



UNIVERSIDAD DE CHILE

FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS

DEPARTAMENTO DE INGENIERÍA INDUSTRIAL

ESTIMACIÓN DE CANIBALIZACIÓN DE CANALES EN RETAIL

MEMORIA PARA OPTAR AL TÍTULO DE INGENIERA CIVIL INDUSTRIAL

CARLA VALENTINA SÁEZ ROJAS

PROFESOR GUÍA:
MARCEL GOIC FIGUEROA

MIEMBROS DE LA COMISIÓN:
ALEJANDRA PUENTE CHANDÍA
RICARDO MONTOYA MOREIRA

SANTIAGO DE CHILE

2020

ESTIMACIÓN DE CANIBALIZACIÓN DE CANALES EN RETAIL

En los últimos años, el retail ha vivido una transformación digital, donde miles de personas han comenzado a preferir comprar online, en consecuencia, se plantea investigar, desde un punto de vista estratégico, la interacción entre los canales de venta de una empresa de retail chilena que comenzó con tiendas físicas (canal offline) y que con los años adicionó su página web (canal online) el cual en los últimos 5 años ha duplicado sus ventas. Para ello, se busca determinar la existencia de canibalización o sinergia positiva entre los canales para los principales tipos de productos (vestuario, decoración y electrónica), para diferentes tipos de clientes y diferenciando por zonas geográficas del país. En palabras simples, se propone precisar que porción de las ventas totales del canal online corresponden a transacciones que anteriormente se hacían por el canal offline. O visto desde otra arista: cuanta de la disminución de ventas en tiendas físicas son debido a que los clientes ahora compran por el sitio web.

Considerando datos transaccionales entre 2016 y 2019 se analizó la situación mediante Modelos Vectoriales Autorregresivos para Series de Tiempo (VAR), enmarcado en el contexto de la metodología Knowledge Discovery in Data Base (KDD). De esta forma se obtuvo que predomina una canibalización (96% de los casos) aunque han afectado en diferentes niveles a cada categoría de producto, tipo de cliente y zona geográfica de Chile. A nivel global, se ha dejado de vender un 2% en las tiendas físicas debido al aumento de ventas por el canal online, siendo la categoría más afectada decoración. Además, en los últimos 3 años el canal offline ha disminuido sus ventas en ~\$166MM por año y según las estimaciones solo el ~16% se debe a las compras por el sitio web, por lo que existen otras razones por las cuales las ventas han disminuido.

Se recomiendan tres ejes para trabajos futuros, en primer lugar, ahondar en la relación entre canales en el periodo 2017-2019 considerando, por ejemplo, categorías de productos más específicas. Por otra parte, replicar la metodología con datos del año 2020 incluyendo también variables que capturen el efecto del Covid-19. En cuando al tercer eje se propone un estudio para identificar qué factores influyen en la decisión del canal de compra, de forma de complementar el estudio de canibalización.

“Todos pueden superar sus circunstancias y alcanzar el éxito si tienen dedicación y pasión por lo que hacen”

Nelson Mandela

Agradecimientos

Me tocó desarrollar mi trabajo de título en medio de una pandemia, entre clases virtuales, teletrabajo y cuarentena, lo que dista bastante de lo que hubiera imaginado un año atrás cuando comencé este desafío. Ha sido un camino con más altos que bajos y sin estos últimos no estaría en este instante tan satisfecha por haber terminado este trabajo y con ello mi etapa universitaria donde me enamoré de mi carrera.

Infinitas gracias a mi familia por su constante apoyo y amor, por siempre haber creído en mi e inspirarme a luchar por todas mis metas y sueños desde que era pequeña, sin lugar a duda son un pilar fundamental en mi vida.

Me parece increíble lo rápido que han pasado los últimos años, mi paso por Beauchef está marcado por millones de recuerdos desde mis días de mechona hasta ahora. Agradezco la compañía de mis amigas y amigos en las infinitas horas de estudio, conversaciones, trabajo, juntas y viajes que me alegraron la vida y me hicieron amar la universidad hasta en los peores momentos. Gracias a mis profesores, sobre todo a quienes guiaron este trabajo que junto a la empresa hicieron que el proceso estuviera lleno de aprendizaje y resultados satisfactorios.

Hoy también agradezco todo aquello que me salió mal, porque así fue como me encontré con grandes oportunidades, maravillosas experiencias y por sobre todo valiosos aprendizajes.

Me faltarían páginas para agradecer a cada una de las personas que con su compañía, apoyo y consejos me han hecho una persona más completa y feliz. Simplemente gracias a todos aquellos que han estado siempre.

Carly.

Tabla de contenido

1	Introducción	1
2	Revisión de Literatura	3
3	Antecedentes y análisis preliminar de los datos	6
3.1	Evolución de las ventas	6
3.2	Clientes	7
3.3	Productos	11
3.4	Análisis preliminar de relación entre canales	13
4	Objetivos	16
4.1	Objetivo general	16
4.2	Objetivos específicos	16
5	Metodología: Knowledge Discovery in Data Base	17
5.1	Abstracción del escenario	17
5.2	Selección de datos, pre- procesamiento y transformación	18
5.3	Minería de datos, evaluación e interpretación	18
5.4	Conocimiento descubierto	24
6	Análisis de resultados	25
6.1	Ventas globales	25
6.2	Ventas clientes premium v/s no premium	28
6.3	Ventas por zonas geográficas	32
6.4	Sensibilidad de los resultados	37
7	Conclusiones	39
8	Trabajo futuro propuesto	42
9	Bibliografía	44
10	Anexos	46
10.1	Puntos críticos IRF	46
10.2	Causalidad de Granger	46
10.3	Elecciones de modelos y funciones de impulso respuesta	47
10.3.1	Ventas globales	47
10.3.2	Clientes premium	49

10.4	Clientes no Premium	53
10.5	Análisis por zona	57
10.5.1	Zona norte	57
10.5.2	Zona RM Oriente	61
10.5.3	Zona RM Poniente	65
10.5.4	Zona Sur	69
10.6	Comparaciones estimaciones de modelos	72
10.6.1	Clientes premium	73
10.6.2	Clientes no premium	74
10.6.3	Categoría vestuario por zonas	74
10.6.4	Categoría electrónica por zonas	75
10.6.5	Categoría decoración por zonas	75

Índice de tablas

Tabla 1: Hallazgos de los principales estudios sobre relación entre canales. Fuente: Huang, Lu & Ba, 2015	3
Tabla 2: Evolución ventas anuales. Fuente: Elaboración propia.	6
Tabla 3: Identificación de ventas y boletas como promedio de los últimos 5 años. Fuente: Elaboración propia.	7
Tabla 4: Diferencia porcentual en cantidad de clientes identificados. Fuente: Elaboración propia	8
Tabla 5: Descripción de zonas y cantidad de tiendas que posee cada una, promedio en periodo 2015-2019. Fuente: Elaboración total.	10
Tabla 6: Proporciones de transacciones por canal en cada zona, periodo 2015-2019. Fuente: Elaboración propia.	10
Tabla 7: Comparación clientes premium vs resto de clientes, en periodo 2015-2019. Fuente: Elaboración propia	11
Tabla 8: Proporción ventas por canal según zona y tipo de producto, promedio 2015-2019. Fuente: Elaboración propia.	12
Tabla 9: Aumento de porcentaje de ventas en canal online entre 2017 y 2019 por tipo de clientes. Fuente: elaboración propia.	14
Tabla 10: Aumento de porcentaje de ventas en canal online entre 2017 y 2019 por zona. Fuente: elaboración propia.	14
Tabla 11: Estimación de coeficientes para cada modelo global. Fuente: Elaboración propia.	25
Tabla 12: Evolución métricas de canibalización del análisis global. Fuente: Elaboración propia.	28
Tabla 13: Composición de cada modelo. Fuente: Elaboración propia.	28
Tabla 14: Comparación estimación de modelos generales según tipo de clientes. Fuente: Elaboración propia.	29
Tabla 15: Comparación de coeficientes de variable navidad y cyber. Fuente: Elaboración propia.	30
Tabla 16: Métricas de canibalización clientes premium vs no premium. Fuente: Elaboración propia.	32
Tabla 17: Caracterización modelos análisis por zona. Fuente: Elaboración propia.	32
Tabla 18: Coeficientes estimaciones por zona en ventas generales. Fuente: elaboración propia.	33
Tabla 19: Coeficientes variable navidad, modelos por zona y categoría de productos. Fuente: Elaboración propia.	34
Tabla 20: Métricas de canibalización del análisis por zona. Fuente: Elaboración propia.	36
Tabla 21: Diferencia de estimaciones al cambiar el horizonte de tiempo. Fuente: Elaboración propia.	38
Tabla 22: Identificación de casos donde hay causalidad en el sentido de Granger. Fuente: Elaboración propia.	46
Tabla 23: Comparación modelos ventas globales generales. Fuente: Elaboración propia.	47
Tabla 24: Modelos VAR y VARX de ventas globales: vestuario. Fuente: elaboración propia	48
Tabla 25: Modelos VAR y VARX para categoría Electrónica. Fuente: Elaboración propia	48
Tabla 26: Modelos VAR y VARX para la categoría Decoración, análisis global. Fuente: Elaboración propia.	49
Tabla 27: Modelos VAR y VARX para ventas generales en clientes Premium, análisis general. Fuente: Elaboración propia.	49
Tabla 28: Modelos VAR y VARX para categoría Vestuario, análisis clientes Premium. Fuente: Elaboración propia.	50
Tabla 29: Modelos VAR y VARX para categoría electrónica, análisis clientes Premium. Fuente: Elaboración propia.	51
Tabla 30: Modelos VAR y VARX para la categoría decoración en clientes Premium, análisis venta Decoración. Fuente: Elaboración propia.	52
Tabla 31: Modelos VAR y VARX para vestuario, clientes no Premium. Fuente: Elaboración propia.	54
Tabla 32: Modelos VAR y VARX para electrónica en clientes no premium. Fuente: Elaboración propia.	55
Tabla 33: Modelos VAR y VARX para decoración en clientes no premium. Fuente: Elaboración propia.	56
Tabla 34: Modelos VAR y VARX para clientes zona norte en venta general. Fuente: Elaboración propia.	57
Tabla 35: Modelos VAR y VARX para clientes zona norte en vestuario. Fuente: Elaboración propia.	58
Tabla 36: Modelos VAR y VARX para clientes zona norte en electrónica. Fuente: Elaboración propia.	59
Tabla 37: Modelos VAR y VARX para clientes en zona norte en decoración. Fuente: Elaboración propia.	60
Tabla 38: Modelos VAR y VARX para la venta general en clientes de zona RM Oriente. Fuente: Elaboración propia.	61
Tabla 39: Modelos VAR y VARX para la categoría vestuario en clientes de zona RM Oriente. Fuente: Elaboración propia.	62
Tabla 40: Modelos VAR y VARX para la categoría electrónica en clientes RM Oriente. Fuente: Elaboración propia.	63
Tabla 41: Modelos VAR y VARX para la categoría decoración en clientes de zona RM Oriente. Fuente: Elaboración propia.	64
Tabla 42: Modelos VAR y VARX para la venta general en clientes de zona RM Poniente. Fuente: Elaboración propia.	65
Tabla 43: Modelos VAR y VARX para la categoría vestuario en clientes RM Poniente. Fuente: Elaboración propia.	66
Tabla 44: Modelos VAR y VARX para la categoría electrónica en clientes zona RM Poniente. Fuente: Elaboración propia.	67
Tabla 45: Modelos VAR y VARX para la categoría decoración en clientes zona RM Poniente. Fuente: Elaboración propia.	68
Tabla 46: Modelos VAR y VARX para venta general en clientes zona Sur. Fuente: Elaboración propia.	69
Tabla 47: Modelos VAR y VARX para la categoría vestuario en clientes zona Sur. Fuente: Elaboración propia.	70

<i>Tabla 48: Modelos VAR y VARX, categoría electrónica clientes zona Sur. Fuente: Elaboración propia.</i>	71
<i>Tabla 49: Modelos VAR y VARX para la categoría decoración en clientes zona Sur. Fuente: Elaboración propia.</i>	72
<i>Tabla 50: Estimación modelos por categorías de producto para clientes premium. Fuente: Elaboración propia</i>	73
<i>Tabla 51: Estimación modelos por categorías de producto para clientes no premium. Fuente: Elaboración propia</i>	74
<i>Tabla 52: Estimación modelos para categoría vestuario por zonas. Fuente: Elaboración propia</i>	74
<i>Tabla 53: Estimación modelos para categoría electrónica por zonas. Fuente: Elaboración propia</i>	75
<i>Tabla 54: Estimación modelos para categoría decoración por zonas. Fuente: Elaboración propia</i>	75

