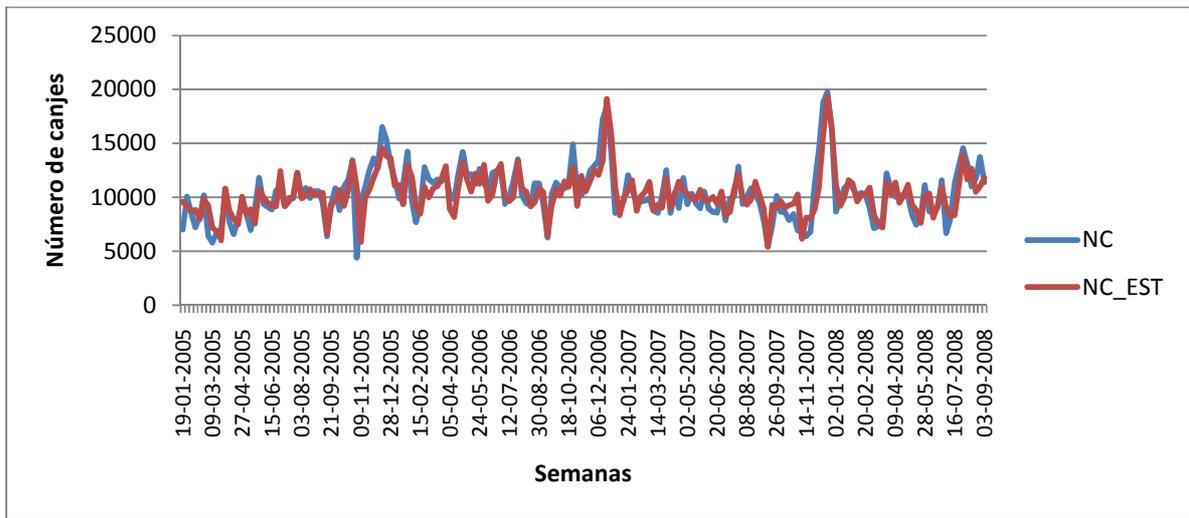
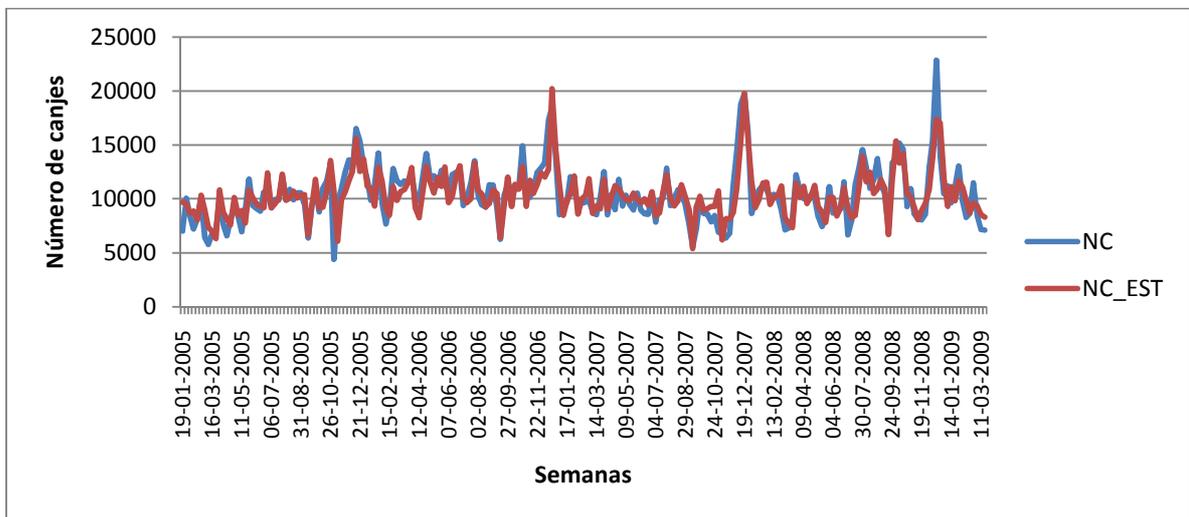


Gráfico 18: Ajuste del Modelo 2 para el nivel 1

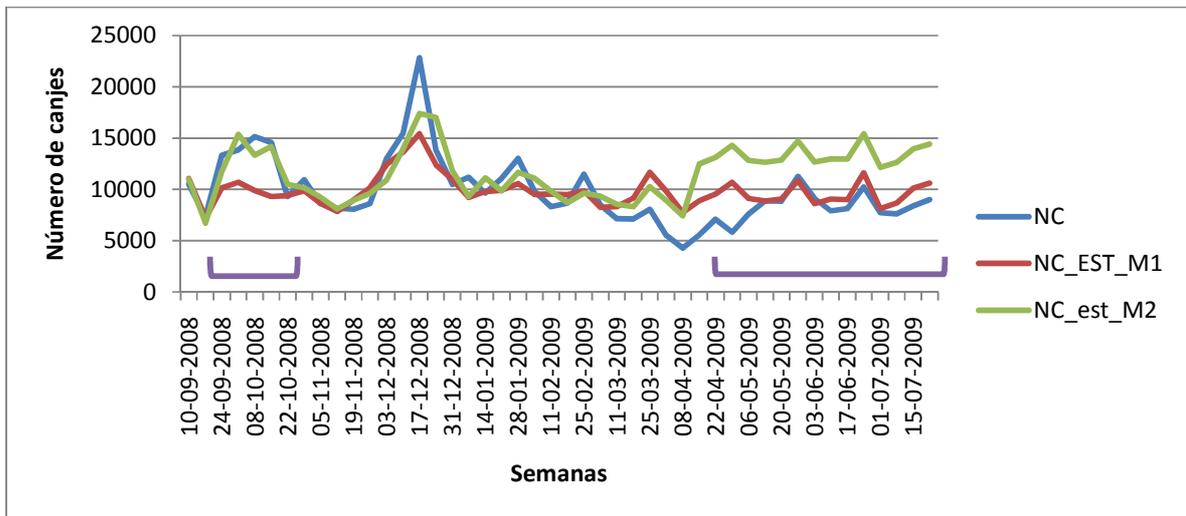


Donde se observa que ambos modelos se ajustan bien a la serie número de canjes semanales del nivel 1 durante el período de calibración.

8.4.2.3 Pronóstico de los modelos

Para determinar la validez de los modelos en cuanto a capacidad predictiva, se realizará el pronóstico del número de canjes de las semanas posteriores a la calibración.

Gráfico 19: Ajuste de los Modelos 1 y 2 para el nivel 1



Se observa que el modelo 1 subestima el número de canjes durante el primer período promocional y se ajusta al nivel de canje real en las semanas posteriores y en el segundo período promocional, mientras que el modelo 2 tiende a sobreestimar los canjes durante el período de calibración.

Tabla 18: Medidas de error Estimación número de canje Nivel 1

	Modelo 1	Modelo 2
MAPE	19,18%	64,85%
NMSE	54,31%	869,87%

A partir de la información anterior, es posible determinar que el modelo 1 presenta un mejor ajuste que el modelo 2 en el período de pronóstico.

Al igual que en el número de canjes totales, el efecto de la primera promoción, incrementa el número de canjes durante el período en aproximadamente 13.000 canjes. Sin embargo, en la segunda promoción los canjes se mantuvieron de acuerdo a los esperado para una semana de canje regular.

El desarrollo de los modelos para los siguientes niveles se encuentran en el Anexo 3: Estimación de canjes por niveles

8.4.3 Conclusiones de las estimaciones del número de canjes por niveles

A través de la estimación del número de canjes por niveles, fue posible identificar que el efecto que tuvo la promoción en cada uno de los niveles fue distinto al obtenido en la estimación agregada.

Para el nivel 2, la primera promoción tuvo un efecto negativo, disminuyendo el número esperado de canjes durante esas semanas. A diferencia de la estimación a nivel agregado y para el primer nivel, el modelo que se ajustó mejor a los datos durante el período de pronóstico fue el 2, que a pesar de ser calibrado con datos pertenecientes a la primera promoción, no tiene entre sus variables significativas a las variables dummies asociadas a la promoción.

El primer período promocional no tuvo efectos para el nivel 3, sin embargo, los canjes experimentaron una baja las semanas posteriores a ésta (efecto que es capturado por la variable post-promoción en el modelo 2), lo que puede deberse a una anticipación de canje, es decir, los clientes que usaron la promoción pudieron acceder a un nivel superior de canje o a realizar una mayor cantidad de canjes en niveles inferiores. Durante la segunda campaña promocional, ambos modelos presentaron un ajuste con la serie número de canjes del nivel 3, por lo que la cantidad de canjes se mantiene de acuerdo a lo esperado para una semana de canje regular.

El efecto de la primera promoción en el nivel 4 fue incrementar los canjes realizados durante esas semanas, efecto que fue seguido por una caída durante las semanas post-promocionales. El modelo que mejor pronosticó los canjes durante el período de la segunda promoción fue el modelo 1, lo que indica que los canjes se mantuvieron de acuerdo a lo esperado para una semana de canje regular.

Los últimos cuatro niveles (5, 6, 7 y 8) presentaron con la primera campaña un aumento en el número de canjes. Posteriormente, con la segunda se produjo un aumento más pequeño, pero sostenido en las semanas.

Es importante considerar que los primeros niveles, que son los que concentran un mayor número de canjes semanalmente, tienen menor porcentaje de error que aquellos con menor participación. Esto se debe a que mientras menor sea el número de canjes, mayor es el porcentaje de error asociado a una unidad.

A partir del análisis anterior es posible entonces identificar que si bien el efecto global de la promoción fue aumentar el número de canjes, no todos los niveles se vieron beneficiados de la misma forma.

La segunda promoción no tuvo mayores efectos, manteniendo los canjes de acuerdo a lo esperado.

8.5 Modelo proporcional de Hazard:

El modelo proporcional de Hazard será usado para estimar si la promoción generó aceleración de compra durante su período de vigencia.

Por este motivo, se estimará la probabilidad de compra de un cliente en una determinada semana dado que ha pasado un tiempo t desde que el cliente no realiza compra.

Para llevar esto a cabo, fue necesario seleccionar para cada nivel todo el proceso de acumulación de puntos hasta llegar a un canje, desde el canje previo (en caso que el cliente ya haya canjeado previamente) o desde su primera compra. La base entonces incorpora semanas en las que el cliente realizó compras y semanas en las que no, las que sirvieron para estimar la línea base del modelo, que captura patrones entre compras de los clientes y que es la que se ve modificada por covariables que incrementan o disminuyen la probabilidad que el cliente compre una determinada semana.

Las covariables a considerar en el modelo serán entonces las siguientes:

- Frecuencia de compra del cliente
- Dist – canje, variable que relaciona la distancia relativa al nivel con el número de canjes.
- Variables dummies
 - Pre-promoción
 - Promoción
 - Post-promoción
 - Navidad
 - Semana santa
 - Fin de mes

A continuación se detalla el desarrollo del modelo para el primer nivel de canje.

8.5.1 Estimación de los parámetros del modelo para el nivel 1

Se cuenta con una base de 601.565 observaciones, las que corresponden a la información transaccional semanal de 6.000 clientes que realizaron al menos un canje en el primer nivel durante el período de enero del 2005 a diciembre del 2008.

En la base, 240.220 observaciones corresponden a compras, aproximadamente el 39% de las observaciones.

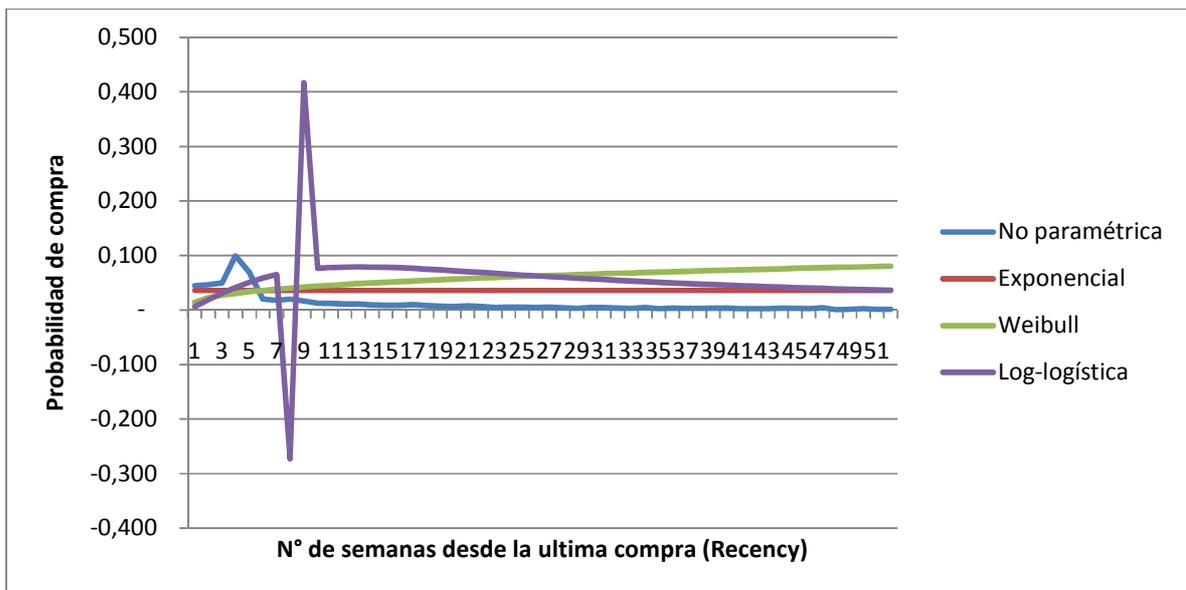
Los parámetros son estimados para la línea base paramétrica y no paramétrica maximizando la función de verosimilitud, obteniéndose los siguientes parámetros:

Tabla 19: Parámetros del modelo proporcional de Hazard para el nivel 1

	NP	EXP	W	LL
Frecuencia de compra	3,537	3,804	4,925	6,227
Pre-promoción	- 0,059			- 0,077
Promoción	- 0,175	- 0,159	- 0,220	- 0,258
Post-promoción	- 0,210	- 0,238	- 0,321	- 0,315
Fin de mes	- 0,103	- 0,126	- 0,115	- 0,118
Navidad	0,182	0,149	0,173	0,394
Semana santa	- 0,052	- 0,053	- 0,030	- 0,075
Dist – Canje	- 0,476	- 0,691	- 0,890	
ALPHA			1,359	1,987
GAMMA		0,036	0,044	0,080

La forma de la línea base paramétrica se encuentra dada por los parámetros estimados α y γ . En la Tabla 20 se exponen las líneas bases del modelo proporcional de Hazard para el nivel 1.