Índice de gráficos

Gráfico 1: Ventas mensuales por canal, periodo 2015-2019. Fuente: elaboración propia.	7
Gráfico 2: Características de los segmentos de clientes según clasificación por canales. Fuente: Elaboración propia.	9
Gráfico 3: Evolución ventas productos primera categorización por canal. Fuente: Elaboración propia	12
Gráfico 4: Evolución de porcentajes de ventas por canal entre 2017 y 2019. Fuente: elaboración propia.	13
Gráfico 5: Métricas de canibalización análisis global. Fuente: Elaboración propia.	27
Gráfico 6: Puntos de corte IRF's de los modelos. Fuente: Elaboración propia.	46
Gráfico 7: IRF con 95% de confianza, modelo de venta general clientes premium. Fuente: Elaboración propia	50
Gráfico 8: IRF con 95% de confianza, modelo de venta vestuario en clientes premium. Fuente: Elaboración propia	51
Gráfico 9: IRF con 95% de confianza, modelo de venta en electrónica clientes premium. Fuente: Elaboración propia	52
Gráfico 10: IRF con 95% de confianza, modelo de venta en decoración para clientes premium. Fuente: Elaboración propia	53
Gráfico 11: Comparación modelos VAR y VARX para venta general, cliente no premium. Fuente: Elaboración propia	iError! Marcador no definido.
Gráfico 12: IRF con 95% de confianza, modelo de venta general clientes no premium. Fuente: Elaboración propia.	54
Gráfico 13: IRF con 95% de confianza, modelo de venta vestuario en clientes no premium. Fuente: Elaboración propia	55
Gráfico 14: IRF con 95% de confianza, modelo de venta electrónica en clientes no premium. Fuente: Elaboración propia	56
Gráfico 15: IRF con 95% de confianza, modelo de venta decoración en clientes no premium. Fuente: Elaboración propia	57
Gráfico 16: IRF con 95% de confianza, modelo de venta general en zona norte. Fuente: Elaboración propia	58
Gráfico 17: IRF con 95% de confianza, modelo de venta vestuario en zona norte. Fuente: Elaboración propia	59
Gráfico 18: IRF con 95% de confianza, modelo de venta electrónica en zona norte. Fuente: Elaboración propia	60
Gráfico 19: IRF con 95% de confianza, modelo de venta decoración en zona norte. Fuente: Elaboración propia	61
Gráfico 20: IRF con 95% de confianza, modelo de venta general en zona oriente. Fuente: Elaboración propia	62
Gráfico 21: IRF con 95% de confianza, modelo de venta vestuario en zona RM Oriente. Fuente: Elaboración propia	63
Gráfico 22: IRF con 95% de confianza, modelo de venta electrónica en zona RM Oriente. Fuente: Elaboración propia	64
Gráfico 23: IRF con 95% de confianza, modelo de venta decoración en zona RM Oriente. Fuente: Elaboración propia	65
Gráfico 24: IRF con 95% de confianza, modelo de venta general en zona RM Poniente. Fuente: Elaboración propia	66
Gráfico 25: IRF con 95% de confianza, modelo de venta vestuario en zona RM Poniente. Fuente: Elaboración propia	67
Gráfico 26: IRF con 95% de confianza, modelo de venta electrónica en zona RM Poniente. Fuente: Elaboración propia	68
Gráfico 27: IRF con 95% de confianza, modelo de venta decoración en zona RM Poniente. Fuente: Elaboración propia	69
Gráfico 28: IRF con 95% de confianza, modelo de venta general en zona RM Sur. Fuente: Elaboración propia	70
Gráfico 29: IRF con 95% de confianza, modelo de venta vestuario en zona Sur. Fuente: Elaboración propia	71
Gráfico 30: IRF con 95% de confianza, venta electrónica en zona RM Poniente. Fuente: Elaboración propia	71
Gráfico 31: IRF con 95% de confianza, modelo de venta decoración en zona RM Sur. Fuente: Elaboración propia	72

ÍNDICE DE ILUSTRACIONES

<i>Ilustración 1: Tipos de clientes según canal de compra. Fuente: elaboración propia.</i>	8
<i>Ilustración 2: Metodología KDD. Fuente: Introduction to Knowledge Discovery and Data Mining. Maimon & Rokach, 2010</i>	17
<i>Ilustración 3: Función de impulso respuesta. Fuente: Elaboración propia</i>	22
<i>Ilustración 4: IRF's ventas globales. Fuente: elaboración propia.</i>	27
<i>Ilustración 5: Comparación métricas premium vs no premium por categoría de producto. Fuente: Elaboración propia.</i>	31
<i>Ilustración 6: Comparación métricas por zona para cada categoría de producto. Fuente: Elaboración propia.</i>	35

1 Introducción

En los últimos años las transacciones comerciales en los canales online han aumentado rápidamente, quitándole protagonismo a las tradicionales tiendas físicas. Ante ello, muchas empresas han incorporado estrategias multicanales para estar en sintonía con el cambio de comportamiento de compra (Kollmann et al. 2012). Chile no es ajeno a esta realidad, según el XI Informe de Tendencias de Medios de Pago (Minsait Payments, 2019) es uno de los líderes de la región, con un 98.3% de penetración del ecommerce. Por otra parte, la Cámara Nacional de Comercio estima que, con respecto al año anterior, las compras online aumentaron un 31.6% en 2017, 39.4% en 2018 y 16% en 2019.

En este contexto de digitalización del comercio, la empresa busca fomentar la venta a través de su canal online, que cuenta con prácticamente la totalidad de los productos comercializados por la firma, pero también introduciendo el concepto de omnicanalidad (integración del canal offline con el online) con el objetivo de generar una sinergia positiva entre ellos (Herhausen, Binder, Schoegel & Herrmann, 2015). En consecuencia, la empresa tiene la iniciativa Click & Collect, que da la opción de comprar online y luego retirar en una tienda a elección sin costo, la ventaja, según estudios de la propia empresa, es que el acudir a una tienda aumenta las probabilidades de comprar en ellas.

Si bien la introducción de un nuevo canal online por parte de las empresas busca llegar a más clientes, es importante señalar que no todas las ventas de los canales digitales son incrementales. En este contexto canibalización hace referencia a la disminución de las ventas del canal offline a causa de que los clientes compran a través del sitio web de la empresa. De forma antónima, se define sinergia positiva como el fenómeno donde la inclusión del nuevo canal genera que las ventas en las tiendas físicas aumenten (Kollmann et al., 2012).

Las ventas del canal físico de la empresa han disminuido un 12% en los últimos 5 años, lo cual se puede deber a distintas razones: las personas están gastando menos, que se encuentren fidelizados por la competencia o bien que estén comprando en la misma empresa, pero por el canal online. Hasta el momento la firma no ha realizado ningún estudio econométrico para determinar cuanta de la disminución de las ventas offline se debe a una

canibalización, por lo que se propone la implementación de Modelos vectoriales autorregresivos para series de tiempo (VAR), que permitan aislar el efecto de interés.

La utilización de modelos VAR en la estimación de canibalización de canales en una empresa de retail chilena no ha sido explorada con anterioridad, por lo que este estudio contribuye a la literatura existente con un nuevo enfoque para abordar el análisis de la relación entre canales.

El estudio se realiza para diferentes tipos de cliente según su comportamiento de compra y zona geográfica donde residen, primero a nivel general de las ventas, y luego para las principales categorías de productos: vestuario, decoración y electrónica.

La investigación es relevante para la empresa debido a que la estimación del nivel de canibalización o sinergia positiva entre los canales impacta en decisiones estratégicas, por ejemplo, evaluar el cierre de tiendas, asignación de fondos para estrategias de marketing según canales de venta o la modificación de los productos expuestos actualmente en las tiendas.

2 Revisión de Literatura

El impacto de la incorporación de un nuevo canal ha sido estudiado en diferentes rubros y desde diferentes puntos de vista a lo largo de los años, variando las conclusiones de un caso a otro. La Tabla 1 sintetiza el principal hallazgo de los estudios más destacados sobre la relación entre los canales de venta.

Autor	Industria	Canales	Hallazgo	Canibalización	Sinergia
Kumar et al. (2005)	Hardware y software	Vendedores, mail, telefónica y online	Consumidores multicanal son más valiosos		✓
Wagner et al. (2013)	Retail	E-channels: sitio web, app y tv	Sinergia y la complementariedad mejora la satisfacción de consumidores		✓
Avery et al. (2012)	Retail	Online y offline	Sinergia cuando el nuevo canal es el offline y canibalización cuando se agrega un canal online	✓	✓
Pauwels and Neslin (2008)	Retail	Venta por catalogo, online y tienda física	Una nueva tienda física genera canibalización solo en catalogo. A nivel general mayores ingresos	✓	✓
Deleersnyder et al. (2002)	Periódicos	Offline y online	Si el nuevo canal online imita bien a los canales ya arraigados, más probable es la canibalización	✓	
Neslin et al. (2006)	Varios	Offline y online	Más canales aumentan las ventas de la empresa		✓
Van Nierop et al. (2011)	Retail	Offline y online informativo	Debido a sitio web menor frecuencia de compra y menos ingresos en offline	✓	
Ansari et al. (2008)	Retail	Offline y online	Clientes multicanal recibieron más marketing y se fortaleció la incidencia de compra.		✓

Tabla 1: Hallazgos de los principales estudios sobre relación entre canales. Fuente: Huang, Lu & Ba, 2015

Abrir un nuevo canal y diseñar una estrategia de omnicanalidad ayuda a que los clientes aumenten su lealtad con la empresa (T. Kollmann, A. Kuchertz & I. Kayseren, 2012). No obstante, también abre la posibilidad de una canibalización parcial o total (Montoya-Weiss, 2003), el fenómeno es más probable cuando el nuevo canal ofrece opciones que ayudan a los clientes a maximizar su utilidad (Deleersnyder et al., 2002). En este contexto, estudios anteriores evidencia que la incorporación de un nuevo canal impacta negativamente en la cantidad de visitas de los clientes a las tiendas físicas (Nierop et al. 2011).

Por otra parte, un nuevo canal puede tener un efecto positivo en canal original y en la empresa en general, debido a que aumenta la distribución de productos

e información de la empresa, ya que los clientes están más expuestos a esfuerzos de marketing. Neslin en su estudio "*Challenges and opportunities in multichannel customer management*" describe que en este contexto la firma tiene a favor del "*Efecto de disponibilidad*" pues se abre una nueva vía para comunicarse con nuevos y viejos clientes (Neslin et al., 2006).

En consecuencia, si bien la empresa plantea como hipótesis que hay un nivel importante de canibalización y que afecta a todos los productos, tipos de clientes y zonas geográficas de Chile, es necesario realizar un estudio para determinar la naturaleza de la relación entre los canales.

Se destacan tres estudios que están directamente relacionados a este trabajo. El primero de ellos estima el efecto de canibalización en un minorista de China que expandió su servicio web a una aplicación móvil. Aplicaron el método cuasiexperimental Diferencia en Diferencia, justificando que el ingreso del nuevo canal marcó un claro antes y después en el tiempo. Se concluyó que el efecto de canibalización puede ser total o parcialmente contrarrestado por una sinergia positiva (Huang, Lu & Ba, 2015). Se descarta replicar este tipo de modelos debido a que, en el contexto del caso de estudio, la introducción del canal online no marcó un claro antes y después en las ventas de la empresa, debido a que era ineficiente en diversas formas y poseía un bajo porcentaje de los productos comercializados por la firma, en consecuencia, el canal online paso años sin ser realmente un sustituto del canal original.

En segundo lugar, se encuentra una investigación enfocada en la industria de los periódicos en Reino Unido y Los Países Bajos. Para determinar la canibalización, desarrollaron un análisis empírico mediante series de tiempo. Del cual dedujeron que un efecto de canibalización es más probable cuando el nuevo canal imita los aspectos favorables del canal original y agrega características atractivas para los clientes (Deleersnyder et al., 2002).

Si bien estimar el fenómeno mediante series de tiempo se considera un buen enfoque, este no incluye directamente la relación entre los canales. Tomando este aspecto, se considera la opción de modelar la situación con una metodología más sofisticada: Vectores autorregresivos para series de tiempo. También conocidos como modelos VAR, fueron propuestos por Christopher Sims en la década de los 80's y consideran un set de series de tiempo convertidas en un sistema de ecuaciones que caracterizan interacciones

sincrónicamente en un grupo de variables de interés, llamadas endógenas (Jame, Stock, Mark & Watson, 2012). Este tipo de análisis son útiles para definir la estructura y magnitud de la relación entre las variables. (Goic, Álvarez & Montoya, 2017). Por otra parte, este tipo de modelos permite estudiar la relación de canales cuando no existe un punto clave en la historia que permite diferenciar un antes y un después en el historial de ventas, siendo también replicable a diferentes escenarios.