Tabla 20: Líneas base del nivel 1



Dependiendo del signo del parámetro y como esté definida la covariable es el efecto que ésta tendrá en la probabilidad de compra.

Así, dado que el parámetro asociado a la frecuencia de compra es positivo y a que esta toma valores entre 0 y 1, donde 1 indica que el cliente compra todas las semanas y 0 que el cliente no compra nunca, el efecto que un cliente compre más seguido aumenta la probabilidad de comprar en una determinada semana, lo mismo ocurre con la variable navidad.

El signo negativo asociado al parámetro de la variable Dist-canje indica que a medida que esta variable toma un valor mayor, menor es la probabilidad de canjear. Sin embargo, si consideramos que Dist-canje toma valores entre 0 y 1 donde 0 indica que ya se alcanzó el puntaje para realizar un canje en el nivel y 1 que aún no se ha acumulado puntaje, valores cercanos a 0 incrementarían la probabilidad de canjear, lo que es un claro indicio de aceleración de compra, que es una de la hipótesis que se deseaba testear.

La variable promoción está asociada a un parámetro negativo, por lo que estar en una semana de promoción disminuye la probabilidad de compra, efecto que podría explicarse pues para canjear usando la promoción los clientes debían cancelar una diferencia en pesos correspondiente a la disminución de puntaje, por lo que podrían haber usado dinero que en situaciones regulares hubiesen usado para realizar compra.

8.5.2 Ajuste de modelo para el nivel 1

La siguiente tabla presenta el ajuste del modelo proporcional de Hazard para las cuatro líneas bases consideradas en el estudio:

Tabla 21: Ajuste del modelo proporcional de Hazard para el nivel 1

	NP	EXP	W	LL
N° de observaciones	601.565	601.565	601.565	601.565
LL	- 2.989.315,59	-509.340,29	-489.620,04	-466.328,38
N ° de parámetros		8,0	9,0	9,0
SBS	5.978.730,3	1.018.787,0	979.359,9	932.776,5

De acuerdo al criterio de información Bayesiana, el criterio que presenta un mayor ajuste a los datos (menor SBS) es el log-logístico, sin embargo, es importante

considerar que el alto SBS del modelo no paramétrico puede deberse a la penalización por la gran cantidad de parámetros a estimar para construir la línea base y no a un mal ajuste del modelo, por lo que es necesario testear la calidad predictiva de los modelos.

La regla de decisión δ a partir del cual se determina si un cliente realiza o no compra en una semana se determinó usando los indicadores de la matriz de confusión, los que pueden verse en la siguiente tabla:

Tabla 22: Indicadores matriz de confusión (base de calibración)

	NP	Exp	W	LL
δ	18%	14%	12%	15%
PREDIC_OK	63,6%	61,4%	60,9%	59,4%
C_OK	58,1%	56,1%	58,2%	58,4%
NC_OK	67,2%	64,9%	62,6%	60,0%
C_ERROR	32,8%	35,1%	37,4%	40,0%
NC_ERROR	41,9%	43,9%	41,8%	41,6%

Al comparar estos indicadores para los distintos modelos, es posible apreciar que el modelo con mejor capacidad predictiva fue el no paramétrico, el que pese a tener el mayor valor en el criterio SBS, estima los parámetros de la línea base a partir de los propios datos.

8.5.3 Pronóstico del modelo proporcional de Hazard para el nivel 1

Luego, con el objetivo de testear los modelos usando información de la nueva campaña promocional, se pronosticó la compra en una nueva base de datos que incorpora el período de la segunda promoción (enero del 2005 a junio del 2009). Los indicadores de la calidad del modelo pueden verse en la tabla:

Tabla 23: Indicadores matriz de confusión (base de pronóstico)

	NP	Exp	W	LL
δ	62,5%	60,6%	59,6%	58,3%
PREDIC_OK	61,5%	60,1%	61,1%	59,9%
C_OK	63,2%	61,0%	58,4%	57,1%
NC_OK	36,8%	39,0%	41,6%	42,9%
C_ERROR	38,5%	39,9%	38,9%	40,1%
NC_ERROR	62,5%	60,6%	59,6%	58,3%

Es posible observar nuevamente que el modelo que presenta mayor ajuste con los datos es el no paramétrico.

El desarrollo del resto de los modelos se encuentra en el Anexo 4: Modelo proporcional de Hazard.

8.5.4 Conclusiones sobre el modelo en los niveles

Al analizar los distintos modelos para cada uno de los niveles fue posible identificar que las variables frecuencia y Dist-canje resultaron significativas para todos los niveles, por lo que independiente del nivel, mientras más frecuentemente compre el cliente, mayor probabilidad de comprar en una determinada semana. Lo mismo ocurre con la variable Dist-canje, mientras más cerca se encuentra un cliente al nivel de puntos requerido, mayor probabilidad de realizar una compra que le permita acercarse aún más al nivel, lo que indica que la estructura del programa de fidelización de la empresa genera aceleración de compra en sus clientes.

Por otra parte, las variables asociadas a la promoción resultaron ser significativas para los cuatro primeros niveles de compra, y en general su efecto fue negativo, es decir, estar en semana de promoción o post-promoción disminuía la probabilidad de compra de los clientes. Esto podría deberse a que aquellos que canjearon usando la promoción destinaron el dinero de la compra del período a pagar el monto requerido para canjear. Es importante además destacar que a medida que el nivel es más bajo, menor es el porcentaje de compras sobre el total de observaciones, lo que podría influir de alguna forma en el coeficiente negativo del parámetro para estos niveles.

El efecto de las variables dummies incorporadas en los modelos (fin de mes, navidad, semana santa) fue consistente entre los distintos niveles para los que resultaron significativas, así la navidad aumenta la probabilidad de comprar en una semana, mientras que fin de mes y semana santa la disminuyen.

Los modelos fueron capaces de predecir de manera correcta aproximadamente un 60% de las observaciones, lo que se consideró satisfactorio, pues resulta difícil predecir semanalmente si el cliente realizará o no compra usando sólo información transaccional y demográfica de éstos.

8.6 Modelos de predicción de canje

Para determinar quiénes de los clientes que canjearon durante los períodos promocionales lo hicieron debido a la promoción y no hubiesen realizado canje de haber sido una semana normal, se planteó un modelo logit para estimar la probabilidad de canje.

Las variables a considerar en el desarrollo del modelo se basan en el análisis realizado a la serie número de canjes de la sección 8.4.1.1, incorporando la variable Dist-canje que incorpora información del nivel de puntos del cliente.

Así, las variables a incorporar en el modelo son las siguientes:

- Fin de mes
- Semana santa
- Fiestas Patrias
- Navidad
- Dist-canje
- Recency de canje
- Frequency de canje
- Variables promocionales

Para el desarrollo de estos modelos, se seleccionaron las observaciones correspondientes a los canjes realizados por los clientes pertenecientes a las muestras de sus respectivos niveles y una cantidad igual de observaciones correspondientes a no canjes.

El modelo a usar es un logit binario, donde la probabilidad de compra está dada por:

$$P(Y = 1) = \frac{e^{\bar{\beta} \cdot \bar{X}}}{1 + e^{\bar{\beta} \cdot \bar{X}}} \quad (44)$$

8.6.1 Estimación de los parámetros para el nivel 1

Los parámetros obtenidos para el primer nivel fueron los siguientes:

Tabla 24: Parámetros del modelo de predicción de canjes

	B	E.T.	Wald	Sig.	Exp(B)
R_canje	0,004	0,000	90,251	0,000	1,004
F_Canje	- 0,795	0,322	6,090	0,014	0,451
Pre-promoción	0,715	0,192	13,875	0,000	2,043
Promoción	2,781	0,244	129,614	0,000	16,136
Post-promoción	0,758	0,215	12,444	0,000	2,134
Fin de mes	0,091	0,041	4,890	0,027	1,095
Fiestas patroas	- 0,443	0,141	9,946	0,002	0,642
Navidad	0,686	0,124	30,763	0,000	1,986
Semana santa	- 0,296	0,136	4,731	0,030	0,744
Dist-canje	- 11,772	0,220	2.870,924	-	0,000
Constante	0,713	0,034	447,163	0,000	2,040

A partir de ellos, es posible interpretar que:

- Mientras mayor es el recency de canje del cliente, mayor será su probabilidad de canje.
- A menor frecuencia de canje, mayor es la probabilidad de canje
- A medida que el cliente se encuentre más cerca del nivel de canje (Dist-canje cercano a 0), aumentan las probabilidades de canje

Con respecto a las variables promocionales, es posible interpretar a partir de los coeficientes que estar en una semana promocional aumenta las posibilidades de canje, lo que es consistente con los resultados obtenidos en la estimación de demanda realizada en la sección 8.4.

8.6.2 Ajuste del modelo para el nivel 1

La siguiente tabla presenta el ajuste del modelo:

Tabla 25: Ajuste del modelo predicción de canje para el nivel 1

OBSERVACIONES	21.416,0
LOG-LIKEHOOD	- 9.846,483
PARÁMETROS	11,0
SBS	19.802,7
L(0)	- 14.872,5
R^2 COX & SNELL	37,5%
R^2 NAGELKERKE	49,9%

Las medidas de R-cuadrado de Cox y R-cuadrado de Nagelkerke, son medidas que permiten determinar el porcentaje de varianza captado por el modelo. En modelos logísticos, es preferible usar el R-cuadrado de Nagelkerke, pues el R-cuadrado de Cox no toma valores entre 0 y 1, por lo que interpretar los coeficientes resulta complejo.

Su fórmula es:

$$R_N^2 = \frac{1 - \left[\frac{L_o}{L_{\text{modelo}}} \right]^{(2/N)}}{1 - L_o^{(2/N)}} \quad (45)$$

Donde:

- L_o : verosimilitud del modelo con solo una constante
- L_{modelo} : verosimilitud del modelo a usar.
- N: tamaño de la muestra

Al igual que para el modelo proporcional de Hazard, la regla de decisión δ , se determinó usando los indicadores de la matriz de confusión. Luego es posible evaluar la capacidad predictiva del modelo en los datos con los cuales se calibró el modelo usando los indicadores:

Tabla 26: Indicadores matriz de confusión (base de calibración)

Logit Nivel 1	
δ	64%
PREDIC_OK	79,7%
C_OK	89,2%
NC_OK	70,3%
C_ERROR	29,7%
NC_ERROR	10,8%

A partir de la tabla anterior se desprende que el modelo predice bien el 89% de los canjes, y el 70,3% de los no canjes, por lo que el porcentaje de predicciones correctamente realizadas por el modelo es de un 79,7%.

8.6.3 Pronóstico del modelo de predicción de canjes

Para evaluar la capacidad predictiva del modelo, se utilizó una base que incluía el proceso de acumulación de puntos hasta el canje. El número de observaciones de la base para el nivel 1 es de 601.565, donde el número de canjes y no canjes se detalla a continuación:

Tabla 27: N° de canjes y no canjes de la base

NO CANJES	591.254
CANJES	10.311

La capacidad predictiva del modelo se evaluó a través de los indicadores de la matriz de confusión:

Tabla 28: Indicadores matriz de confusión en (base de pronóstico)

Logit Nivel 1	
δ	69,3%
PREDIC_OK	89,5%
C_OK	68,9%
NC_OK	31,1%
C_ERROR	10,5%
NC_ERROR	69,3%

Donde se observa que las proporciones de canjes y no canjes correctamente identificados de la base son similares a los de la base con la que se calibró el modelo, sin embargo, como en esta base el número de no canjes es mayor, la proporción de predicciones correctas varía.

El desarrollo del resto de los modelos se encuentra en el Anexo 5: Modelo de predicción de canjes.

Luego, es posible usando estos modelos determinar el porcentaje de clientes que no hubiese canjeado durante el período de no haber estado en vigencia la promoción. Para ésto, se estimó la probabilidad de canje de aquellos clientes que canjearon usando la promoción en el primer período y luego se calculó los porcentajes de casos en que el modelo predijo que canjearían y el porcentaje de casos en que el modelo dijo que no lo harían.

Tabla 29: Porcentaje de canjes y no canjes predichos para clientes que usaron la promoción en período 1

Nivel	1	2	3	4	5	6	7
No	26%	48%	35%	59%	29%	77%	100
Sí	74%	52%	65%	41%	71%	23%	0

Posteriormente, se estimaron nuevamente las probabilidades de canje, pero considerando a la semana como una normal, es decir, las variables asociadas a la promoción con valor 0, y se calcularon nuevamente los porcentajes de casos predichos correcta e incorrectamente por el modelo.

Tabla 30: Porcentaje de canjes y no canjes predichos para clientes que usaron la promoción en período 1 sin incorporar variables promocionales.

Nivel	1	2	3	4	5	6	7
No	79%	75%	69%	84%	96%	77%	100
Sí	21%	25%	31%	16%	4%	23%	0

De las tablas se observa que para el nivel 1, el modelo predijo que el 70% de los clientes canjearía durante la semana de promoción, sin embargo, si esa semana no hubiese sido una semana de promoción, el modelo pronosticó que el solo el 20% de esos clientes hubiese canjeado.

En la sección 11.6 se presenta el cálculo para los clientes que canjearon en el período promocional 2.

De esta forma, es posible afirmar que más del 60% de los clientes que usó la promoción no hubiese canjeado en ese período de no haber estado la promoción.

8.7 Análisis descriptivo de los clientes que usaron la promoción

Durante el primer período en que se desarrolló la promoción Puntos con pesos (2 al 19 de octubre del 2008), 54.547 clientes realizaron canje, de los cuales, 13.453 canjearon usando la promoción (24,6% de los canjeadores del período).

El 52% de las personas que canjeó durante el período usando la promoción no habían realizado canjes durante los dos últimos años, mientras que el 68% de éstas no había canjeado durante el último año.

En el período, se realizaron 58.454 canjes, de los cuales el 23,5% (13.751 canjes) se realizó a través de la promoción. A continuación se presenta una tabla donde se observa la cantidad de canjes realizados con promoción por niveles:

Tabla 31: Canjes promoción Puntos con pesos por nivel, primer período

Nivel	N° de canjes	% de canjes
1	11.192	81,4%
2	1.222	8,9%
3	741	5,4%
4	469	3,4%
5	111	0,8%
6	14	0,1%
7	2	0,0%
Total	13751	100,0%

El 17 de abril del 2009 fue relanzada esta promoción e incorporada de manera permanente como una modalidad de canje del programa de fidelización. Desde el inicio de este segundo período, hasta el 31 de julio del 2009 se realizaron 202.603 canjes, de los cuales 12% (24.342 canjes) fueron realizados usando la promoción.

Durante el segundo período canjearon 169.325 personas, de las cuales 23.955 (14,15%) canjearon usando la promoción. De estas 23.955 personas, un 58,63% no había realizado canjes durante el último año, mientras que el 39,84% no lo había hecho en los últimos 2 años.

Tabla 32: Canjes promoción Puntos con pesos por nivel, segundo período

Nivel	N° de canjes	% de canjes
1	15.453	63%
2	5.146	21%
3	1.685	7%
4	1.113	5%
5	704	3%
6	176	1%
7	62	0%
8	3	0%
Total	24.342	100%

Dado que la mayoría de los clientes que usó la promoción no había realizado canjes al menos un año antes de ésta, se plantea la pregunta de cuál es el porcentaje de clientes que canjea una determinada semana y que no ha realizado canjes durante el último año. Para ello, se calculó semana a semana el porcentaje de clientes que canjeó sin haberlo realizado durante el año previo. A partir de ese análisis, fue posible determinar que antes del primer período promocional, del total de los clientes que canjeaba durante un período, aproximadamente un 46,09% no había canjeado durante el último año. En el primer período de promoción, este número alcanzó el 53,35% (promedio de las 3 semanas de su extensión), mientras que en período promocional dos, este alcanzó un 43,04% (promedio de las 14 semanas).

A continuación se presenta un análisis descriptivo donde se comparará, para cada período promocional, a los clientes que canjearon usando Puntos con pesos con aquellos clientes que canjearon normalmente.

8.7.1 Sexo

Tabla 33: Porcentaje de hombres y mujeres que canjearon en período promocional

Sexo	Período 1		Período 2	
	Usa promoción	No usa promoción	Usa promoción	No usa promoción
F	58,43%	48,53%	53,51%	49,46%
M	41,56%	51,47%	46,49%	50,54%

En la tabla anterior se observa que la proporción entre hombres y mujeres de los clientes que usan la promoción favorece al sexo femenino en los dos períodos de promoción, mientras que aquellos clientes que no usaron la promoción se encuentran equiparados.

8.7.2 Edad

Gráfico 20: Proporción de clientes que usaron la promoción en el período 1 por rango de edad

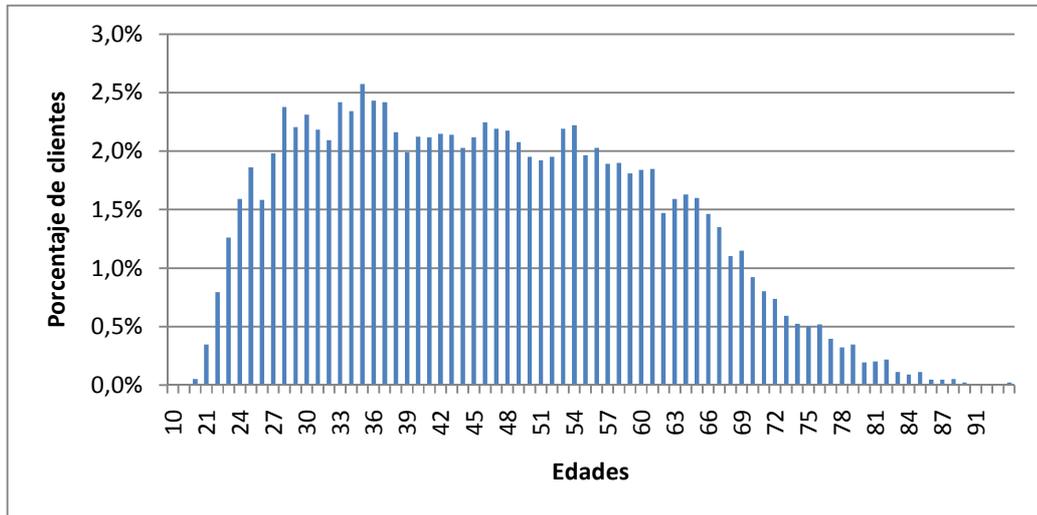
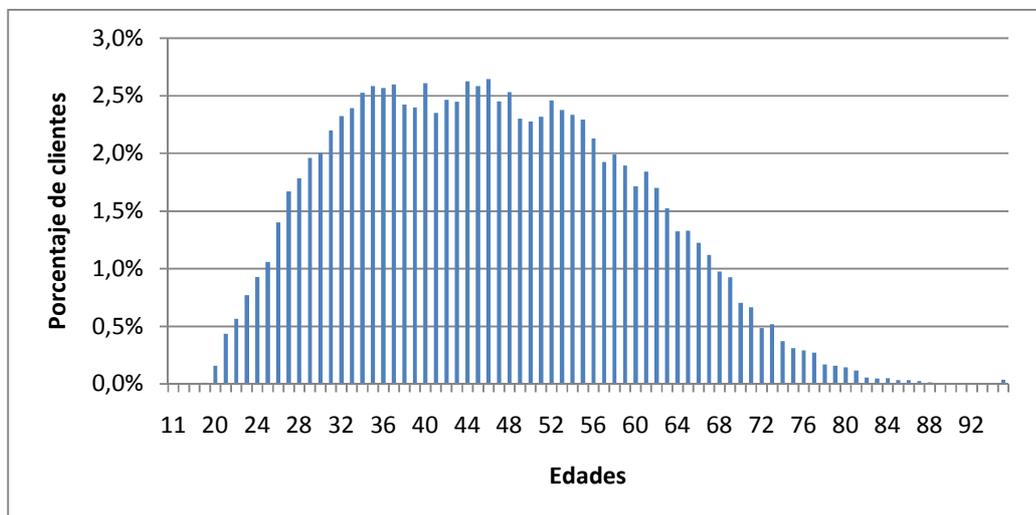


Gráfico 21: Proporción de clientes que no usaron la promoción en el período 1 por rango de edad



En el período 1, no se observan diferencias significativas en el rango de edades, sin embargo, hay un mayor porcentaje de clientes que usó la promoción entre 24 y 34 años y entre 58 y 66 que aquellos que no usan la promoción, y disminuye el porcentaje de canjeadores entre los 32 y 52 años.

Gráfico 22: Proporción de clientes que usaron la promoción en el período 2 por rango de edad

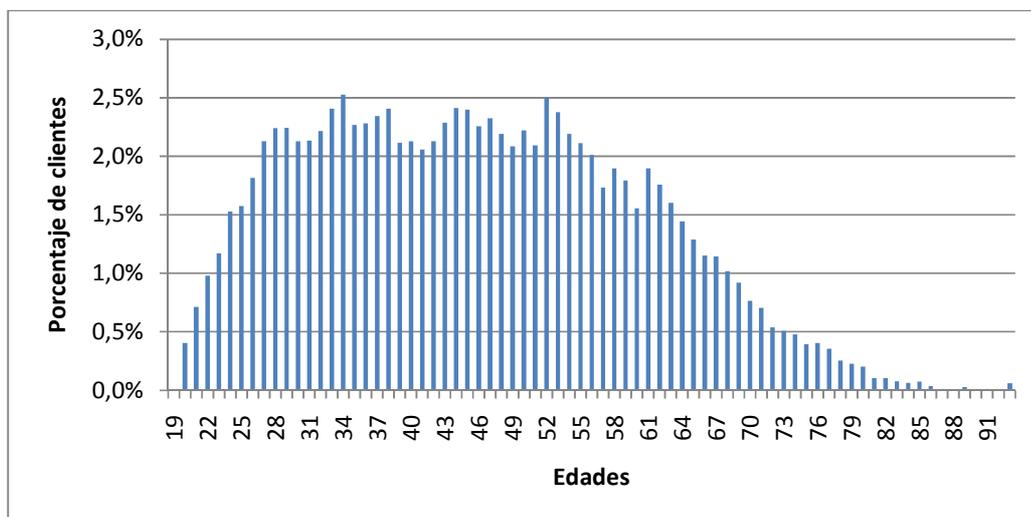
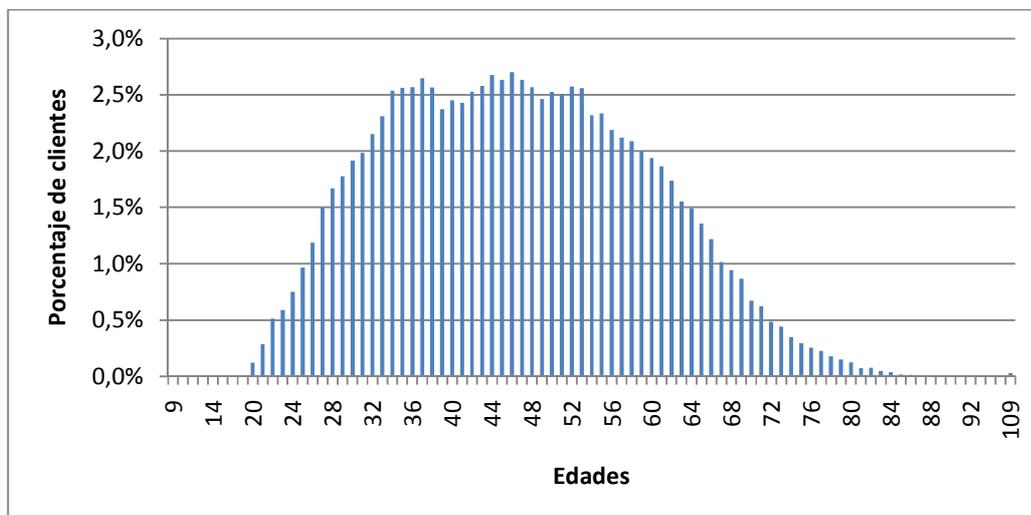


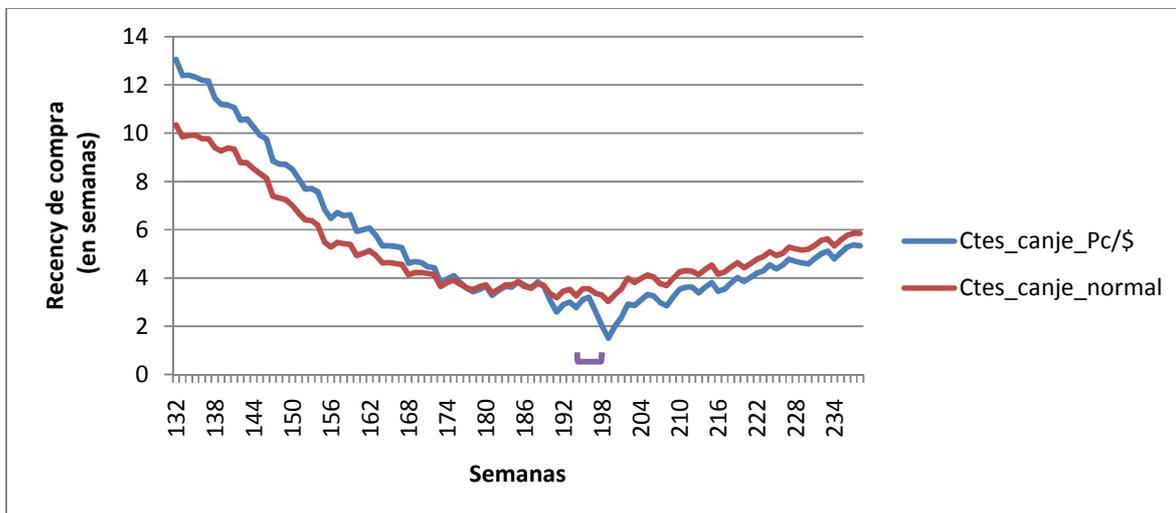
Gráfico 23: Proporción de clientes que no usaron la promoción en el período 1 por rango de edad



En el período 2 sucede algo similar, no hay diferencias significativas en el rango de las edades, sin embargo, hay un mayor porcentaje de clientes que usó la promoción entre 27 y 34 años y entre 55 y 64 que aquellos que no usan la promoción.