3 Antecedentes y análisis preliminar de los datos

3.1 Evolución de las ventas

En las ventas históricas de la empresa el 2017 marcó un antes y un después en la evolución. Tal y como se muestra en la Tabla 2, hasta dicho año la empresa había estado aumentando sus ventas generales [MM\$¹], pero desde 2018 los ingresos han disminuido. Posibles razones son que que la economía de Chile en general ha tendido a la baja en los últimos años o que los clientes estén comprando en la competencia. En la misma tabla, también se detalla la evolución de las ventas por canal, se observa que los ingresos mediante el sitio web han ido aumentando en los años y se destaca que tuvo su mayor crecimiento en el año 2017, lo cual se correlaciona con que la venta offline ha estado disminuyendo desde dicho año, siendo el decrecimiento más importante en las tiendas físicas en 2019. Lo anterior sugiere cuestionarse si existe una canibalización del canal online a las tiendas físicas.

	Venta gral [MM \$]	Diferencia	Venta offline [MM \$]	Diferencia	Venta online [MM \$]	Diferencia
2015	\$ 1.4	-	\$ 1.2	-	\$ 0.2	-
2016	\$ 1.6	13%	\$ 1.4	11%	\$ 0.2	32%
2017	\$ 1.7	3%	\$ 1.3	-3%	\$ 0.3	34%
2018	\$ 1.6	-3%	\$ 1.2	-7%	\$ 0.4	19%
2019	\$ 1.5	-5%	\$ 1.1	-12%	\$ 0.4	17%

Tabla 2: Evolución ventas anuales. Fuente: Elaboración propia.

A pesar de que las estadísticas anuales son importantes, en ellas se esconden detalles que solo quedan a la vista cuando se realiza un análisis más desagregado en el tiempo. Al estudiar el comportamiento de las ventas, como se muestra en el Gráfico 1, se aprecian dos estacionalidades. La primera de ellas es el mes de Diciembre, donde debido a la navidad se registran las mayores ventas del año en el canal offline. En segundo lugar, se tienen los eventos *Cyber* que consisten en ventas especiales en el canal online con ofertas de gran impacto, marcando peaks cada vez mayores, disminuyendo la brecha entre los canales. Octubre de 2019 ha sido el único mes donde el sitio web vendió más que las tiendas físicas. Sin embargo, la situación no es solo mérito del canal online pues, en dicho mes comenzó un periodo de movilización histórico en Chile conocido como "*Estallido social*" que obligó a cerrar algunos centros comerciales y a reducir los horarios en otros, sumado a que en los primeros días del mes se realizó un evento *cyber*.

¹ Miles de millones

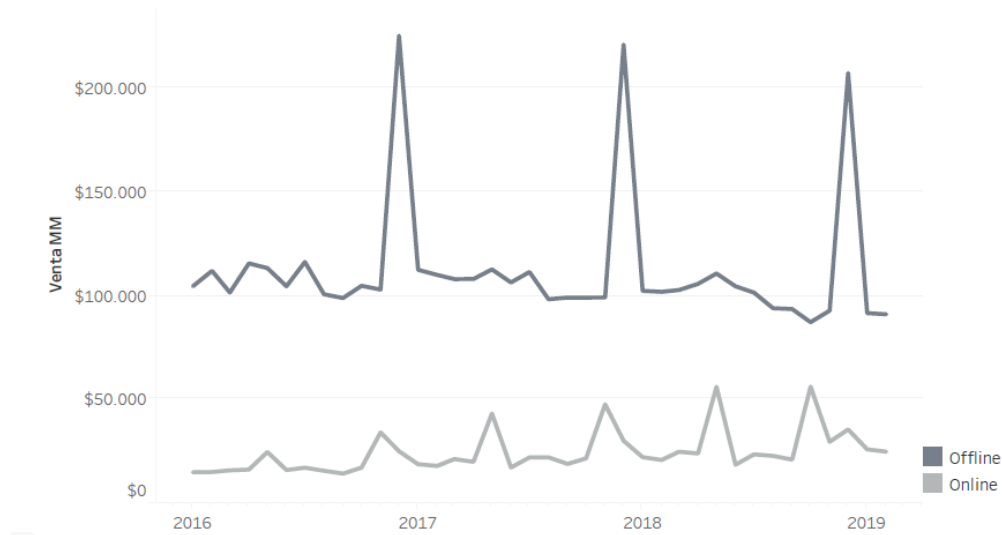


Gráfico 1: Ventas mensuales por canal, periodo 2015-2019. Fuente: elaboración propia.

De esta forma, el Gráfico 1 sugiere que es importante considerar una agrupación de ventas mensuales junto con la consideración de incluir en los modelos una distinción para los meses con eventos cyber y navidad.

3.2 Clientes

Los clientes son un pilar fundamental en toda empresa por lo que es importante conocerlos. Más de 9 millones de personas han realizado al menos una compra en la empresa en los últimos 5 años. Aquellos que indican un RUT válido al momento de pagar, son reconocidos como clientes identificados; para ellos es posible realizar cruces de información desde diferentes fuentes logrando caracterizarlos con atributos sociodemográficos, entre otros aspectos. En la Tabla 4, se detallan los porcentajes de ventas [\$MM] y boletas que han sido identificadas en promedio en los último 5 años.

Ventas identificadas			Boletas identificadas		
General	Offline	Online	General	Offline	Online
82%	78%	100%	76%	73%	100%

Tabla 3: Identificación de ventas y boletas como promedio de los últimos 5 años. Fuente: Elaboración propia.

Sobre la cantidad de clientes totales de la empresa que son identificados a través de los años, estos han tendido a aumentar. Sin embargo, si se compara contra el año anterior, la diferencia ha sido cada vez más pequeña (Tabla 5), lo cual se correlaciona con la disminución de los ingresos de la empresa a partir del año 2017. Por otra parte, en la misma tabla se aprecia como los clientes que compran en el canal offline han disminuido desde 2018 y aquellos que compran online han siempre aumentado, aunque en menor medida en los últimos dos años.

	# Clientes empresa	Variación	# Clientes offline	Variación	# Clientes online	Variación
2015	5.12	-	5.0		0.7	-
2016	5.52	7.2%	5.4	7.0%	0.9	20.8%
2017	5.76	4.2%	5.6	3.4%	1.3	26.3%
2018	5.78	0.3%	5.5	-0.7%	1.5	17.1%
2019	5.79	0.1%	5.4	-1.9%	1.8	14.5%

Tabla 4: Diferencia porcentual en cantidad de clientes identificados. Fuente: Elaboración propia.

Los clientes identificados son categorizados anualmente en 3 categorías según el historial de canales que utilizaron en el año anterior, tal y como se muestra en la Ilustración

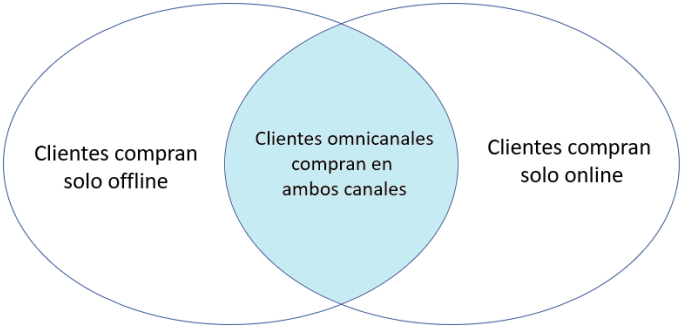


Ilustración 1: Tipos de clientes según canal de compra. Fuente: elaboración propia.

El Gráfico 2 resume las dos principales características de la clasificación de clientes según canales de compra: el nivel de venta que ha representado en el último tiempo y su importancia relativa en la cantidad de clientes (porcentaje indicado sobre cada línea de venta).

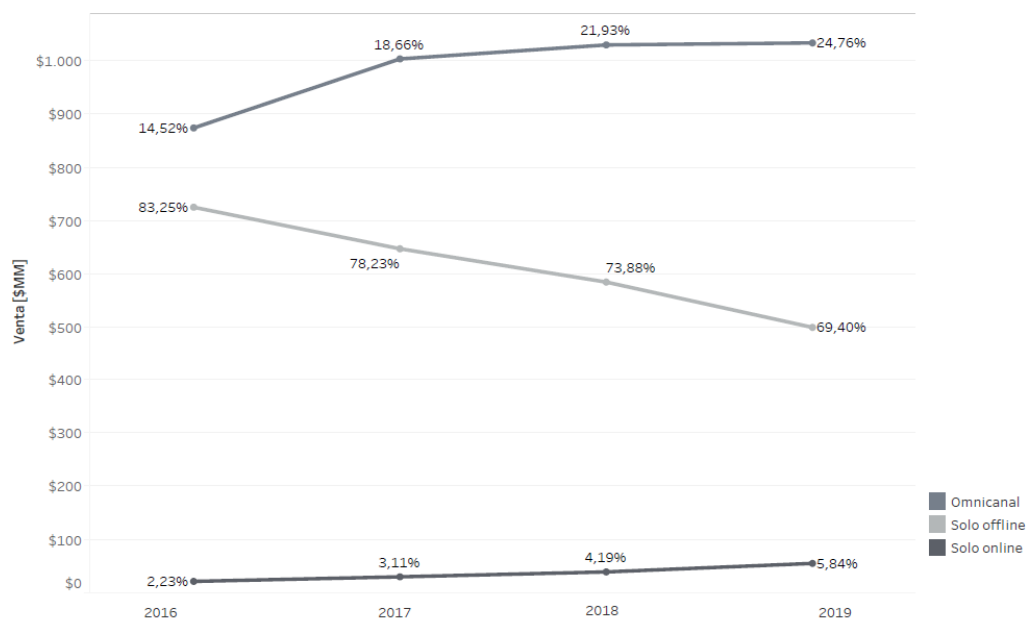


Gráfico 2: Características de los segmentos de clientes según clasificación por canales. Fuente: Elaboración propia.

Desde el gráfico anterior se desprende que los clientes que más han comprado a la empresa son los omnicanales, además éstos junto con los solo online han ido aumentando en el tiempo a costa de las ventas que antes se realizaban a clientes solo offline, evidenciando un cambio en el comportamiento de compra en cuanto a los canales. En línea con lo anterior y considerando el aporte monetario de cada segmento, se destaca que, desde 2017 los clientes omnicanales han ganado protagonismo superando con creces las ventas offline, lo cual sugiere la hipótesis de que existe canibalización entre los canales.

En lo que sigue, se muestran estadísticas relevantes de los clientes. Los valores fueron calculados como promedio de los últimos 5 años (2015-2019).

La empresa está presente a lo largo de todo el país, es por ello por lo que clasifican a los clientes en 4 grandes zonas. Cabe destacar que más del 60% de las ventas de la empresa se realizan en la Región Metropolitana.

Zona	Territorio	Cantidad de tiendas
Norte	Desde región I hasta IV más XV	15
RM Oriente	Lo Barnechea, Vitacura, Las Condes, Providencia, La Reina y Ñuñoa	8
RM Poniente	Todas las comunas de RM que no están incluidas en la zona Oriente	12
Sur	Todas aquellas comunas y regiones de Chile no incluidas en las zonas anteriores	8

Tabla 5: Descripción de zonas y cantidad de tiendas que posee cada una, promedio en periodo 2015-2019. Fuente: Elaboración total.

Considerando la Tabla 6, se tiene que en todo el país el canal más utilizado es el offline. Sin embargo, las proporciones del uso de canales difiere de una zona a otra. Se destaca RM Oriente que posee un mayor porcentaje de ventas realizadas por el canal online que el resto de las zonas, siendo Sur la que posee un menor nivel. Junto con lo anterior la Tabla 6 se detalla la proporción de ventas por canal en cada zona para cada tipo de productos, se destaca que las mayores diferencias se tienen en decoración donde la distribución de ventas en cada canal rodea el 50% con excepción de RM Oriente donde predomina las compras en las tiendas físicas (70%).

	Venta general [\$MM]		Venta vestuario[\$MM]		Venta electrónica [\$MM]		Venta decoración [\$MM]	
	Offline	Online	Offline	Online	Offline	Online	Offline	Online
Norte	78%	22%	91%	9%	68%	32%	46%	54%
RM Oriente	87%	13%	97%	3%	62%	38%	70%	30%
RM Poniente	79%	21%	91%	9%	69%	31%	53%	47%
Sur	80%	20%	91%	9%	72%	28%	51%	49%

Tabla 6: Proporciones de transacciones por canal en cada zona, periodo 2015-2019. Fuente: Elaboración propia.

A partir de la información anterior, se propone que el análisis debe considerar estimaciones para cada una de las zonas, pues el comportamiento de compra y las características de cada zona difieren entre sí. En conjunto con lo anterior, los datos sugieren incluir en los modelos la cantidad de tiendas disponible en los periodos de tiempo y zonas a analizar, lo que también tiene el beneficio de permitir incorporar el hecho de que durante el "Estallido social" algunas tiendas permanecieron cerradas o con horarios acotados con mayor impacto en la Región Metropolitana.