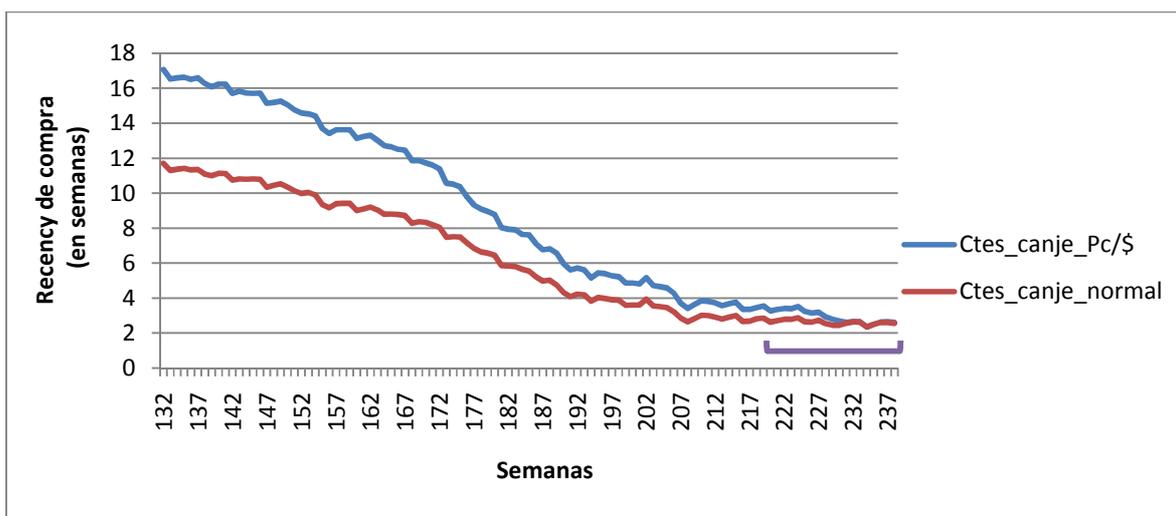
8.7.3 Recency de compra

Gráfico 24: Recency de compra promedio de clientes canjeadores del período 1



Se observa que los clientes que canjearon usando promoción³ presentaban, en promedio, un recency de compra mayor en el período previo a la promoción que aquellos clientes que no la usaron. Sin embargo, posterior a la promoción, éste se mantuvo menor que el de los clientes que realizaron canje normal en el período. Esto puede ser efecto de la promoción, la que a través de la recompensa, incentivó a los clientes a participar más activamente en el programa en períodos posteriores.

Gráfico 25: Recency de compra promedio de clientes canjeadores del período 2



³ De aquí en adelante se denomina Ctes_canje_Pc/\$ a los clientes que canjearon en período promocional usando la promoción y Ctes_canje_Normal al de los clientes que realizaron canje normal en el período de la promoción.

Se observa que aquellos clientes que canjearon usando promoción tenían un recency de compra mayor en el período previo a la promoción, y que a medida que comenzó el segundo período se fue acortando hasta llegar a ser casi iguales con los clientes que no la usaron.

8.7.4 Frequency de compra

Gráfico 26: Frecuencia de compra promedio de clientes canjeadores del período 1

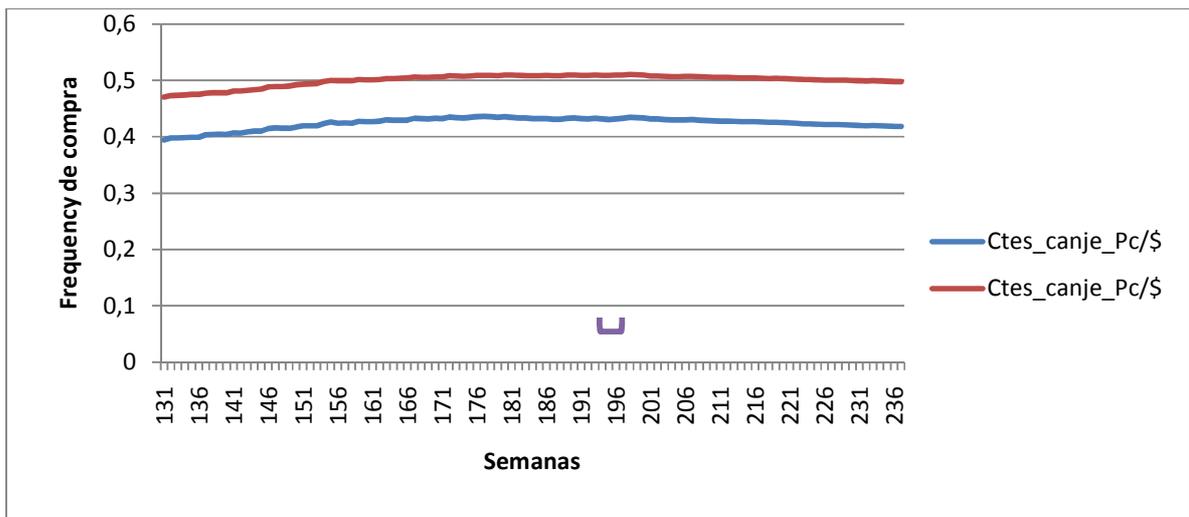
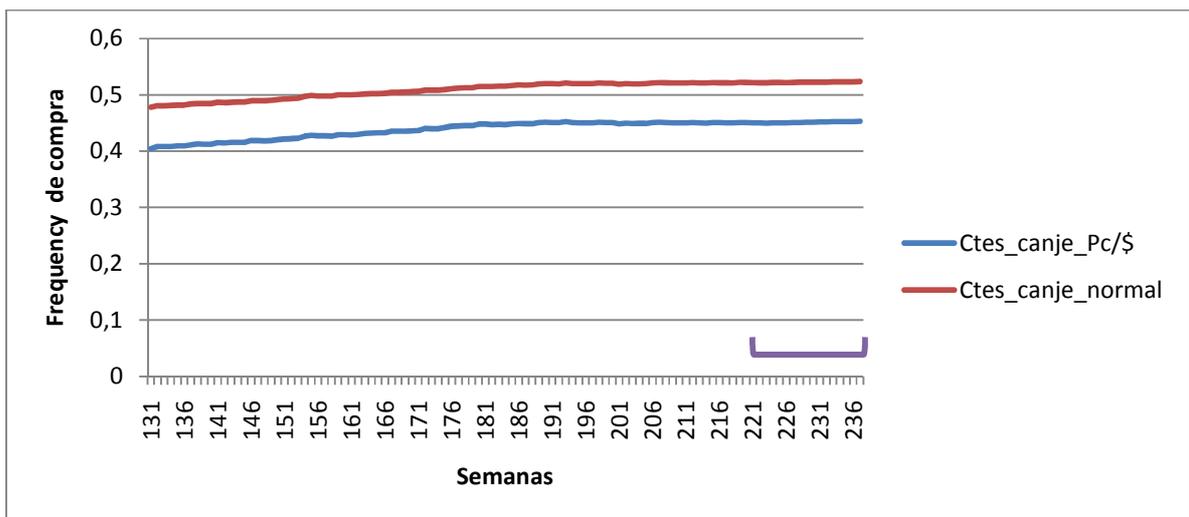


Gráfico 27: Frecuencia de compra promedio de clientes canjeadores del período 2



Se observa, que aquellos clientes que no usaron la promoción compran más frecuentemente que aquellos que la usaron.

8.7.5 Monto de compra

Gráfico 28: Monto de compra promedio de clientes canjeadores del período 1

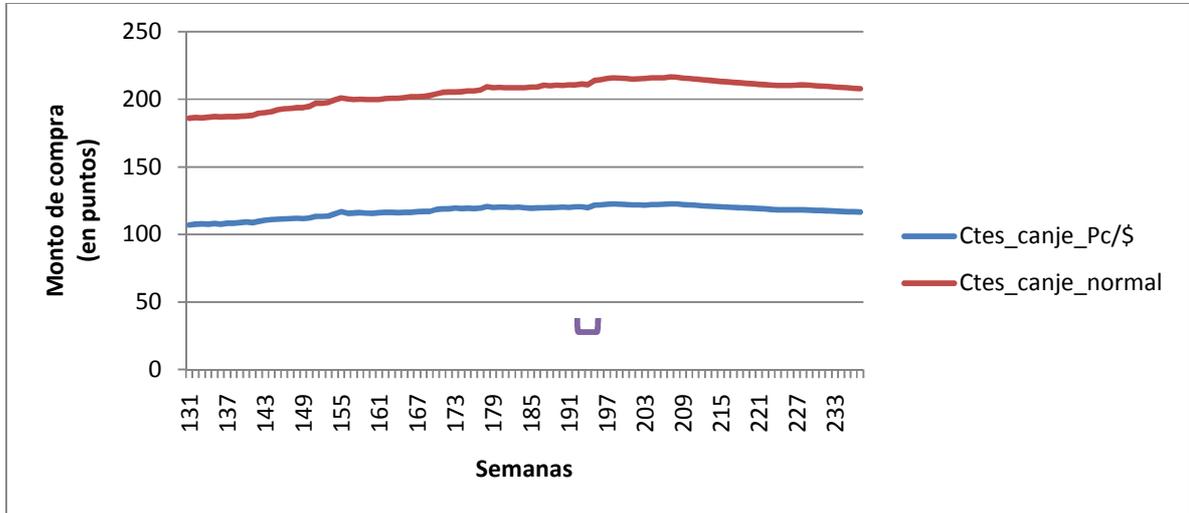
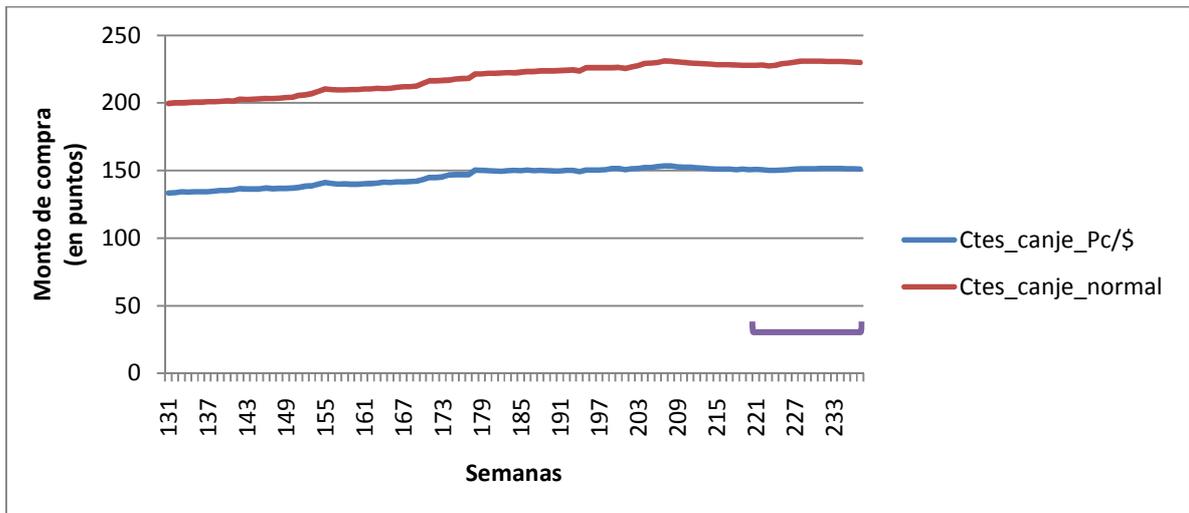


Gráfico 29: Monto de compra promedio de clientes canjeadores del período 2



Se observa para ambos períodos que aquellos clientes que canjearon usando promoción tienen menores montos de compra promedio semanal que aquellos que no la usaron. Esta diferencia es en promedio es de 50 puntos que para el caso de compra se traduce en un monto de \$6.000.

8.7.6 Recency de canje

Gráfico 30: Recency de canje promedio de clientes canjeadores del período 1

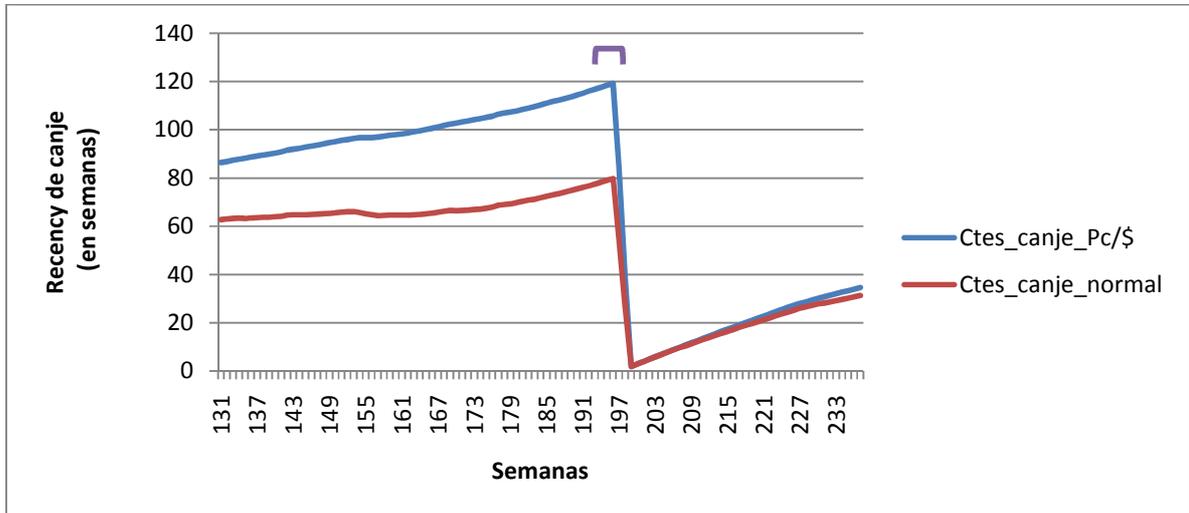
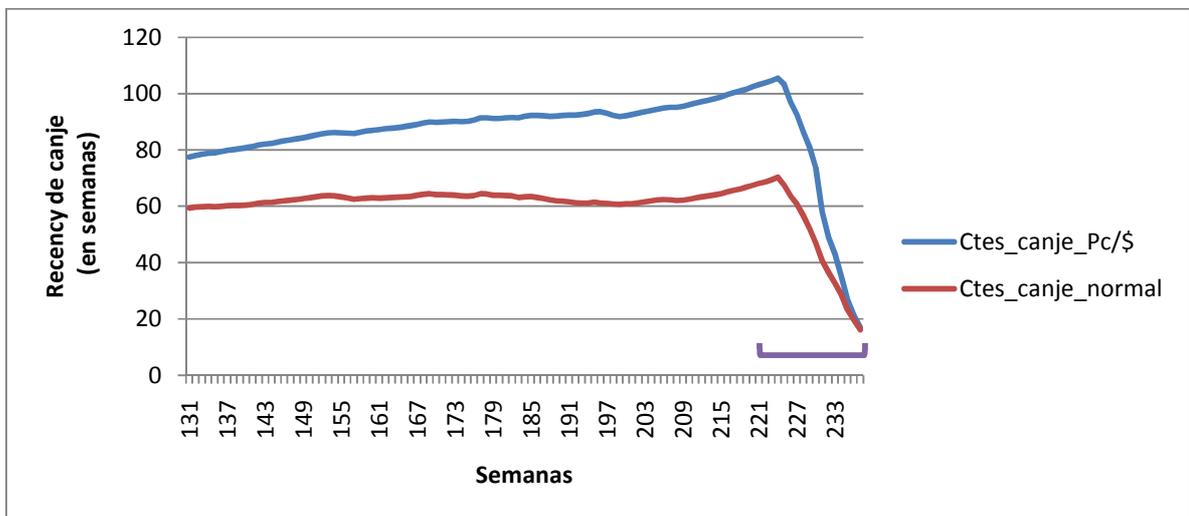


Gráfico 31: Recency de canje promedio de clientes canjeadores del período 2



De los gráficos es posible ver que aquellos clientes que canjearon usando Puntos con pesos tenían, previo a la promoción, un recency de canje mucho mayor a aquellos clientes que no usaron la promoción. Se observa que el recency promedio de los aquellos clientes que usaron la promoción es superior a las 80 semanas, lo que indica que éstos canjean menos de una vez al año, que es consecuente con los resultados obtenidos en la sección 8.7.

8.7.7 Frequency de canje

Gráfico 32: Frecuencia de canje promedio de clientes canjeadores del período 1

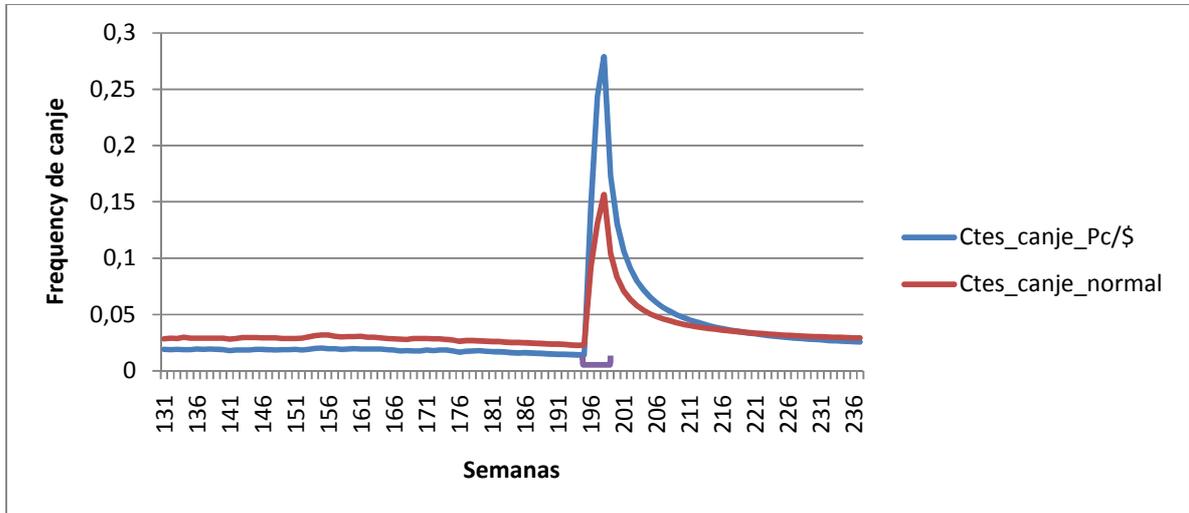
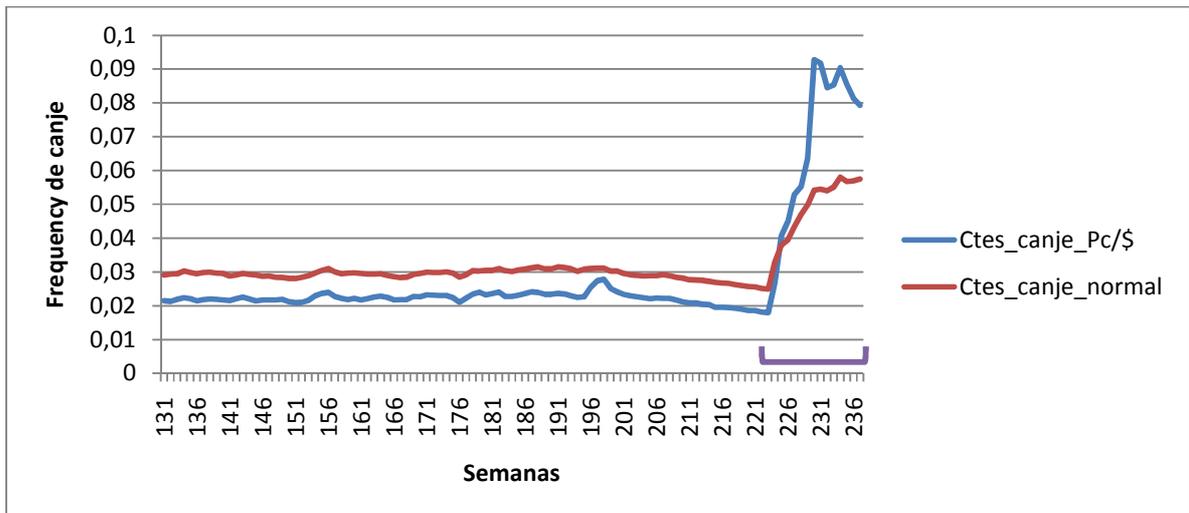


Gráfico 33: Frecuencia de canje promedio de clientes canjeadores del período 2



A partir de los gráficos es posible observar que la frecuencia de canje para los canjeadores que no usaron la promoción es mayor que aquellos que sí la usaron. Sin embargo, dado que gran parte de los clientes no había realizado canje durante los últimos años, su frecuencia de canje aumenta sobrepasando a la de los clientes que no usaron la promoción.

8.7.8 Monto de canje

A continuación se presenta el monto promedio semanal de los clientes que canjearon en los períodos promocionales:

Gráfico 34: Monto de canje promedio de clientes canjeadores del período 1

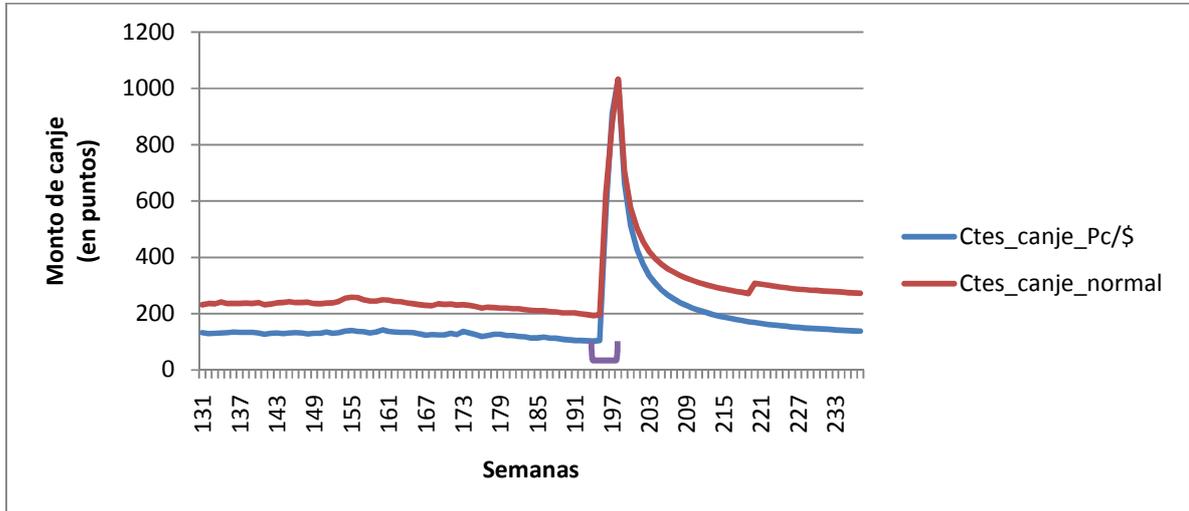
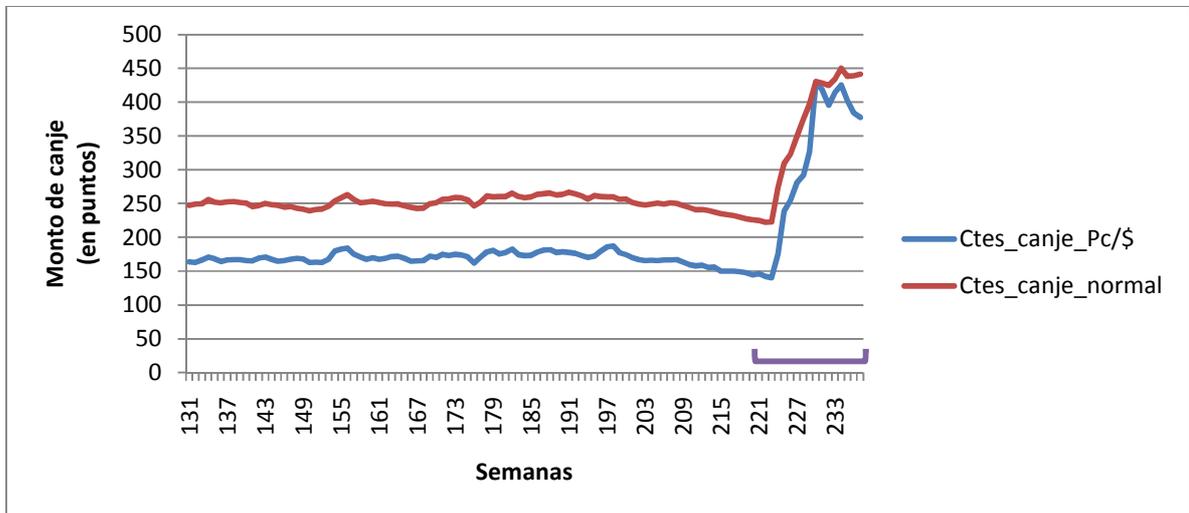


Gráfico 35: Monto de canje promedio de clientes canjeadores del período 1



Se observa para ambos períodos que aquellos clientes que no usan la promoción acumulan una mayor cantidad de puntos canjeados promedio semanalmente que aquellos que sí la usaron, lo que puede deberse a que han realizado un mayor número de canjes.

Luego de este análisis, es posible decir que aquellos clientes que canjearon utilizando la promoción son:

- En su mayoría mujeres.
- Compran en promedio dos de cada cinco semanas y sus montos promedio de semanal oscila entre los \$12.000 y \$18.000 (100 – 150 puntos entregados).
- Su frecuencia de canje es baja, en su gran mayoría canjean menos de una vez al año.
- Canjean principalmente en los niveles más bajos de puntos.

9. Conclusiones

9.1 Conclusiones sobre la estimación de demanda

Mediante modelos de series de tiempo, que combinan información histórica y variables externas, fue posible estimar el número de canjes agregado y por niveles, información que sirvió de línea base para cuantificar, en unidades de productos canjeados, el efecto de la promoción.

A través de un análisis descriptivo de los canjes fue posible observar un comportamiento estacional en la serie y determinar aquellas variables externas que influyen en los canjes, como las semanas correspondientes a fin de mes, navidad, semana santa y fiestas patrias. También se identificó que los canjes se ven directamente influenciados por los puntos entregados en períodos anteriores, los que están directamente relacionados con la compra de los clientes.

Los resultados de este modelo arrojan un aumento en el número de canjes totales durante el primer período promocional, incrementándose en más de 7.300 unidades. La realización de una segunda campaña promocional y la incorporación de esta modalidad de canje que combina puntos y pesos de manera permanente al programa, permitió contar con la información necesaria para determinar el efecto de la promoción en el largo plazo, observándose que al no tener ésta una caducidad, la urgencia por usarla disminuyó, lo que llevó a que en el segundo período el número de canjes se ajustara al de una semana normal.

El efecto de la promoción para cada nivel fue distinto al agregado. El primer período incrementó en algunos niveles el número de canje (1, 4, 5, 6, 7 y 8), los mantuvo de acuerdo a lo esperado en otros (nivel 3) y también los sobreestimó (nivel 2). En el segundo período, se observó que para los niveles que concentran un mayor porcentaje de canjes semanalmente (1, 2, 3 y 4) éstos se mantuviesen de acuerdo a lo esperado para una semana normal y para los niveles 5, 6, 7 y 8 se observase un

pequeño aumento en el canje. A partir de lo cual es posible concluir que, en el largo plazo, los canjes se mantienen de acuerdo a lo esperado para una semana normal.