Otro subconjunto de clientes importantes, son los Premium, quienes cumplen con haber comprado un monto igual o superior a \$500.000 el año anterior y haber registrado al menos 7 visitas² a la empresa en el mismo periodo. En los últimos 5 años, en promedio este tipo de clientes corresponde al 12% de los clientes identificados y son responsables del 53% de las ventas, he ahí su gran importancia para la empresa. Sus principales atributos son comparados con el resto de los clientes en la Tabla 7.

² Visita se define como una combinación única de RUT (válido), número de boleta y fecha.

		Premium	Resto de clientes
Visitas	Promedio por año	15	3
Zona	Norte	20%	23%
	RM Oriente	10%	6%
	RM Poniente	41%	39%
	Sur	29%	32%
Proporciones ventas por canal	Offline	72%	82%
	Online	28%	18%
Porcentaje ventas por categoría	Vestuario	43%	60%
	Decoración	13%	9%
	Electrónica	44%	31%

Tabla 7: Comparación clientes premium vs resto de clientes, en periodo 2015-2019. Fuente: Elaboración propia

Se destaca que, en promedio, los clientes Premium realizan 5 veces más visitas que el resto de los clientes, además están en mayor proporción en la Región Metropolitana. Es por ello que es importante hacer una distinción entre los tipos de clientes en los análisis de canibalización, pues sus comportamientos de compra son diferentes.

3.3 Productos

La empresa ofrece gran variedad de productos los cuales se clasifican en: decoración, electrónica y vestuario. Es esta última la más importante, pues representa más del 55% de las ventas, seguida por electrónica y luego decoración.

El Gráfico 3, indica los niveles de ventas mensuales por canal para cada una de las tres principales categorías de productos. Se destaca la categoría decoración que en 2015 era vendida mayoritariamente por el medio offline, sin embargo, este canal perdió protagonismo y en 2019 las ventas fueron principalmente online. Lo anterior, se debe a una combinación de razones, entre las que se destaca: la posible canibalización entre canales y el hecho de que en las tiendas físicas se exhiban en menor cantidad este tipo de productos. Por otra parte, en las otras categorías tipos de productos las ventas online han disminuido la brecha con las ventas offline pero sigue siendo superior.

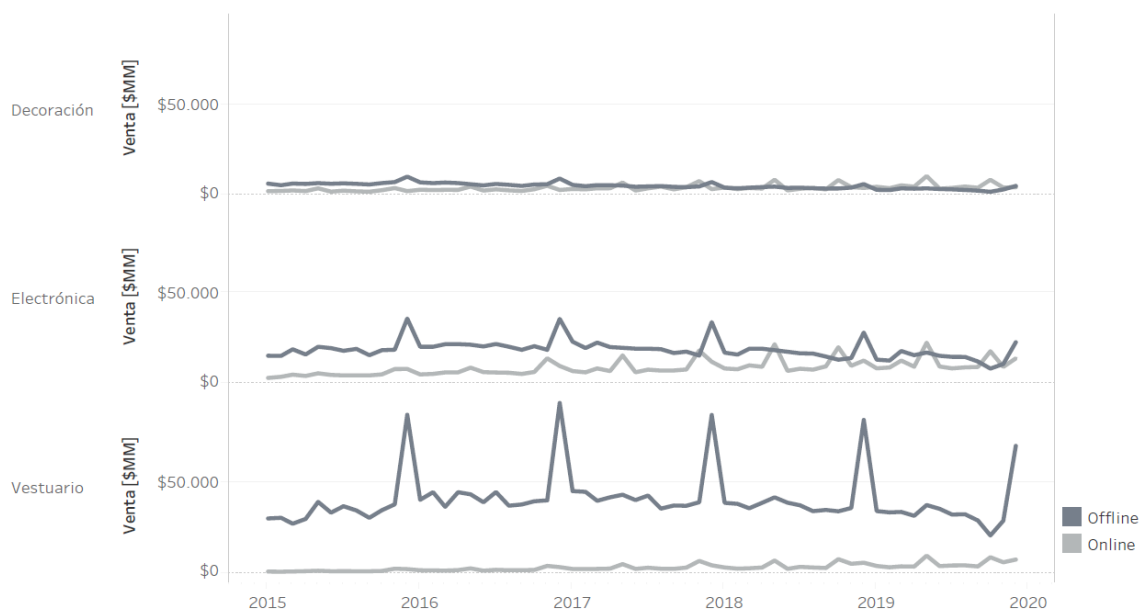


Gráfico 3: Evolución ventas por tipo productos en cada canal. Fuente: Elaboración propia

La Tabla 8 expone el porcentaje de ventas por canal para las diferentes combinaciones tipo de producto y zona geográfica.

	Zona	Offline	Online
Decoración	Norte	41%	59%
	RM Oriente	45%	55%
	RM Poniente	43%	57%
	Sur	46%	54%
Electrónica	Norte	66%	34%
	RM Oriente	51%	49%
	RM Poniente	64%	36%
	Sur	70%	30%
Vestuario	Norte	90%	10%
	RM Oriente	87%	13%
	RM Poniente	87%	13%
	Sur	90%	10%

Tabla 8: Proporción ventas por canal según zona y tipo de producto, promedio 2015-2019. Fuente: Elaboración propia.

Se mantiene la tendencia de que los productos de decoración se venden más en el sitio web y que las otras dos categorías más por el canal offline. Sin embargo, las proporciones son diferentes en las zonas. El caso más drástico es en la venta de electrónica, donde RM Oriente vende 51% Offline y la zona Sur casi 20 puntos porcentuales más.

En base a la información anterior sobre los tipos de productos comercializados por la empresa, se distinguen que entre sí sus ventas se comportan de manera distinta, lo cual es importante considerar en el desarrollo de modelos.

3.4 Analisis preliminar de relación entre canales

Los antecedentes de la empresa evidencian que las ventas del canal offline han disminuido a través de los años y las realizadas por el medio online han aumentado. Además, la dinámica entre los canales de compra es distinta para cada tipo de cliente, zona geográfica y categoría de productos, por lo que se propone un análisis de canibalización para cada uno de ellos, además de uno global que incorpore todas las ventas de la empresa, lo anterior para el periodo entre 2017 y 2019. Como primer acercamiento se analizan las variaciones de las ventas en cada canal para cada tipo de producto, cliente y zona en el periodo de interés.

A continuación, en el Gráfico 4 se observa la evolución del porcentaje de ventas atribuible a cada uno de los canales entre el 2017 y 2019, además de destacar el aumento de los puntos porcentuales en el caso de la venta online.

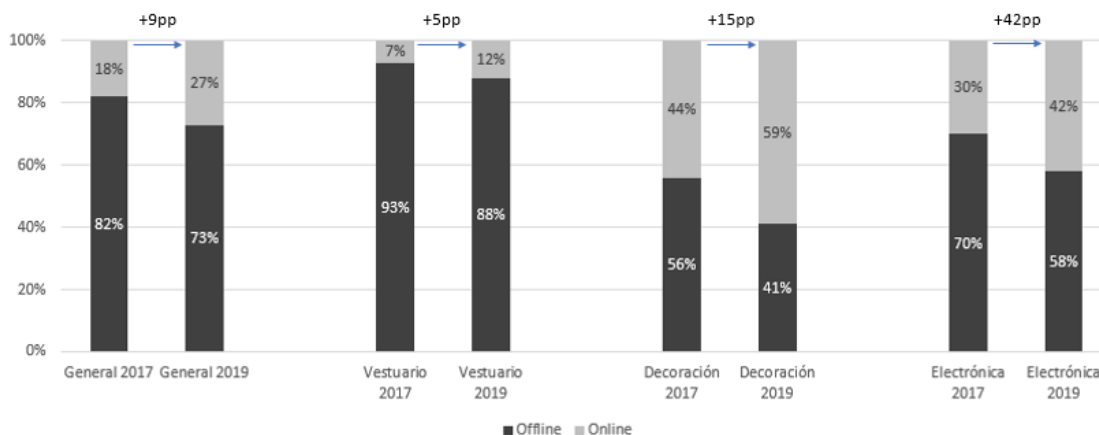


Gráfico 4: Evolución de porcentajes de ventas por canal entre 2017 y 2019. Fuente: elaboración propia.

Los datos del gráfico anterior sugieren la hipótesis de que existe una canibalización general y que esta además es mayor en electrónica, luego en decoración y finalmente en vestuario.

Por otra parte, se tiene la Tabla 9 la cual muestra el claro aumento del porcentaje de ventas del canal online en cada una de las categorías tanto para los clientes premium como los que no pertenecen a dicho grupo. Los datos orientan a que podría existir una canibalización que sería mayor en decoración para ambos tipos de clientes.

	Categoría	Dif. % vtas online
Clientes Premium	General	11 pp
	Vestuario	10 pp
	Decoración	15 pp
	Electrónica	11 pp
Clientes no Premium	General	8 pp
	Vestuario	6 pp
	Decoración	14 pp
	Electrónica	9 pp

Tabla 9: Aumento de porcentaje de ventas en canal online entre 2017 y 2019 por tipo de clientes. Fuente: elaboración propia.

Al realizar el mismo análisis, pero a nivel de zonas, se ve la clara tendencia de que el canal online ha tenido un aumento en el porcentaje de ventas de la empresa. Al igual que el caso anterior, las mayores variaciones se aprecian en la categoría decoración en cada una de las zonas.

	Categoría	Dif. % vtas online
Norte	General	10 pp
	Vestuario	7 pp
	Decoración	14 pp
	Electrónica	10pp
RM Oriente	General	9 pp
	Vestuario	8 pp
	Decoración	12 pp
	Electrónica	7 pp
RM Poniente	General	9 pp
	Vestuario	8 pp
	Decoración	13 pp
	Electrónica	8 pp
Sur	General	9 pp
	Vestuario	7 pp
	Decoración	17 pp
	Electrónica	10 pp

Tabla 10: Aumento de porcentaje de ventas en canal online entre 2017 y 2019 por zona. Fuente: elaboración propia.

El análisis preliminar de los datos sugiere que existe canibalización en todas las combinaciones de productos-clientes-zonas e incluso muestra que es mayor en ciertos casos. Sin embargo, un estudio tan simple deja fuera otros factores relevantes, como la cantidad de tiendas abiertas en cada mes y la consideración del efecto de estacionalidades de navidad y eventos cyber. Es por ello que se considera la hipótesis de que existen una predominancia de canibalización por sobre la sinergia positiva entre los canales, pero se propone un análisis más sofisticado para obtener estimaciones más exactas.

4 Objetivos

4.1 Objetivo general

Identificar y estimar, el impacto de ventas del canal online de una empresa de retail sobre las ventas de su canal offline, para la toma de decisiones estratégicas.

4.2 Objetivos específicos

- Determinar y cuantificar el efecto de canibalización y/o sinergia positiva a nivel de ventas generales de la empresa.
- Establecer y estimar el efecto de canibalización y/o sinergia positiva a nivel de categoría de productos.
- Determinar y valorar el efecto de canibalización y/o sinergia positiva por zona geográfica y tipo de clientes.

5 Metodología: Knowledge Discovery in Data Base

Con el fin de encontrar insights relevantes y cumplir los objetivos planteados se utiliza la metodología KDD, descrita en la Ilustración 1. Principalmente consiste en un proceso de 9 etapas ordenadas de forma de extraer información relevante de grandes bases de datos.

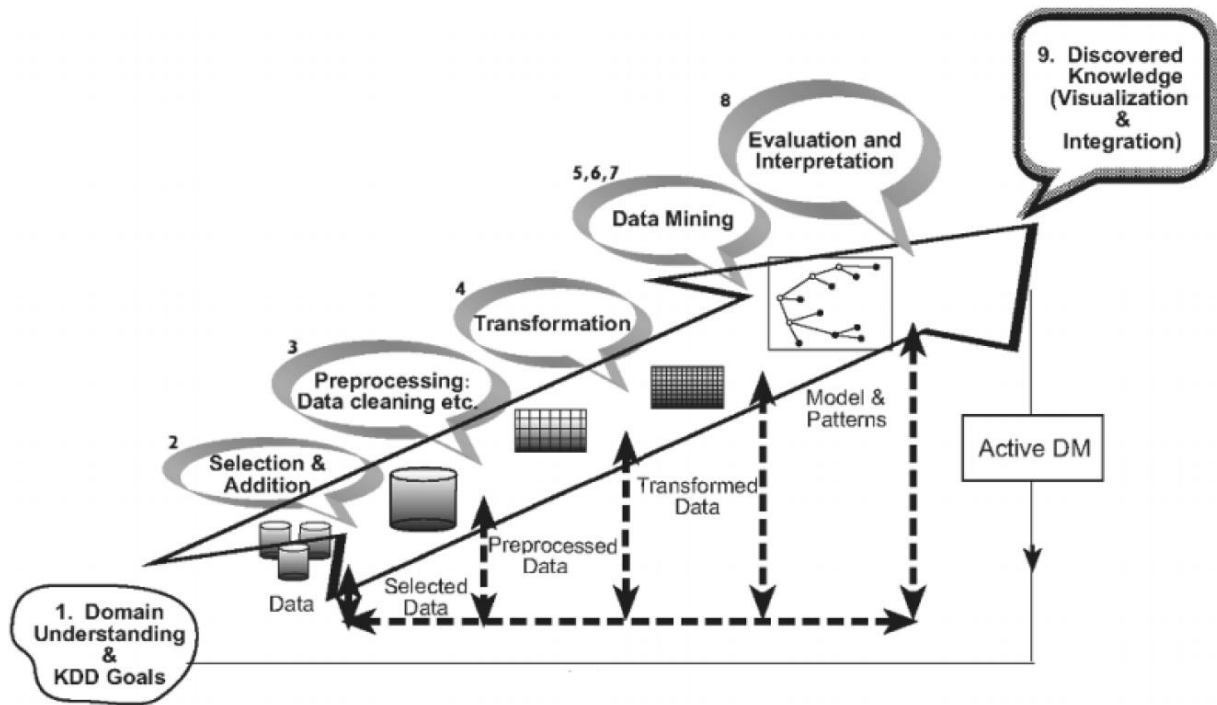


Ilustración 2: Metodología KDD. Fuente: *Introduction to Knowledge Discovery and Data Mining*. Maimon & Rokach, 2010

5.1 Abstracción del escenario

Consiste en entender el contexto de los datos, para luego proponer soluciones viables a la problemática planteada y definir metas al respecto. En esta línea y dadas las características y antecedentes del estudio se definen las siguientes métricas.

$$\text{Variación sobre venta Online}_a = \frac{\Delta Y_{\text{Offline},a} [Y_{\text{Online},a}]}{Y_{\text{Online},a}} \quad \text{Expresión (1)}$$

$$\text{Variación sobre venta Offline}_a = \frac{\Delta Y_{\text{Offline},a} [Y_{\text{Online},a}]}{Y_{\text{Offline},a}} \quad \text{Expresión (2)}$$

Donde el índice $a \in \{2017, 2018, 2019\}$ es el año. $\Delta Y_{Offline,a}[Y_{Online,a}]$ representa la variación de las ventas en el canal offline en el año a debido a las ventas del canal online en el mismo periodo. $Y_{Online,a}$ e $Y_{Offline,a}$ indican las ventas en el año a en cada uno de los canales.

De esta forma, la Expresión 1 se interpreta como el porcentaje de las ventas del sitio web que son atribuibles a la variación en el canal offline. En cambio, la Expresión 2 representa el porcentaje en que han variado las ventas en tienda debido al comportamiento del canal online.

Cuando los valores de las métricas son positivos, estamos frente a una sinergia entre canales, es decir, en ambos las ventas han aumentado. Caso contrario se trata de una canibalización y cuando esto ocurre se tiene que $(100\% - |Variación\ sobre\ venta\ Online_a|)$ corresponde al porcentaje de las ventas incrementales que aporta el sitio online en el año a .

5.2 Selección de datos, pre- procesamiento y transformación

Los datos utilizados son resultado del cruce de información proveniente de diferentes fuentes oficiales de la empresa, de forma de tener los datos transaccionales y a la vez información sociodemográfica, para así poder hacer distinción entre si el cliente es identificado o no y poder agrupar por tipo de cliente, zonas geográficas y tipo de producto comprado de forma mensual entre Enero de 2016 y Diciembre 2019. Dicha información fue manipulada y preparada en dos programas computacionales: Tableau y RStudio. Se buscó limpiar la información corroborando que los datos coincidieran con la información que utiliza la empresa en diferentes reportes y estructurarla de forma que estuviera en línea con el formato correspondiente a los modelos VAR.

5.3 Minería de datos, evaluación e interpretación

Estimar el efecto del aumento de las ventas del canal online sobre el offline, incluye complejos análisis dinámicos entre los canales que también podrían depender del pasado. Esto junto con la estructura de los datos disponible en

formato serie de tiempo, apoyan la propuesta de utilizar Modelos Vectoriales Autorregresivos, donde las ventas en los canales son las variables endógenas del sistema.

Todo modelo VAR debe tener especificado su orden, el que referencia a la cantidad de rezagos que considera. Una extensión son los modelos VARX(p), donde adicionalmente se incluyen variables exógenas que ayudan a precisar la relación entre las endógenas. Importante mencionar que este tipo de modelos puede ser estimado de la misma forma que un VAR (Tsay, 2014).

Un modelo de orden p con bivalente, VARX (p), se detalla en la Expresión (3), la cual muestra de forma general la estructura de un de esta naturaleza. Para el análisis de cada categoría de producto, tipo de cliente y zona geográfica se utilizó un modelo personalizado, tal y como se explicará más adelante, según una serie de métricas se determinó la cantidad de rezagos y variables exógenas para cada situación en específica.

$$\begin{bmatrix} Y_{Offline,t} \\ Y_{Online,t} \end{bmatrix} = \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} + \sum_{j=1}^{j=p} \begin{bmatrix} \beta_{t-j,11} & \beta_{t-j,12} \\ \beta_{t-j,21} & \beta_{t-j,22} \end{bmatrix} \begin{bmatrix} Y_{Offline,t-j} \\ Y_{Online,t-j} \end{bmatrix} + \begin{bmatrix} a_{t,11} & a_{t,12} & a_{t,13} \\ a_{t,21} & a_{t,22} & a_{t,23} \end{bmatrix} \begin{bmatrix} X_{navidad,t} \\ X_{cyber,t} \\ X_{tiendas,t} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{bmatrix} \text{ Expresión (3)}$$

Donde, $Y_{Offline,t}$ e $Y_{Online,t}$ corresponden a las variables endógenas y representan las ventas en cada uno de los canales en el periodo t y β_{t-j} son los coeficientes de los rezagos de dichas variables. Debido a las estacionalidades que indicaron los datos en los análisis preliminares de las ventas se tienen dos variables exógenas de control, $X_{navidad,t}$ y $X_{cyber,t}$, la cuales son binarias que toman valor 1 si en el periodo t es navidad o si hay eventos cyber respectivamente. Se cuenta también con una tercera variable exógena, $X_{tiendas,t}$, que indica la cantidad de cantidad de tiendas físicas disponibles que corresponde a una numérica que representa el promedio de las tiendas que estuvieron abiertas en el periodo t, con ello se busca aislar el efecto que provoca que durante el periodo estudiado se hayan abierto y cerrado tiendas a lo largo de Chile. Además de ilustrar el fenómeno ocurrido en Octubre del 2019 (Estallido Social). Finalmente, $\alpha_{t,nm}$ son los coeficientes de las variables exógenas y ε_t son errores asociados.

Para estimar un modelo autorregresivo es necesario que cada serie de tiempo sea estacionaria, en otras palabras, que tenga una matriz de covarianza y media que no varíen en el tiempo, lo que es comprobable mediante la prueba de Dickey-Fuller (Shumway & Stoffer, 2006). En este caso el 96% de las series

son estacionarias con un 95% de confianza y el 100% de ellas lo son con un nivel de confianza 90%, por lo que se trabaja con este último.

Los modelos deben cumplir ciertos supuestos para que la estimación sea adecuada, consistente y tenga coeficientes insesgados. Primero, que los residuos cumplan con distribuir normal y que sean homocedásticos, es decir, la distribución condicional de los errores dada las variables de las series de tiempo sean constante y que no dependan de dicha variable. En caso contrario los errores son heterocedásticos. Además, no debe existir autocorrelación en los residuos. Por otra parte, el modelo debe ser estable, es decir, que las raíces del polinomio característico de la matriz de coeficientes β_{t-j} de la Expresión (3), se encuentran dentro del círculo unitario.

Para la aplicación de este tipo de modelos primero se define un 80% de la data (38 semanas) para calibración y un 20% para validación (10 semanas). Luego se estiman y comparan una serie de modelos VAR y VARX que difieren entre sí en la cantidad de rezagos y variables exógenas. Ahora bien, ¿Cómo se define cual es el orden adecuado de un modelo VAR y las variables exógenas? Primero se considera si cumplen o no los supuestos mencionados con anterioridad. Luego se calculan las siguientes métricas para evaluar el ajuste de los modelos, así como su capacidad de predecir (Stock & Watson,2012)(Chai & Draxler,2014).

- Criterio de Información de Akaike (AIC):

$$AIC(k) = -2 * \ln(L) + 2k \quad \text{Expresión (4)}$$

Donde k es la cantidad de retardos considerados y L es el valor máximo de la función de verosimilitud del modelo. Se busca el mínimo valor posible.

- Criterio de Información de Bayes (BIC): también se busca el mínimo valor posible, y su fórmula es:

$$BIC(k) = -2 * \ln(L) + k * \ln(n) \quad \text{Expresión (5)}$$

Donde n es número de observaciones de la data y los otros parámetros mantienen la notación del criterio anterior.

- Log Verisimilitud = Logaritmo natural del máximo valor de la función de verosimilitud del modelo. Entre mayor sea, mejor se considera.
- Error absoluto medio (MAE, por sus siglas en inglés): calculado de la siguiente forma:

$$MAE = \frac{\sum_{i=1}^n |\varepsilon_i|}{n} \quad \text{Expresión (6)}$$

Donde n corresponde al número de observaciones y ε_i al error asociado a la predicción del modelo. Se busca que sea el mínimo posible.

- Error cuadrático medio (RMSE): El ideal es obtener el mínimo valor, su fórmula de cálculo es como sigue:

$$RMSE = \frac{1}{n} \sum_1^n [Y_i^e - Y_i]^2 \quad \text{Expresión (7)}$$

Donde Y_i corresponde al valor real de la variable en la observación i e Y_i^e a la predicción hecha por el modelo. Además, n corresponde al número de observaciones totales de la muestra.

En la sección de Anexo se detallan las comparaciones de los modelos para cada estimación realizada.

Una vez elegido el "mejor" modelo VAR o VARX, se aplica para obtener vectores de impulso respuesta (IRF) los cuales permiten estimar la reacción hipotética de una variable endógena en el tiempo debido a un impulso en otra variable de endógena en el instante s formando el vector de la expresión 8. Importante considerar que la función no es exacta y se encuentra asociada a un intervalo de confianza (Stock & Watson, 2012).

$$\begin{bmatrix} \Delta Y_{Offline,s} [\Delta Y_{Online,s}] \\ \dots \\ \Delta Y_{Offline,s+n} [\Delta Y_{Online,s}] \end{bmatrix} \quad \text{Expresión(8)}$$

Donde, $\Delta Y_{Offline,t} [\Delta Y_{Online,T}]$ corresponde a la variación del canal offline en el tiempo t debido a un impulso en el canal online en el periodo T . En la sección de Anexo se encuentran las IRF's de los análisis realizados.

La Ilustración 3 muestra gráficamente una función de impulso respuesta estimada según el modelo VAR correspondiente. El área bajo la curva corresponde a la respuesta acumulada de la variable endógena ventas offline debido al impulso en el tiempo 0 en la variable venta online, es decir, representa las ventas [\$MM] perdidas o ganadas en las tiendas físicas debido al impulso en las ventas en el sitio web. En consecuencia, es necesario establecer el horizonte de tiempo a considerar. Se establece como punto crítico, aquel que antecede un conjunto de, al menos, 5 periodos consecutivos donde se cumpla:

$$\left| \frac{1 - IRF_{t+1}}{IRF_t} \right| < 10\% \quad \text{Expresión (9)}$$

Existe un $\sim 15\%$ donde no se alcanzan los 5 periodos consecutivos. En ellos se tomó como punto crítico aquel que antecede al conjunto donde la condición se cumplió por la mayor cantidad de periodos consecutivos. El detalle de dichos puntos se encuentra en la sección Anexo 10.1. Con el fin de poder realizar una posterior comparación entre métricas de canibalización, se plantea un único punto de corte que corresponda al promedio de los puntos críticos de los análisis ($t=14$).