9.2 Conclusiones sobre el modelo proporcional de Hazard

El modelo proporcional de Hazard es el modelo de supervivencia a través del cual se estimó la probabilidad de comprar en una determinada semana, dado el tiempo transcurrido desde la última ocurrencia del evento.

Para ésto, se modeló el comportamiento entre compras de los clientes a través de una línea base, la cual podía ser generada a partir de los datos (no paramétrica) o ser impuesta, asumiendo que los clientes siguen un patrón entre compras que puede ser determinado a través de alguna función de distribución conocida (exponencial, weibull, log-logística, entre otras).

Mediante la incorporación de las covariables al modelo, fue posible identificar a aquellas que influyen en la decisión de compra de los cliente. De esta manera, se identificó que mientras mayor es la frecuencia de compra, mayor es la probabilidad de comprar, y que ciertas semanas del año como fin de mes y semana santa tienen un efecto negativo en la compra, mientras que estar en semana de navidad incrementa la probabilidad de comprar.

En particular, para el caso en estudio, se deseaba testear dos hipótesis: 1) cómo afecta el comportamiento de compra de los clientes el encontrarse cerca o lejos de una recompensa y 2) determinar si los clientes aceleraron o no la compra para poder acceder a la promoción.

Con respecto a la segunda hipótesis, se encontró que ésta afectaba negativamente la probabilidad de compra de los clientes para los niveles más bajos de canje, lo que se explica pues más del 55% de los clientes que canjeó usando la promoción no había realizado canjes durante el último año, probablemente porque no haber alcanzado a acumular el mínimo de puntos para canjear, por lo que pueden haber decidido sacrificar la compra del período, privilegiando el canje.

Para determinar el efecto de la primera hipótesis, se incorporó al modelo una variable que cuantifica la cercanía o lejanía al nivel de canje, asociando el número de canjes acumulados durante el último año y la distancia relativa al nivel, encontrándose un parámetro negativo asociado a ésta variable, el cual indica que a medida que el cliente se encuentra más cerca de llegar a la recompensa, mayor es la probabilidad que realice una compra, lo cual es un claro indicio que la estructura del programa de fidelización de la empresa genera aceleración de compra en sus clientes.

De los distintas especificaciones de línea base usada para predecir la compra, la que presentó un mejor ajuste a los datos en todos los niveles fue la log-logística. Sin embargo, al analizar la capacidad predictiva de los modelos se encontró que el modelo no paramétrico entregaba mejores pronósticos.

El modelo proporcional de Hazard fue capaz de predecir correctamente el 60% de las compras y no compras, lo que debido a lo complejo que resulta predecir compra semanalmente usando sólo información transaccional, fue considerado como satisfactorio.

9.3 Conclusiones sobre el modelo de predicción de canje

Mediante la aplicación de un modelo logit, se estimó la probabilidad de canje con el objeto de determinar si los clientes que canjearon usando la promoción lo hubiesen hecho sin ella durante la misma semana.

A partir de este modelo, se observó que estar en semanas como fin de mes, navidad y promoción incrementa la probabilidad de canjear, mientras que estar en semana santa o fiestas patrias, la disminuye. Otras variables que influyen directamente en la decisión de canje es la distancia al nivel, donde a medida que el cliente se encuentre más próximo al nivel de canje, la probabilidad de canjear aumenta.

Las variables asociadas a la promoción, resultaron tener un efecto positivo, incrementando la posibilidad de canjear. Este efecto se vio reflejado en que de no haber estado la promoción vigente durante el mismo período, más del 50% de los clientes que usaron la promoción no hubiesen canjeado.

9.4 Conclusiones generales

Tras aplicar varias metodologías para evaluar el impacto de una promoción de canje de productos, es posible concluir que la promoción produjo efectos a nivel de canje y de compra.

Si bien la promoción no logró generar aceleración de compra en los clientes, logró incrementar el número de canjes, aumento que se vio reflejado posteriormente en una disminución del recency de compra de aquellos clientes que usaron la promoción, quienes previo ésta presentaban un recency mayor a los clientes que canjearon regularmente en el mismo período, llegando a ser iguales durante el período promocional y manteniéndose con un recency de compra menor semanas posteriores a la promoción.

Se observó también que, en promedio, los clientes que usaron la promoción compran gastan montos menores y tienen períodos más largos entre una compra y otra, razón por la cual acumulan una menor cantidad de puntos y por lo tanto realizar una menor cantidad de canjes.

Luego, si se considera que con el monto en pesos que el cliente paga al canjear en Puntos con pesos la empresa es capaz de cubrir todo o parte del costo del producto,

y que se ha observado un efecto post-promoción de compra positivo en aquellos clientes que utilizaron la promoción, se sugiere a la empresa que continúe realizándola.

9.5 Limitaciones de los modelos

Una limitación importante a considerar en el desarrollo de los modelos es la no incorporación de heterogeneidad para el desarrollo de estos. No todos los clientes responden de la misma forma ante una determinada acción de marketing, por lo que estimar un solo modelo que incorpore distintos tipos de cliente puede inducir a errores, pues los efectos pueden superponerse, eliminando, por ejemplo, la posibilidad de encontrar segmentos de clientes para los cuales la promoción sí generó aceleración de compra con el cual se pueda desarrollar un perfil más detallado del cliente que utilizó la promoción.

9.6 Trabajos futuros

Dado que el objetivo de los programas de fidelización es incentivar relaciones de largo plazo entre la empresa y el cliente, y aumentar la participación en las canastas de los éstos, se propone para trabajos futuros el diseño de acciones de marketing para los clientes más valiosos de la empresa y que puedan ser aplicadas en este tipo de programas. Un caso interesante para testear puede ser el incorporar recompensas más inmediatas que motiven al cliente a participar más frecuentemente del programa. Un ejemplo de esto puede ser la entrega de cupones de descuento para líneas de productos en las que el cliente compre o en empresas de alianzas, de manera que el cliente se vea recompensado y a su vez realice compras que beneficien a la empresa.

También se sugiere incorporar heterogeneidad al modelo proporcional de Hazard para estimar la probabilidad de compra, usando por ejemplo, clases latentes. De esta manera se podrá agrupar a los clientes de acuerdo a la sensibilidad de las variables a considerar en el modelo, por ejemplo la distancia psicológica a la recompensa, y pronosticar, en base a los parámetros propios de cada grupo, la compra.

10. Bibliografía

- [1] Anguita O., F. A.. 2008. Evaluación de promociones en tiendas de conveniencia. Memoria para optar al título de Ingeniero Civil Industrial, Santiago, Universidad de Chile, Facultad de Ciencias Físicas y Matemáticas, Departamento de Ingeniería Industrial.
- [2] Baek, H.. 2006. Estimación de customer lifetime value a nivel de clientes usando variables socio-demográficas y transaccionales. Memoria para optar al título de Ingeniero Civil Industrial, Santiago, Universidad de Chile, Facultad de Ciencias Físicas y Matemáticas, Departamento de Ingeniería Industrial.
- [3] Durán P., M. 2008. Estudio del comportamiento de compra y canje de los miembros de un club de lealtad. Memoria para optar al título de Ingeniero Civil Industrial, Santiago, Universidad de Chile, Facultad de Ciencias Físicas y Matemáticas, Departamento de Ingeniería Industrial.
- [4] Espinoza M., I. M. 2008. Estimación de demanda de premios, para un club de fidelización de acumulación de puntos. Memoria para optar al título de Ingeniero Civil Industrial, Santiago, Universidad de Chile, Facultad de Ciencias Físicas y Matemáticas, Departamento de Ingeniería Industrial.
- [5] Garrido R., M.. 2009. Estimación del valor de clientes de un club de fidelización de una tienda de retail. Memoria para optar al título de Ingeniero Civil Industrial, Santiago, Universidad de Chile, Facultad de Ciencias Físicas y Matemáticas, Departamento de Ingeniería Industrial.
- [6] Guadagni, P., Little, J.. 1983. A Logit Model Scanner Data. *Marketing Science*, 2(3): 203-238.
- [7] Kamakura, W. y Mela C.. 2004. "Choice Models and Relationship Management". Summary Paper for the sixth Choice Symposium.
- [8] Kivetz, R. y Simonson, I.. 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.
- [9] Kivetz, R., Urminsky, O. y Zheng, Y.. 2006. The Goal-Gradient Hypothesis Resurrected: Purchase Acceleration, Illusionary Goal Progress, and Customer Retention. *Journal of Marketing Research*, 43(1): 39-58.
- [10] Lewis, M.. 2004. The Influence of Loyalty Programs and Short-Term Promotions on Customer Retention. *Journal of Marketing Research*. 41(3): 281-292.
- [11] Liu, Y.. 2007. The long-term impact of loyalty programs on consumer purchase behavior and loyalty. *Journal of Marketing*, 71 (4): 19-35.

- [12] Mela, C., Gupta, S. y Lehmann, D.. 1997. The Long-term Impact of Promotion and Advertising on Consumer Brand Choice. *Journal of Marketing Research*, 34(1), 248-261.
- [13] Rodríguez, C.. 2001. Modelos Box-Jenkins Aplicación de su metodología a la producción de azúcar en Cuba. *Economía y Desarrollo*, 128(1): 167-180.
- [14] Seetharaman, P., 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.

11. Anexos

11.1 Anexo 1: Cálculo del tamaño muestral

Para determinar el tamaño muestral se utiliza la siguiente fórmula:

$$n = \frac{N * Z_{\alpha/2}^2 * P * (1-P)}{(N-1) * e^2 + Z_{\alpha/2}^2 * P * (1-P)} \quad (39)$$

Donde:

- N: tamaño de la población (Se asume que se distribuye normalmente).
- $Z_{\alpha/2}$: nivel de confianza.
- P: proporción de una categoría. Se elige $P = 0,5$ para maximizar el tamaño muestral.
- e = error máximo.

Se escoge un nivel de confianza del 95%, un error máximo del 2% y una proporción de categoría del 0,5, que maximiza el tamaño muestral, el que puede verse en la siguiente tabla:

Tabla 34: Tamaño muestral por niveles

	Nivel 1	Nivel 2	Nivel 3	Nivel 4	Nivel 5	Nivel 6	Nivel 7	Nivel 8
N	675.040	199.062	145.171	24.179	2.083	1.988	431	324
Z_{95%}	1,96	1,96	1,96	1,96	1,96	1,96	1,96	1,96
P	0,50	0,50	0,50	0,50	0,50	0,50	0,50	0,50
E	0,02	0,02	0,02	0,02	0,02	0,02	0,02	0,02
n	2.392,49	2.372,40	2.361,95	2.184,20	1.115,61	1.087,78	365,54	285,58

El tamaño muestral escogido para cada nivel fue el siguiente

Tabla 35: Tamaño de las muestras utilizadas por niveles

Nivel 1	6.000,00
Nivel 2	6.000,00
Nivel 3	5.500,00
Nivel 4	5.500,00
Nivel 5	1.300,00
Nivel 6	1.300,00
Nivel 7	370,00
Nivel 8	290,00