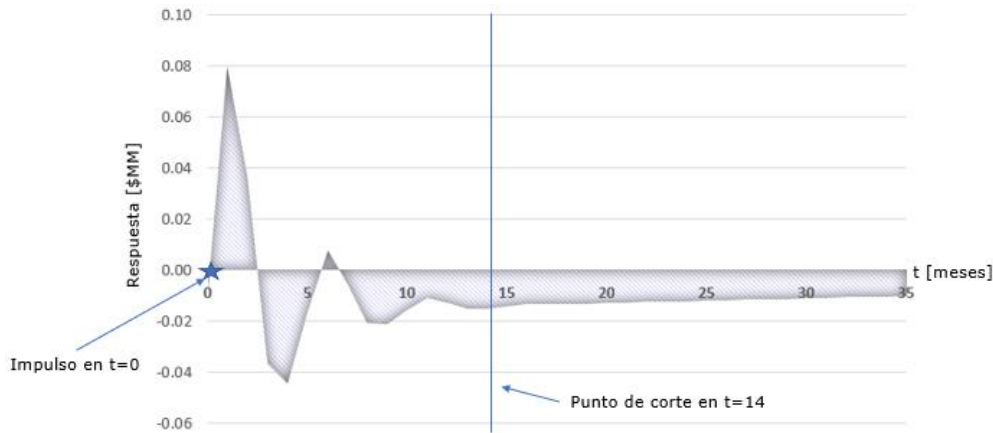


Ilustración 3: Función de impulso respuesta. Fuente: Elaboración propia

Para normalizar la variación del canal offline se calcula la respuesta mensual promedio de la venta en tienda, como se establece en la expresión 10.

$$Rsta\ mensual\ promedio = \frac{\text{Área bajo la curva}}{\text{Punto de corte}} \quad \text{Expresión (10)}$$

El paquete (vars) utilizado en R Studio para estimar tanto los modelos como la IRF asociada toma por defecto un impulso correspondiente a una desviación estándar en las ventas online (δ_a). Para obtener la respuesta hipotética que ha tenido el canal offline debido a las ventas mensuales promedio del canal online se plantea una proporcionalidad lineal que deriva en la Expresión 11.

$$Variación_{offline,a} = \frac{Rsta\ mensual\ promedio_a \cdot Venta\ mensual\ promedio_{online,a}}{\delta_a} \quad \text{Expresión (11)}$$

Donde a representa el año $\in \{2017, 2018, 2019\}$. Una vez estimada la Expresión 11, es posible calcular las métricas definidas en la subsección 5.1 y por ende el nivel de canibalización o sinergia positiva entre los canales.

Importante destacar que los resultados que se obtienen mediante la metodología planteada indican, con cierto nivel de confianza, como las ventas [MM] en el canal offline han variado considerando las transacciones que se han realizado por el medio online. En caso de evidenciar que hay ventas que han migrado de las tiendas físicas al sitio web, se trataría de una canibalización de canales, sin importar si los clientes por su cuenta han preferido comprar digitalmente o si la empresa consciente del cambio de comportamiento de compra haya o no decidido cambiar la oferta de productos de los canales para hacer más atractivo el online.

En cuanto a la interpretación e implicancias directas de los coeficientes asociados a los modelos VARX, se presenta el concepto Causalidad de Granger postulada en el año 1969. El sentido clásico de causalidad utilizado en econometría hace referencia a que *"En un experimento aleatorizado, controlado e ideal en el que los diferentes valores de X se aplicaban de forma experimental y se observaba el efecto consiguiente sobre Y"*. (Stock & Watson, 2012) En palabras simples: X causa Y bajo ciertas variables de control y con cierta confianza, lo cual no se puede aplicar a modelos VARX. En cambio, cuando X es un predictor útil de Y dadas las variables de control del modelo entonces se denomina que X causa Y en el sentido de Granger (Stock & Watson, 2012) lo cual se testea mediante una prueba específica en R Studio.

En síntesis, el sentido clásico de causalidad evidencia que existe una relación causal directa entre las variables, en cambio la causalidad de Granger se basa en la utilidad de las correlaciones entre las variables para predecir, sin establecer si existe una relación más estrecha y clara entre las variables de un modelo VARX. De no poder comprobarse que existe causalidad bajo la hipótesis de Granger, entonces los resultados indican una relación de correlación simple entre las variables.

En la sección de Anexos 10.2 que presenta en detalle en qué análisis se encontró causalidad en el sentido de Granger y en cuales simple correlación. Se destaca el hecho de que al ver si la venta online es útil para predecir las ventas offline, se encontró que solo en un 39% hay evidencia de que con 95% de confianza hay causalidad en el sentido de Granger. En cambio, en la otra dirección, solo en el 14% de los casos la venta offline causa en el sentido de

Granger la venta offline con 95% de confianza. Es decir, en ambos casos predomina una correlación simple entre las variables.

5.4 Conocimiento descubierto

En base a los insight obtenidos en los pasos previos y en base a los resultados e interpretaciones de los modelos de data mining se concluye acerca de las situaciones planteadas en los objetivos del trabajo de título.

6 Análisis de resultados

Considerando la metodología explicada, se realizan 3 tipos de análisis: ventas globales, clientes premium vs no premium y finalmente por zona geográfica. Cada uno de ellos se realizó a nivel general de las ventas, así como también para las tres principales categorías de productos: vestuario, decoración y electrónica.

6.1 Ventas globales

Aplicando la metodología descrita, primero se compararon diferentes modelos (detallados en Anexo 10.3.1) para luego elegir aquellos que según las métricas definidas representan mejor la relación entre las ventas [MM\$] en cada uno de los canales. En la Tabla 11 se distinguen la cantidad de rezagos y variables exógenas asociadas a cada modelo, así como también las estimaciones de sus coeficientes junto con su significancia. De ella se desprende que efectivamente la venta en cada canal depende, con al menos 90% de confianza, de las ventas de algunos de los meses anteriores en ambos canales.

Ventas globales								
Variable dependiente:								
	Modelo general		Modelo vestuario		Modelo electrónica		Modelo decoración	
	Online	Offline	Online	Offline	Online	Offline	Online	Offline
Online.l1	-0.09	-0.40**	0.29**	0.17	0.08	-0.01	0.06	-0.04
Offline.l1	0.04	0.20***	-0.03**	0.04	-0.04	0.11**	-0.05	0.10
Online.l2	0.05	-0.57***	0.36***	0.11	0.15**	-0.05	0.17	-0.22***
Offline.l2	0.01	0.13***	-0.02	0.09**	-0.01	0.11**	-0.03	-0.11
Online.l3	0.11*	-0.23***	0.20*	-0.77**	0.08	-0.10	0.19	-0.18**
Offline.l3	0.04**	0.11***	-0.002	0.01	0.08	0.31***	0.20	0.07
Online.l4	0.07	-0.08			0.09	-0.16**	-0.01	-0.16
Offline.l4	-0.02	0.10***			-0.04	0.18***	-0.13	0.07
Online.l5	0.36***	-0.01			0.21***	-0.07	0.21	-0.11
Offline.l5	-0.01	0.11***			0.11*	0.23***	0.27*	0.10
Online.l6	0.44**	0.40**					0.19	0.02
Offline.l6	-0.02	0.06					-0.26	-0.09
Online.l7	0.08	0.41*					0.01	-0.01
Offline.l7	0.02	0.16***					-0.08	0.24*
const	-45.36***	-102.41***	-5.00	26.84	-54.18**	-38.40	-3.30	3.45
navidad	13.82**	126.29***	1.57	87.07***	5.93***	25.97***	-0.23	4.76***
cyber	21.26***	-7.00**	6.03***	2.84	14.14***	-3.43**	4.83***	-0.26
tiendas			0.02	-1.27	0.74	0.44		
Adjusted R ²	0.95	0.99	0.75	0.98	0.93	0.95	0.87	0.94
F Statistic	33.87***	423.42***	12.22***	163.78***	33.90***	47.37***	13.11***	28.73***

Note:

*p<0.1; **p<0.05; ***p<0.01

Tabla 11: Estimación de coeficientes para cada modelo global. Fuente: Elaboración propia.

En base a la tabla anterior, se destaca que la correlación con las ventas del mismo canal en los periodos anteriores, esta es principalmente positiva. En cambio, las correlaciones estadísticamente significativas entre las ventas en uno de los canales en el tiempo t , con las del otro en los meses anteriores, las relaciones en general son negativas. Además, se interpreta que, a nivel general, las ventas del canal offline en el periodo t disminuyeron \$0.4MM por cada \$1MM de ventas en el sitio web en el mes $t-1$.

Sobre las variables exógenas, se destaca que en los meses de navidad ambos canales venden más que el resto del año, siendo las tiendas físicas las que tienen un mayor incremento, pues con 99% de confianza venden \$126MM más en Diciembre que en los otros meses. En cambio, en el sitio web se venden casi \$14MM por efecto de navidad. A nivel general, los eventos *cyber* tienen una correlación distinta con las ventas en cada canal. El sitio web vende \$21MM más por mes cuando ocurren estos eventos, de esta forma con 99% de confianza, se afirma que los dos *cyber* del año 2019 representaron un 10% de las ventas del canal online. En simultaneo, las tiendas físicas disminuyeron sus ventas en \$7MM en ambos meses que ocurrieron los eventos, sin embargo, las perdidas no logran contrarrestar las ganancias del canal online.

Acerca de la variable exógena que indica la cantidad de tiendas abiertas, ésta está presente solo en los modelos de vestuario y electrónica y en ninguno de los dos tiene un efecto estadísticamente significativo.

Para comprender con detalle las implicancias de los coeficientes estimados por el modelo VARX, a continuación, se calculan las funciones de impulso respuestas vinculadas a cada modelo. Con ellas se simula la reacción del canal offline ante una variación en las ventas en el sitio web. En la Ilustración 4 se exponen las 4 IRF's asociadas a los análisis globales con sus respectivos puntos críticos según su estabilidad. En todas las gráficas predomina una respuesta negativa tanto a corto como largo plazo.

Considerando el punto de corte de $t=14$ se calcularon las Expresiones (10) y (11) obteniendo los porcentajes de canibalización para el análisis global según las dos métricas definidas con anterioridad.

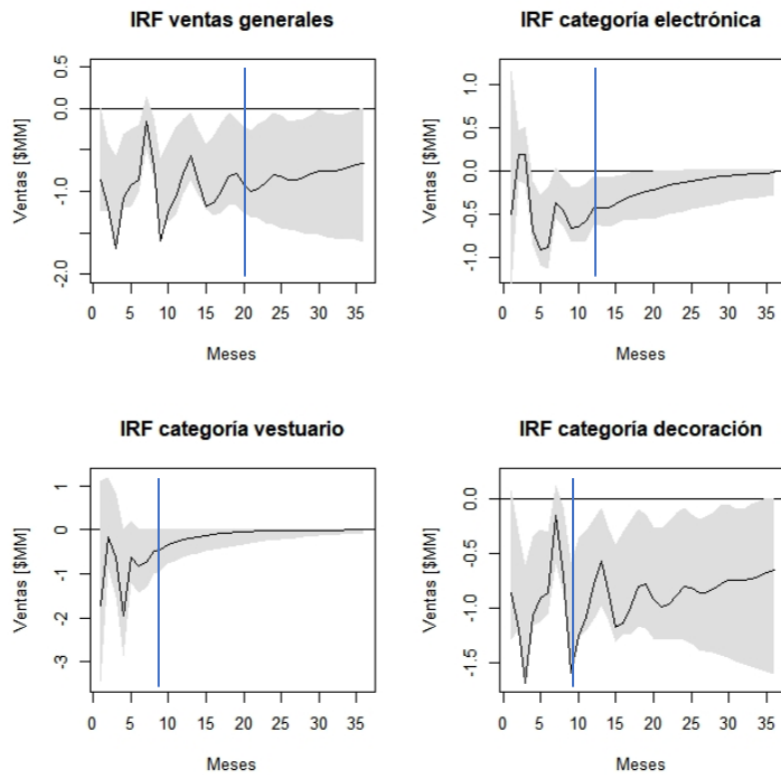


Ilustración 4: IRF's ventas globales. Fuente: elaboración propia.

El Gráfico 5 ilustra los valores promedio de ambas métricas de canibalización. Se infiere que en los últimos 3 años las tiendas físicas han dejado de vender un 2% debido al comportamiento de las ventas en el sitio web, equivalente a ~\$26MM por año. En este contexto, la categoría que más se ha dejado de transar en el medio offline ha sido decoración. Desde otro punto de vista, la canibalización entre los canales ha significado que en promedio el 6% de las ventas anuales del canal online provienen del otro canal y, por tanto, el 94% de las ganancias del sitio web son incrementales de la empresa.

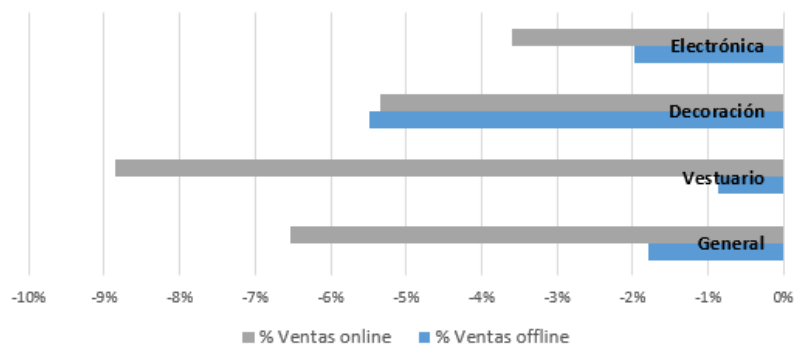


Gráfico 5: Métricas de canibalización análisis global. Fuente: Elaboración propia.

La Tabla 12 muestra la evolución de las estimaciones en los últimos 3 años. El signo negativo indica que se trata de una canibalización y el módulo la magnitud del fenómeno. Se observa como patrón que la métrica sobre las ventas offline se ha mantenido prácticamente constante, la única variación que hay es decoración que, en modulo, aumenta 1 punto porcentual el año 2019. En cambio, el módulo de la otra métrica ha tendido a la baja.