11.2 Anexo 2: Número de canjes semanales por mes

Gráfico 36: Número de canjes semanales, mes de febrero

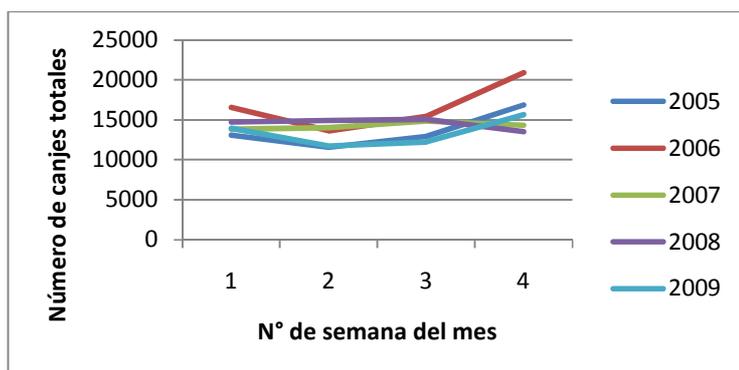


Gráfico 37: Número de canjes semanales, mes de abril

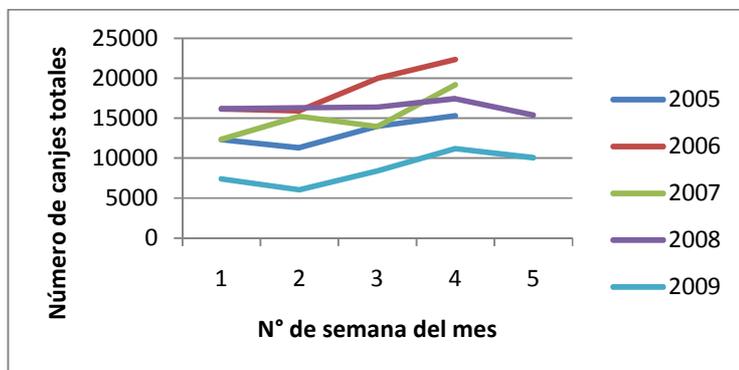


Gráfico 38: Número de canjes semanales, mes de mayo

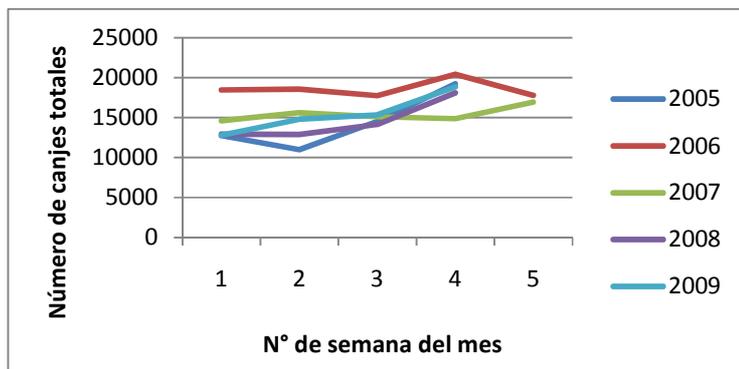


Gráfico 39: Número de canjes semanales, mes de junio

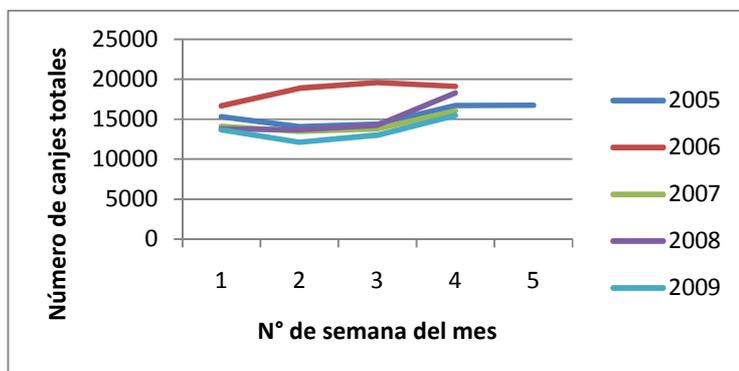


Gráfico 40: Número de canjes semanales, mes de julio

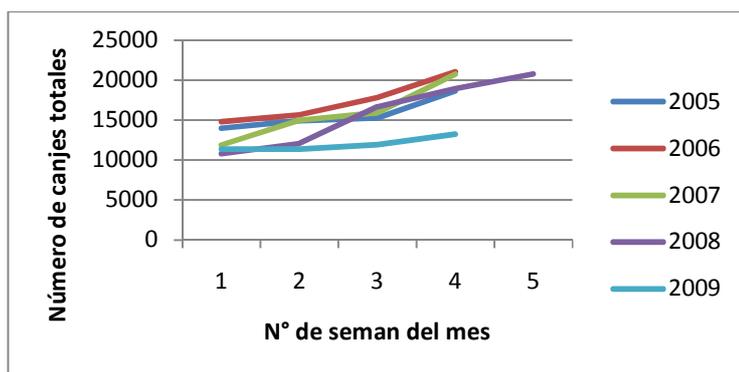


Gráfico 41: Número de canjes semanales, mes de agosto

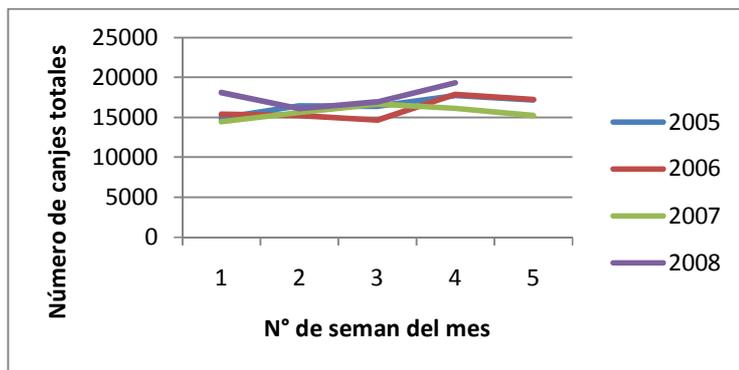


Gráfico 42: Número de canjes semanales, mes de octubre

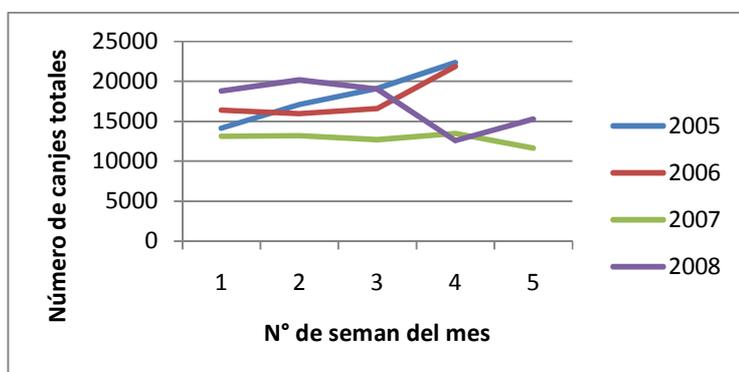
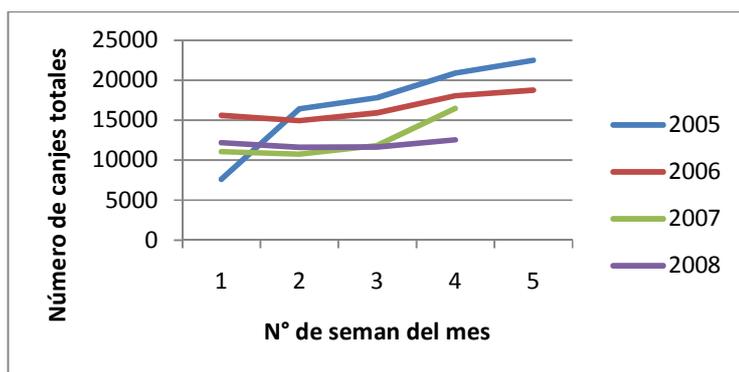


Gráfico 43: Número de canjes semanales, mes de noviembre



11.3 Anexo 3: Estimación de canje por niveles

Tabla 36: Resumen del efecto de la primera Promoción por nivel

	Pre-promoción	Promoción	Post-promoción
Global	Ajusta	Subestima	Ajusta
Nivel 1	Ajusta	Subestima	Ajusta
Nivel 2	Sobreestima	Sobreestima	Sobreestima
Nivel 3	Ajusta	Ajusta	Sobreestima
Nivel 4	Ajusta	Subestima	Sobreestima
Nivel 5	Ajusta	Subestima	Subestima
Nivel 6	Ajusta	Subestima	Subestima
Nivel 7	Ajusta	Subestima	Sobreestima
Nivel 8	Subestima	Subestima	Ajusta

Tabla 37: Parámetros por nivel del Modelo 1

Nivel	1	2	3	4	5	6	7	8
AR1	0,73	0,90	0,98	0,91	0,96	0,95	1,00	0,99
MA1		0,27	0,50	0,47	0,64	0,65	0,87	0,82
MA2		0,23	0,16					
Seasonal AR1	0,42	0,50	0,55	0,36		0,28		
PE_1	8,36E-06	1,35E-06	1,88E-06	5,44E-07	7,45E-08	1,01E-07		
PE_2				4,65E-07	5,49E08			
Patrias	- 3.060,28	- 4.585,66	- 811,04					
Navidad	2.413,49	3.474,26	511,55					
S. santa	- 1.542,30	- 2.413,11	- 450,38					
fdm_1	1.304,37	1.923,17				9,14		
fdm_2	1.390,27	1.822,55			10,95			
fdm_3	2.037,29	3.095,72	517,29	126,28	18,84	18,61	4,16	5,84
fdm_4	1.448,82	1.987,77			14,36	13,25	2,78	5,76
fdm_5	1.464,91	1.861,01		114,51	14,81	11,93	3,00	4,14
fdm_6	2.250,17	3.302,11			15,62	15,19		
fdm_7	2.283,27	3.788,05	591,25	79,14	20,84	14,76	3,08	
fdm_8						14,69	3,27	
fdm_9					13,31			
fdm_10	2.122,87	3.142,86	488,49	82,61	10,05	15,28		
fdm_11								
fdm_12	1.560,08	3.230,79				11,43	5,27	
constante	8.297,62	12.826,43	1.800,31	151,75				

Tabla 38: Parámetros por nivel del Modelo 2

Nivel	1	2	3	4	5	6	7	8
AR1	0,69	0,96	0,97	0,97	0,98	0,96	1,00	0,98
MA1		0,36	0,43	0,50	0,67	0,65	0,75	0,76
MA2		0,30	0,18	0,21			0,15	
Seasonal AR1	0,38	0,41	0,51	0,30	0,18	0,34	0,30	
PE_1	5,60E-06	1,91E-06	1,89E-06	5,30E-07	4,75E-08	7,83E-08		
PE_2			1,53E-06	5,05E-07	5,21E-08	3,86E-08		
Patrias	- 3.092,30	- 810,10	- 839,55	- 109,35				
Navidad	3.624,04	621,52	641,30	96,19	12,74			- 4,33
S. santa	- 1.375,41		- 366,51					
Pre-promo								
Promo	3.740,16				18,83			
Post-promo			- 547,55	- 135,50				
fdm_1	1.631,61	432,91	327,07		11,90	11,38		
fdm_2	1.717,82				12,24			
fdm_3	2.064,70		515,43	118,55	19,65	17,77	3,95	5,77
fdm_4	1.454,35				12,16	12,77		5,68
fdm_5	1.485,64			97,80	15,40	11,59		4,13
fdm_6	2.171,74	586,90			15,76	14,06		
fdm_7	2.313,12	703,18	579,62	78,14	21,62	14,26		
fdm_8						14,73		
fdm_9	2.107,51				18,35	13,85		3,20
fdm_10	2.346,35	679,85	510,27	81,26		12,89		
fdm_11					10,36	13,87		
fdm_12	1.218,64	505,86						
constante	8.840,42	2.181,44	1.510,62	163,79				

Tabla 39: Ajuste por nivel de los Modelos 1 y 2

Nivel		Modelo 1		Modelo 2	
		MAPE	NMSE	MAPE	NMSE
1	Calibración	11,19%	32,53%	11,13%	31,15%
	Pronóstico	19,18%	54,31%	64,85%	869,87%
2	Calibración	12,33%	45,07%	12,88%	35,57%
	Pronóstico	51,07%	175,72%	26,98%	83,06%
3	Calibración	11,40%	30,47%	11,43%	28,53%
	Pronóstico	35,76%	84,34%	35,16%	92,88%
4	Calibración	16,66%	39,99%	15,99%	38,74%
	Pronóstico	23,56%	83,45%	31,49%	102,49%
5	Calibración	35,78%	37,41%	33,26%	31,44%
	Pronóstico	42,38%	131,54%	35,50%	111,12%
6	Calibración	26,86%	48,80%	26,92%	40,46%
	Pronóstico	33,55%	131,41%	36,00%	114,05%
7	Calibración	53,61%	81,71%	52,97%	70,97%
	Pronóstico	46,94%	130,76%	37,71%	126,96%
8	Calibración	58,83%	82,83%	60,15%	82,58%
	Pronóstico	59,31%	123,49%	51,27%	176,13%

11.4 Anexo 4: Modelo proporcional de Hazard

11.4.1 Nivel 2

Tabla 40: Parámetros del modelo proporcional de Hazard para el nivel 2

	NP	EXP	W	LL
Frecuencia de compra	3,202	3,478	4,807	5,981
Promoción	- 0,145	- 0,143	- 0,226	- 0,190
Post-promoción	- 0,230	- 0,268	- 0,385	- 0,333
Fin de mes	- 0,079	- 0,095	- 0,079	- 0,073
Navidad	0,169	0,142	0,181	0,441
Semana santa	- 0,088	- 0,089	- 0,072	- 0,124
Dist – Canje	- 0,432	- 0,588	- 0,799	- 0,975
ALPHA			1,460	2,246
GAMMA		0,045	0,054	0,107

Tabla 41: Ajuste del modelo proporcional de Hazard para el nivel 2

	NP	EXP	W	LL
N° de observaciones	377.202	377.202	377.202	377.202
LL	- 2.153.562,70	- 341.838,99	- 320.444,83	- 301.741,17
N ° de parámetros		8,0	9,0	9,0
SBS	4.307.210,1	683.780,7	641.005,2	603.597,9

Tabla 42: Indicadores matriz de confusión (base de calibración)

	NP	Exp	W	LL
δ	23%	18%	15%	21%
PREDIC_OK	64,1%	62,9%	62,0%	59,7%
C_OK	63,0%	60,1%	61,2%	57,3%
NC_OK	65,0%	65,5%	62,8%	61,8%
C_ERROR	35,0%	34,5%	37,2%	38,2%
NC_ERROR	37,0%	39,9%	38,8%	42,7%

Tabla 43: Indicadores matriz de confusión (base de pronóstico)

	NP	Exp	W	LL
δ	23%	18%	15%	21%
PREDIC_OK	63,5%	62,4%	60,9%	58,5%
C_OK	61,1%	58,2%	58,8%	53,8%
NC_OK	65,6%	66,2%	62,9%	62,7%
C_ERROR	34,4%	33,8%	37,1%	37,3%
NC_ERROR	38,9%	41,8%	41,2%	46,2%

11.4.2 Nivel 3

Tabla 44: Parámetros del modelo proporcional de Hazard para el nivel 3

	NP	EXP	W	LL
Frecuencia de compra	3,055	3,478	4,831	5,983
Pre-promoción			- 0,108	
Promoción	- 0,071	- 0,087	- 0,146	- 0,094
Post-promoción	- 0,180	- 0,197	- 0,247	- 0,242
Fin de mes	- 0,079	- 0,091	- 0,074	- 0,073
Navidad	0,128	0,104	0,146	0,373
Semana santa	- 0,086	- 0,082	- 0,065	- 0,130
Dist – Canje	- 0,481	- 0,727	- 1,029	- 1,209
ALPHA			1,461	2,349
GAMMA		0,048	0,056	0,121

Tabla 45: Ajuste del modelo proporcional de Hazard para el nivel 3

	NP	EXP	W	LL
N° de observaciones	404.393,0	404.393,0	404.393,0	404.393,0
LL	- 2.566.867,4	- 382.619,1	- 357.312,6	- 332.544,4
N ° de parámetros	7,0	8,0	10,0	9,0
SBS	5.133.820,500	765.341,547	714.754,322	665.204,932

Tabla 46: Indicadores matriz de confusión (base de calibración)