	% sobre ventas offline			% sobre ventas online		
	2017	2018	2019	2017	2018	2019
General	-2%	-2%	-2%	-8%	-6%	-5%
Vestuario	-1%	-1%	-1%	-11%	-9%	-7%
Decoración	-5%	-5%	-6%	-6%	-5%	-4%
Electrónica	-2%	-2%	-2%	-4%	-3%	-3%

Tabla 12: Evolución métricas de canibalización del análisis global. Fuente: Elaboración propia.

Finalmente, considerando el antecedente de las ventas offline general decrecieron en 2017, 2018 y 2019 en montos equivalentes a \$278.3MM, \$81.4MM y \$138.6MM respectivamente, la estimación de canibalización provocada por el canal online solo explica el 1%, 30% y 18% de ellas.

6.2 Ventas clientes premium v/s no premium

Si bien los clientes premium corresponden a ~12% de la totalidad de los clientes, ellos representan más del 50% de las ventas, por lo que es de interés para la empresa comparar la relación entre los canales cuando ellos compran versus aquellos que no pertenecen al grupo de elite. A continuación, en la Tabla 13 se caracterizan los modelos utilizados para cada estimación a nivel general de las ventas.

		Rezagos	Variables exogenas		
Premium	General	5	Navidad	Cyber	Cant. Tiendas
	Vestuario	5	Navidad	Cyber	-
	Electrónica	5	Navidad	Cyber	-
	Decoración	7	Navidad	Cyber	-
No Premium	General	5	Navidad	Cyber	-
	Vestuario	5	Navidad	Cyber	-
	Electrónica	5	Navidad	Cyber	Cant. Tiendas
	Decoración	6	Navidad	Cyber	-

Tabla 13: Composición de cada modelo. Fuente: Elaboración propia.

La Tabla 14 contiene la comparación de las estimaciones de los modelos de ventas generales según el tipo de cliente. Queda en evidencia que la venta de los clientes no premium en las tiendas físicas no esta correlacionada significativamente con las ventas en el otro canal en los periodos anteriores. En cambio, para los clientes premium se interpreta que, con 95% de confianza, se interpreta que por cada \$1MM vendido por el sitio web en el mes t-4, el canal offline ha disminuido sus ingresos en \$0.21MM en el mes t.

	<i>Variable dependiente:</i>			
	Premium		No Premium	
	Online	Offline	Online	Offline
Online.t1	0.10	0.18*	0.16**	0.12
Offline.t1	-0.03	-0.09	-0.01	-0.004
Online.t2	0.18**	0.03	0.19**	0.03
Offline.t2	-0.05	-0.02	-0.01	0.07*
Online.t3	0.12*	-0.13	0.14*	-0.16
Offline.t3	0.04	0.10	0.02	0.09*
Online.t4	0.08	-0.21**	0.07	-0.13
Offline.t4	-0.02	0.13*	-0.03	0.07
Online.t5	0.19**	-0.10	0.29***	-0.15
Offline.t5	0.08	0.16**	-0.003	0.13***
const	-40.10	60.54	-11.87***	-15.42**
navidad	4.32**	35.54***	3.22***	42.67***
cyber	16.69***	-1.93	10.06***	-1.49
tiendas	0.47	-1.39		
Adjusted R ²	0.90	0.94	0.89	0.97
F Statistic	22.78***	38.59***	23.22***	98.29***

Note:

*p<0.1; **p<0.05; ***p<0.01

Tabla 14: Comparación estimación de modelos generales según tipo de clientes. Fuente: Elaboración propia.

Por otra parte, el nivel de compra de los clientes no premium son los que más aumentan a causa de la navidad en comparación con el resto del año, las ventas a estos clientes son \$42.67MM más, en cambio para el grupo de elite el aumento es de \$35.54MM. La correlación de las ventas online con el mes de navidad también es positiva y significativa, pero se nota que es mayor en los clientes no premium y en ambos casos menor que la correlación con las tiendas físicas. En cuanto a los meses con *cyber*, la Tabla 14 indica que son los clientes premium los más sensibles a las promociones. En comparación con los otros meses, ellos gastan \$16.7MM más en el canal online, en cambio el resto de los clientes \$10.06MM. Estos eventos tienen en ambos tipos de compradores una correlación negativa con la venta en el medio offline, sin embargo, estos no son estadísticamente significativos.

Si se realiza una distinción entre los diferentes tipos de productos, se tiene la Tabla 15 la cual realiza una comparación de los coeficientes de la variable navidad y cyber para cada categoría de productos.

Tipo de cliente	Canal	Coeficientes navidad			Coeficientes cyber		
		Vestuario	Electrónica	Decoración	Vestuario	Electrónica	Decoración
Clientes premium	Online	0.83*	3.50***	0.28	3.35***	9.99***	3.28***
	Offline	22.08***	12.84***	0.95**	0.01	-2.09*	-0.15
Clientes no premium	Online	0.87**	2.18***	0.18	3.04***	4.88***	1.68***
	Offline	32.95***	9.79***	1.54***	-0.32	-1.07*	-0.05

Tabla 15: Comparación de coeficientes de variable navidad y cyber. Fuente: Elaboración propia.

En base a la tabla anterior, se interpreta que el mes de navidad tiene una correlación positiva y estadísticamente significativa con las ventas en el canal offline para ambos tipos de clientes y en todos los productos. Tanto en vestuario como en decoración la relación es más fuerte en los clientes no premium, donde en Diciembre sus compras aumentan en \$32.95MM y \$1.54MM respectivamente en comparación con el resto del año. En cambio, los clientes premium son los que más incrementan sus compras en electrónica en navidad (\$12.84MM).

Sobre las correlaciones existentes entre los eventos cyber para los diferentes tipos de productos, los datos de la Tabla 15 sugieren que la correlación con la venta online es siempre positiva y estadísticamente significativa con 99% de confianza. En cambio, con la venta en las tiendas físicas la relación es contraria y solo significativa en los productos de electrónica, donde los clientes Premium, que han sido más sensibles a las promociones por los eventos, compraron \$2.09MM menos en cada mes con cyber.

En la sección de Anexo 10.3.1 se encuentran las estimaciones completas para todas las combinaciones tipo de producto y tipo de cliente, así como también las funciones de impulso respuesta respectivas.

A continuación, la Ilustración 5 compara gráficamente los valores promedios de las dos métricas consideradas, donde los valores negativos indican que se trata de una canibalización.

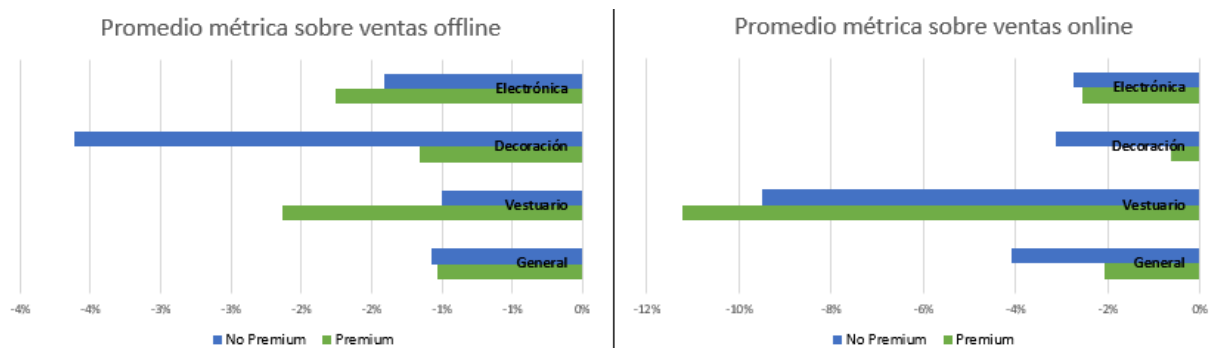


Ilustración 5: Comparación métricas premium vs no premium por categoría de producto. Fuente: Elaboración propia.

Considerando la métrica sobre las ventas offline se observa que, a nivel general, el nivel de canibalización es similar para ambos tipos de clientes. Sin embargo, al analizar con mayor profundidad las ventas agrupandolas por tipo de producto el comportamiento es diferente. Tanto en vestuario como en electrónica los premium tiene una mayor magnitud en la canibalización que en el resto de los clientes, siendo la categoría vestuario la más afectada. En cambio, en decoración son los clientes no premium poseen una magnitud del fenómeno ~4 veces mayor a los clientes de elite.

Desde otro punto de vista, se destaca que la métrica sobre las ventas online en vestuario indica que de todas las ventas del sitio web un ~11% se deben a la disminución de compras en las tiendas físicas por parte de los clientes premium y un ~9% para los clientes de elite.

En la Tabla 16, se presentan la evolución de las métricas en los últimos 3 años. En aquella que considera las ventas offline, el valor absoluto de los porcentajes son bajos (<5%), es decir, la magnitud del fenómeno canibalización es menor. Al igual que en el análisis global, la métrica offline se ha mantenido constante, las únicas variaciones han sido en decoración y electrónica en los clientes no premium. Desde otra perspectiva, la métrica sobre las ventas online mantiene el patrón de disminuir o mantenerse a través de los años.

		% sobre ventas offline			% sobre ventas online		
		2017	2018	2019	2017	2018	2019
Premium	General	-1%	-1%	-1%	-2%	-2%	-2%
	Vestuario	-2%	-2%	-2%	-14%	-11%	-9%
	Decoración	-1%	-1%	-1%	-1%	-1%	-1%
	Electrónica	-2%	-2%	-2%	-3%	-2%	-2%
No Premium	General	-1%	-1%	-1%	-5%	-4%	-3%
	Vestuario	-1%	-1%	-1%	-13%	-9%	-7%
	Decoración	-3%	-3%	-4%	-4%	-3%	-3%
	Electrónica	-2%	-1%	-1%	-3%	-3%	-2%

Tabla 16: Métricas de canibalización clientes premium vs no premium. Fuente: Elaboración propia.

6.3 Ventas por zonas geográficas

Para las estimaciones de canibalización/sinergia positiva para los clientes según la zona geográfica donde residen y haciendo distinción por las categorías de productos que compran, se utilizaron los modelos caracterizados en la Tabla 17. A diferencia de los análisis anteriores, en este contexto la mayoría de los modelos elegidos, según sus métricas asociadas, incorporan la variable exógena cantidad de tiendas disponible por zona.

		Rezagos	Variables exogenas		
Norte	General	5	Navidad	Cyber	-
	Vestuario	5	Navidad	Cyber	Cant. Tiendas
	Electrónica	5	Navidad	Cyber	-
	Decoración	5	Navidad	Cyber	Cant. Tiendas
RM Oriente	General	5	Navidad	Cyber	Cant. Tiendas
	Vestuario	5	Navidad	Cyber	Cant. Tiendas
	Electrónica	3	Navidad	Cyber	Cant. Tiendas
	Decoración	4	Navidad	Cyber	-
RM Poniente	General	5	Navidad	Cyber	Cant. Tiendas
	Vestuario	5	Navidad	Cyber	Cant. Tiendas
	Electrónica	5	Navidad	Cyber	-
	Decoración	5	Navidad	Cyber	Cant. Tiendas
Sur	General	6	Navidad	Cyber	-
	Vestuario	5	Navidad	Cyber	-
	Electrónica	3	Navidad	Cyber	Cant. Tiendas
	Decoración	5	Navidad	Cyber	-

Tabla 17: Caracterización modelos análisis por zona. Fuente: Elaboración propia.

Las estimaciones de los modelos VARX a nivel de ventas generales en las diferentes zona son expuestas en la Tabla 18.

Ventas generales: comparación zonas

	<i>Variable dependiente:</i>							
	Norte		RM Oriente		RM Poniente		Sur	
	Online	Offline	Online	Offline	Online	Offline	Online	Offline
Online.11	0.14**	0.10	0.19*	0.11	0.12	0.04	-0.07	0.25
Offline.11	-0.01	-0.09**	-0.05	0.004	-0.04	-0.05	0.06**	-0.02
Online.12	0.20***	0.02	0.22**	-0.02	0.18**	-0.11	0.08	0.09
Offline.12	-0.01	0.02	-0.11	-0.10	-0.05	-0.01	0.05*	0.09
Online.13	0.13*	-0.08	0.17*	-0.17**	0.12*	-0.21**	0.08*	-0.15
Offline.13	0.04	0.07	0.07	0.10	0.01	0.04	0.08***	0.17**
Online.14	0.08	-0.17	0.06	-0.19**	0.10	-0.24***	0.03	-0.16
Offline.14	-0.01	0.09*	-0.02	0.07	-0.03	0.07	-0.04	0.16**
Online.15	0.28***	-0.19	0.17	-0.07	0.20**	-0.20*	0.38***	-0.19
Offline.15	0.02	0.14***	-0.0003	0.12*	0.03	0.14**	-0.03	0.24***
Online.16							0.44***	-0.32
Offline.16							-0.04	0.03
const	-6.29***	-3.05	-4.06	5.00	-13.58	13.93	-10.51***	-10.65
navidad	1.60***	16.77***	0.92**	6.08***	3.31***	31.31***	3.33***	19.14***
cyber	4.45***	-0.97	3.73***	-0.01	11.69***	-2.22	5.18***	-0.34
tiendas			0.03	-0.35*	0.19	-0.99		
Adjusted R ²	0.91	0.97	0.85	0.94	0.89	0.97	0.95	0.95
F Statistic	29.07***	88.29***	14.78***	36.48***	21.72***	74.66***	47.86***	40.30***

Note:

*p<0.1; **p<0.05; ***p<0.01

Tabla 18: Coeficientes estimaciones por zona en ventas generales. Fuente: elaboración propia.