	NP	Exp	W	LL
δ	26%	20%	16%	21%
PREDIC_OK	65,3%	64,5%	63,5%	60,6%
C_OK	67,8%	63,6%	64,9%	62,5%
NC_OK	62,6%	65,4%	62,1%	58,6%
C_ERROR	37,4%	34,6%	37,9%	41,4%
NC_ERROR	32,2%	36,4%	35,1%	37,5%

Tabla 47: Indicadores matriz de confusión (base de pronóstico)

	NP	Exp	W	LL
δ	26%	20%	16%	21%
PREDIC_OK	64,9%	64,1%	62,7%	59,9%
C_OK	64,9%	60,6%	61,3%	59,4%
NC_OK	65,0%	67,8%	64,2%	60,4%
C_ERROR	35,0%	32,2%	35,8%	39,6%
NC_ERROR	35,1%	39,4%	38,7%	40,6%

11.4.3 Nivel 4

Tabla 48: Parámetros del modelo proporcional de Hazard para el nivel 4

	NP	EXP	W	LL
Frecuencia de compra	2,607	3,200	4,956	5,831
Pre-promoción			- 0,105	
Promoción			- 0,220	0,086
Post-promoción	- 0,158	- 0,227	- 0,525	- 0,100
Fin de mes	- 0,049	- 0,052	- 0,038	- 0,039
Navidad	0,102	0,088	0,126	0,328
Semana santa	- 0,056	- 0,052	- 0,044	- 0,090
Dist – Canje	- 0,481	- 0,768	- 1,205	- 1,451
ALPHA			1,604	2,789
GAMMA		0,057	0,063	0,165

Tabla 49: Ajuste del modelo proporcional de Hazard para el nivel 4

	NP	EXP	W	LL
N° de observaciones	280.079	280.079	280.079	280.079
LL	-2109856,122	- 282.662,234	- 252.324,289	- 227.901,269
N ° de parámetros		7,0	10,0	9,0
SBS	4.219.784,7	565.412,3	504.774,0	455.915,4

Tabla 50: Indicadores matriz de confusión (base de calibración)

	NP	Exp	W	LL
δ	38%	26%	20%	27%
PREDIC_OK	68,8%	68,2%	66,5%	62,7%
C_OK	70,2%	68,9%	67,7%	65,1%
NC_OK	66,4%	67,0%	64,4%	58,7%
C_ERROR	33,6%	33,0%	35,6%	41,3%
NC_ERROR	29,8%	31,1%	32,3%	34,9%

Tabla 51: Indicadores matriz de confusión (base de pronóstico)

	NP	Exp	W	LL
δ	38%	26%	20%	27%
PREDIC_OK	67,4%	66,3%	63,6%	60,1%
C_OK	65,4%	63,4%	61,0%	58,6%
NC_OK	70,7%	71,0%	67,8%	62,6%
C_ERROR	29,3%	29,0%	32,2%	37,4%
NC_ERROR	34,6%	36,6%	39,0%	41,4%

11.4.4 Nivel 5

Tabla 52: Parámetros del modelo proporcional de Hazard para el nivel 5

	NP	EXP	W	LL
Frecuencia de compra	2,230	3,129	5,053	5,486
Promoción				0,211
Post-promoción			- 0,098	
Fin de mes	- 0,027	- 0,027		
Navidad				0,183
Dist – Canje	- 0,432	- 0,809	- 1,310	- 1,470
ALPHA			1,623	3,064
GAMMA		0,061	0,061	0,210

Tabla 53: Ajuste del modelo proporcional de Hazard para el nivel 5

	NP	EXP	W	LL
N° de observaciones	57.812	57.812	57.812	57.812
LL	- 430.464,99	- 61.329,60	- 54.252,68	- 47.104,07
N ° de parámetros		4,0	5,0	6,0
SBS	860.961,8	122.659,2	108.505,4	94.208,1

Tabla 54: Indicadores matriz de confusión (base de calibración)

	NP	Exp	W	LL
δ	45%	27%	21%	31%
PREDIC_OK	74,6%	73,3%	70,3%	65,7%
C_OK	77,0%	74,7%	70,8%	67,8%
NC_OK	68,7%	69,7%	69,2%	60,6%
C_ERROR	31,3%	30,3%	30,8%	39,4%
NC_ERROR	23,0%	25,3%	29,2%	32,2%

Tabla 55: Indicadores matriz de confusión (base de pronóstico)

	NP	Exp	W	LL
δ	45%	27%	21%	31%
PREDIC_OK	70,4%	68,7%	65,7%	61,2%
C_OK	64,4%	61,3%	57,0%	55,4%
NC_OK	79,4%	79,9%	79,1%	70,1%
C_ERROR	20,6%	20,1%	20,9%	29,9%
NC_ERROR	35,6%	38,7%	43,0%	44,6%

11.4.5 Nivel 6

Tabla 56: Parámetros del modelo proporcional de Hazard para el nivel 6

	NP	EXP	W	LL
Frecuencia de compra	2,013	2,963	4,853	5,311
Fin de mes	- 0,044	- 0,045	- 0,041	- 0,039
Dist – Canje	- 0,455	- 0,815	- 1,232	- 1,558
ALPHA			1,618	3,305
GAMMA		0,069	0,065	0,245

Tabla 57: Ajuste del modelo proporcional de Hazard para el nivel 6

	NP	EXP	W	LL
N° de observaciones	63.179	63.179	63.179	63.179
LL	- 498.902,10	- 68.026,27	- 59.637,37	- 49.545,70
N° de parámetros		4,0	5,0	5,0
SBS	860.961,8	136.096,8	119.330,0	99.146,7

Tabla 58: Indicadores matriz de confusión (base de calibración)

	NP	Exp	W	LL
δ	52%	31%	20%	30%
PREDIC_OK	74,3%	72,5%	72,0%	67,2%
C_OK	75,6%	73,3%	74,5%	71,1%
NC_OK	70,4%	70,3%	64,6%	55,9%
C_ERROR	29,6%	29,7%	35,4%	44,1%
NC_ERROR	24,4%	26,7%	25,5%	28,9%

Tabla 59: Indicadores matriz de confusión (base de pronóstico)

	NP	Exp	W	LL
δ	52%	31%	20%	30%
PREDIC_OK	71,1%	68,7%	67,4%	62,1%
C_OK	64,0%	60,2%	61,4%	58,2%
NC_OK	83,0%	83,1%	77,7%	68,6%
C_ERROR	17,0%	16,9%	22,3%	31,4%
NC_ERROR	36,0%	39,8%	38,6%	41,8%

11.4.6 Nivel 7

Tabla 60: Parámetros del modelo proporcional de Hazard para el nivel 7

	NP	EXP	W	LL
Frecuencia de compra	2,003	3,099	5,579	5,776
Dist – Canje	- 0,544	- 0,934	- 1,506	- 1,794
ALPHA			1,793	3,139
GAMMA		0,059	0,057	0,196

Tabla 61: Ajuste del modelo proporcional de Hazard para el nivel 7

	NP	EXP	W	LL
N° de observaciones	12.980	12.980	12.980	12.980
LL	- 88.331,58	- 13.870,42	- 11.687,37	- 10.695,11
N ° de parámetros		3,0	4,0	4,0
SBS	175.837,3	27.769,3	23.412,6	21.428,1

Tabla 62: Indicadores matriz de confusión (base de calibración)

	NP	Exp	W	LL
δ	52%	31%	20%	30%
PREDIC_OK	77,5%	76,1%	74,4%	70,8%
C_OK	79,3%	77,3%	75,6%	72,6%
NC_OK	71,9%	72,4%	70,7%	65,1%
C_ERROR	28,1%	27,6%	29,3%	34,9%
NC_ERROR	20,7%	22,7%	24,4%	27,4%

Tabla 63: Indicadores matriz de confusión (base de pronóstico)

	NP	Exp	W	LL
δ	52%	31%	20%	30%
PREDIC_OK	72,8%	69,3%	68,3%	65,8%
C_OK	66,6%	59,4%	59,3%	58,9%
NC_OK	84,0%	86,9%	84,3%	78,1%
C_ERROR	16,0%	13,1%	15,7%	21,9%
NC_ERROR	33,4%	40,6%	40,7%	41,1%

11.4.7 Nivel 8

Tabla 64: Parámetros del modelo proporcional de Hazard para el nivel 8

	NP	EXP	W	LL
Frecuencia de compra	1,805	2,541	5,122	5,349
Dist – Canje	- 0,261	- 0,457	- 1,004	- 0,870
ALPHA			1,875	4,815
GAMMA		0,090	0,076	0,358

Tabla 65: Ajuste del modelo proporcional de Hazard para el nivel 8

	NP	EXP	W	LL
N° de observaciones	12.662	12.662	12.662	12.662
LL	- 93.948,18	- 13.534,53	- 10.570,04	- 6.647,36
N ° de parámetros		3,0	4,0	4,0
SBS	175.837,3	27.097,4	21.177,9	13.332,5

Tabla 66: Indicadores Indicadores matriz de confusión (base de calibración)

	NP	Exp	W	LL
δ	61%	40%	33%	39%
PREDIC_OK	77,6%	73,6%	66,8%	64,4%
C_OK	79,9%	74,5%	67,1%	69,2%
NC_OK	67,5%	69,6%	65,5%	43,7%
C_ERROR	32,5%	30,4%	34,5%	56,3%
NC_ERROR	20,1%	25,5%	32,9%	30,8%

Tabla 67: Indicadores matriz de confusión (base de pronóstico)

	NP	Exp	W	LL
δ	61%	40%	33%	39%
PREDIC_OK	75,9%	72,4%	68,1%	65,4%
C_OK	75,3%	69,5%	64,5%	69,4%
NC_OK	77,7%	81,7%	79,7%	52,8%
C_ERROR	22,3%	18,3%	20,3%	47,2%
NC_ERROR	24,7%	30,5%	35,5%	30,6%

11.5 Anexo 5: Modelo de predicción de canjes

Tabla 68: Parámetros del modelo de predicción de canje por nivel

Nivel	2	3	4	5	6	7	8
R_canje	0,003	0,004	0,003	0,003	0,003	0,012	
Pre-promoción	1,371	0,816	0,700	0,700	0,700		
Promoción	2,088	1,379	2,121	2,121	2,121		
Post-promoción	0,853	0,166	0,154	0,154	0,154		
Fin de mes	0,139	- 0,521	- 0,494	- 0,494	- 0,494		
Fiestas patrias	- 0,425	0,710	0,624	0,624	0,624		
Navidad	0,545	- 0,326	- 0,417	- 0,417	- 0,417	2,041	
Dist-canje	- 13,156	- 14,000	- 16,027	- 16,027	- 16,027	- 16,993	- 18,055
Constante	0,807	0,798	0,885	0,885	0,885	1,177	0,516

Tabla 69: Ajuste del modelo de predicción de canje por nivel

Nivel	2	3	4	5	6	7	8
Observaciones	15.644,0	17.699,0	14.722,0	3.262,0	3.498,0	754,0	836,0
Log-likelihood	- 7.003,51	- 7.790,21	- 6.374,15	- 6.374,15	- 1.434,23	- 303,473	- 360,963
Parámetros	9,0	9,0	9,0	9,0	9,0	3,0	4,0
SBS	14.094,0	15.668,5	12.834,7	12.821,1	2.941,9	626,8	748,8
L(0)	- 10.843,6	- 12.268,0	- 10.204,5	- 10.204,5	- 2.424,6	- 522,6	- 579,5
R ² Cox & Snell	38,8%	39,7%	40,6%	90,4%	43,2%	44,1%	40,7%
R ² Nagelkerke	51,7%	52,9%	54,1%	90,6%	57,6%	58,8%	54,3%

Tabla 70: Indicadores matriz de confusión (base de calibración)

Nivel	2	3	4	5	6	7	8
δ	67%	65%	63%	68%	68%	65%	65%
PREDIC_OK	81,3%	81,5%	82,2%	83,0%	82,6%	82,8%	81,3%
C_OK	90,5%	91,5%	92,2%	89,7%	89,9%	92,3%	90,9%
NC_OK	72,0%	71,4%	72,2%	76,3%	75,2%	73,2%	71,8%
C_ERROR	28,0%	28,6%	27,8%	23,7%	24,8%	26,8%	28,2%
NC_ERROR	9,5%	8,5%	7,8%	10,3%	10,1%	7,7%	9,1%

Tabla 71: Indicadores matriz de confusión (base de pronóstico)

Nivel	2	3	4	5	6	7	8
δ	67%	65%	63%	68%	68%	65%	65/
PREDIC_OK	70,8%	71,1%	70,5%	76,0%	75,2%	69,5%	68,8%
C_OK	90,5%	91,5%	92,5%	90,5%	90,3%	93,4%	92,2%
NC_OK	70,4%	70,7%	70,0%	75,7%	74,8%	68,9%	68,0%
C_ERROR	29,6%	29,3%	30,0%	24,3%	25,2%	31,1%	32,0%
NC_ERROR	9,5%	8,5%	7,5%	9,5%	9,7%	6,6%	7,8%

11.6 Anexo 6: Porcentaje de canjes y no canjes predichos para clientes que usaron la promoción en período 1

Tabla 72: Porcentaje de canjes y no canjes predichos para clientes que usaron la promoción en período 2.

Nivel	1	2	3	4	5	6	7	8
No	12%	38%	26%	52%	45%	75%	65%	50%
Sí	88%	62%	74%	48%	55%	25%	35%	50%

Tabla 73: Porcentaje de canjes y no canjes predichos para clientes que canjearon usando la promoción en período 2 sin incorporar variables promocionales.

Nivel	1	2	3	4	5	6	7	8
No	54%	66%	54%	77%	91%	82%	65%	50%
Sí	46%	34%	46%	23%	9%	18%	35%	50%