En la zona norte y en la sur las ventas del canal offline no dependen de forma significativa de las transacciones de los meses anteriores en el sitio web. Lo análogo ocurre con las ventas en el canal online en la zona norte. Mientras tanto en RM Oriente y Poniente los coeficientes significativos indican una relación principalmente negativa.

La variable exogena que indica la cantidad de tiendas en cada zona solo tiene una correlación significativa con las ventas en el canal offline en la zona oriente. La estimación sugiere que cuando se ha abierto una nueva tienda, el canal offline, con 90% de confianza, ha disminuido sus ingresos en \$0.35MM en el mes de la inauguración. Si bien lo anterior suena contraintuitivo se plantean la siguiente hipótesis: los resultados pueden incluir el hecho de que la venta offline en dicha zona a estado disminuyendo en los últimos años aunque se hayan abiertos nuevas tiendas, sumado al hecho de que un nuevo punto de venta físico puede generar una canibalización dentro del mismo canal.

A nivel general de las ventas, según las estimaciones de la Tabla 19 en los meses de navidad las correlaciones son positivas y estadísticamente significativas ($\alpha=0.05$) para ambos canales a lo largo de Chile. En las tiendas

físicas, la zona poniente es la que aumenta más su consumo en Diciembre (\$31.31MM) con respecto a los otros meses, seguido por la zona sur (\$19.14MM) y norte (\$16.77MM). El orden de magnitud de la correlación es menor en oriente (\$6.08MM), prácticamente un 80% menos que el incremento en Poniente en el mes de navidad. En cuanto a la variable exógena, se asegura que con 99% de confianza los meses con *cyber* aumentan sus ventas online de forma significativa, el mayor incremento es en RM Poniente con \$11.69MM. Por otra parte, las correlaciones con el canal offline si bien en general son negativas, no son significativas.

A continuación, en la Tabla 19 se detallan los coeficientes de las variables navidad y cyber para las diferentes combinaciones zona-productos. Las estimaciones completas para los modelos se encuentran en la sección de Anexos 10.3.3.

Zona	Canal	Coeficientes navidad			Coeficientes cyber		
		Vestuario	Electrónica	Decoración	Vestuario	Electrónica	Decoración
Norte	Online	0.29*	4.91***	-0.01	1.04***	\$-\$0.61**	1.04***
	Offline	11.57***	1.30***	0.37***	\$-\$0.39	2.38***	-0.01
RM Oriente	Online	0.30***	0.37	0.16	0.79***	1.80***	0.81***
	Offline	4.46***	1.37***	0.50***	0.40*	\$-\$0.06	0.03
RM Poniente	Online	0.77*	2.48***	-0.18	2.76***	6.96***	1.83***
	Offline	21.20***	9.38***	1.06***	-0.44	\$-\$1.56*	-0.12
Sur	Online	0.35	0.78	-0.1	1.54***	3.09***	1.47***
	Offline	13.56***	6.35***	0.79***	-0.29	\$-\$0.15	-0.1

Tabla 19: Coeficientes variable navidad, modelos por zona y categoría de productos. Fuente: Elaboración propia.

A partir de la tabla anterior se afirma que, con 99% de confianza, la categoría de productos que más aumenta su venta offline por navidad es vestuario seguida por electrónica y luego decoración. En cambio, los clientes por la página web aumentan más sus compras en electrónica que en los otros productos, presentando correlaciones estadísticamente significativas en la zona norte y RM Poniente.

En base a lo coeficientes presentados en la Tabla 19, se infiere que, con 90% de confianza, los eventos cyber mantienen una correlación positiva y significativa con la venta en la zona norte en productos de electrónica y en Oriente-vestuario, siendo la primera de mayor magnitud que la segunda. Se interpreta que las tiendas físicas aumentan su venta en \$2.38MM en los meses de cyber para la categoría electrónica, lo cual es suficiente para contrarrestar la

correlación negativa con norte-electrónica pero en el sitio web. Por otra parte, la misma categoría pero en RM Poniente se vende \$1.56MM menos en el canal offline durante los meses con dichos eventos.

En cuanto a la relación del canal online en los meses de *cyber*, se observa el mismo patrón que en los análisis anteriores. A excepción del caso norte-electrónica mencionado con anterioridad, en todas las zonas y todos los productos existe un aumento en las ventas durante los meses con cyber. El caso más destacable, es en RM Poniente donde las compras en electrónica aumentan en casi \$7MM, considerando que ocurren dos eventos al año, el incremento equivale al 16% de la venta del canal en 2019.

La Ilustración 6, contiene las gráficas de las dos métricas de canibalización para todas las combinaciones de zonas y tipo de productos, estimadas como promedio de los últimos 3 años.

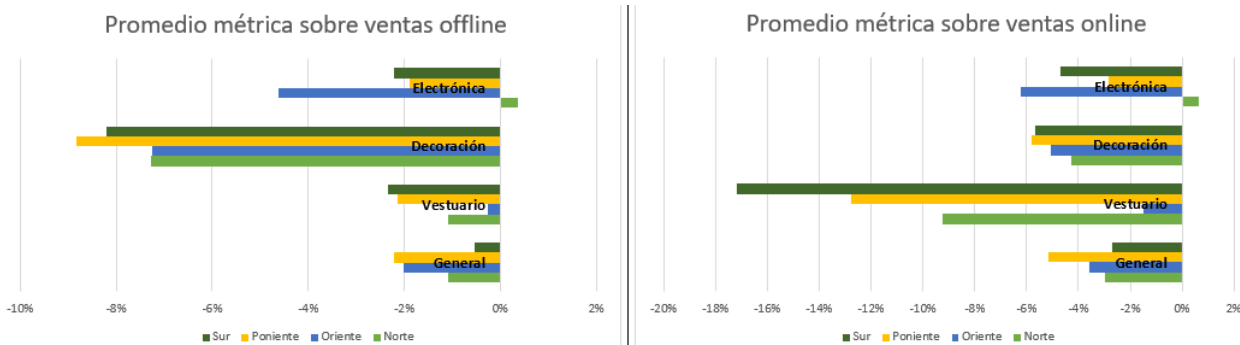


Ilustración 6: Comparación métricas por zona para cada categoría de producto. Fuente: Elaboración propia.

A nivel general las métricas indican que la zona poniente es la que posee mayor canibalización, seguida por oriente, norte y finalmente sur. Al realizar un análisis por separado para cada tipo de productos las observaciones son diferentes. Considerando a la zona norte como punto de inicio, primero, se destaca lo que ocurre en la zona norte en las compras en electrónica, donde a diferencia de todos los casos estudiados, la relación de las ventas indica una leve sinergia positiva que en ambas métricas no supera un 1%. Sobre los otros tipos de productos en dicha zona, la canibalización, con respecto a las ventas online, es mayor en vestuario con casi 10%, seguida por decoración que presenta el menor valor de todas las zonas.

Según la métrica offline, para todas las zonas, aunque en diferente magnitud, la categoría que más se ha dejado de vender es decoración. Sin embargo, lo que ocurre con los otros productos es diferente en cada zona. En oriente electrónica se ha dejado más de vender offline que vestuario que no alcanza a superar el 1% de canibalización. En cambio, la zona sur presenta prácticamente la misma magnitud del fenómeno en vestuario y electrónica.

Desde otro punto de vista, en la métrica que considera las ventas online, la zona sur destaca con su comportamiento en vestuario, donde las ventas "migradas" desde las tiendas físicas equivalen a ~18% manteniendo una brecha de 16 puntos porcentuales con el nivel de canibalización estimado para oriente.

A continuación, en la Tabla 20 se presenta la evolución de ambas métricas a través de los últimos 3 años. Al igual que en los casos anteriores, el signo negativo indica canibalización y los positivos (norte-electrónica) referencian una sinergia, siendo el módulo de la métrica la magnitud de los efectos encontrados.

		% sobre ventas offline			% sobre ventas online		
		2017	2018	2019	2017	2018	2019
Norte	General	-1%	-1%	-1%	-4%	-3%	-3%
	Vestuario	-1%	-1%	-1%	-12%	-9%	-7%
	Decoración	-7%	-7%	-9%	-5%	-4%	-3%
	Electrónica	0%	0%	0%	1%	1%	1%
RM Oriente	General	-2%	-2%	-2%	-4%	-4%	-3%
	Vestuario	0%	0%	0%	-2%	-1%	-1%
	Decoración	-6%	-7%	-9%	-5%	-5%	-5%
	Electrónica	-4%	-4%	-5%	-5%	-4%	-4%
RM Poniente	General	-2%	-2%	-3%	-6%	-5%	-5%
	Vestuario	-2%	-2%	-2%	-16%	-12%	-10%
	Decoración	-8%	-8%	-10%	-7%	-5%	-5%
	Electrónica	-2%	-2%	-2%	-3%	-3%	-3%
Sur	General	-1%	-1%	-1%	-2%	-2%	-1%
	Vestuario	-2%	-2%	-2%	-24%	-18%	-13%
	Decoración	-7%	-8%	-10%	-7%	-5%	-5%
	Electrónica	-2%	-2%	-2%	-6%	-4%	-4%

Tabla 20: Métricas de canibalización del análisis por zona. Fuente: Elaboración propia.

De la tabla anterior se destaca que a través de los años la relación entre los canales ha sido aproximadamente nula para las ventas de electrónica en norte y vestuario en RM oriente, es decir, en estos casos la creciente venta del sitio

web no ha influido en la disminución de las compras en las tiendas físicas. También se subraya lo ocurrido en la zona sur en 2017 donde, de todas las ventas del canal online en vestuario, casi un cuarto de ellas (24%) se relacionan con la disminución de ingresos ligados a dicha categoría en las tiendas físicas.

Por otra parte, se mantienen los patrones encontrados para ambas métricas; la canibalización medida con respecto a las ventas del canal offline tienden a mantenerse (63% de los casos) o aumenta, en módulo, entre el 2017 y el 2019. En cambio, la otra métrica marca la tendencia a disminuir (75% de los casos) o mantenerse en dicho periodo de tiempo.

6.4 Sensibilidad de los resultados

Los resultados dependen de la estipulación del plazo $t=14$ en el horizonte de tiempo en la valuación de las IRF's, por lo que se analizó cómo varían los resultados al considerar diferentes plazos.

Al establecer $t=6$ tanto la métrica offline como la online varían en promedio 0%, alcanzando peaks en Zona RM Oriente venta electrónica (+2 pp) y RM Oriente en categoría vestuario en 2017 (-3 pp) respectivamente.

Cuando se considera $t=18$ las variaciones en promedio también son de 0% en ambas métricas, aquella que considera las ventas offline alcanza su mayor cambio en la Zona RM Poniente específicamente en la venta en decoración en 2017 (+4 pp). En cambio, en la métrica sobre la venta online en ningún caso las variaciones superan 1 punto porcentual.

A continuación, en la Tabla 23 se presentan en detalle cómo varían los resultados para cada métrica en cada caso analizado al cambiar el horizonte de tiempo de $t=14$ a $t=6$ o 18 .

		Métrica % sobre ventas offline						Métrica % sobre ventas offline					
		Dif 14 - 6 meses			Dif 18 -14 meses			Dif 14 - 6 meses			Dif 18 -14 meses		
		2017	2018	2019	2017	2018	2019	2017	2018	2019	2017	2018	2019
Global	General	0%	0%	0%	0%	0%	0%	-1%	-1%	0%	0%	0%	0%
	Vestuario	0%	0%	0%	0%	0%	0%	0%	0%	0%	-1%	-1%	0%
	Decoración	0%	0%	-1%	2%	1%	0%	-1%	0%	0%	0%	0%	0%
	Electrónica	-1%	-1%	-1%	1%	0%	0%	-2%	-1%	-1%	0%	0%	0%
Premium	General	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
	Vestuario	0%	0%	0%	1%	1%	0%	-3%	-2%	-1%	1%	1%	1%
	Decoración	1%	1%	1%	0%	0%	0%	1%	1%	1%	0%	0%	0%
	Electrónica	-1%	-1%	-1%	1%	0%	0%	-1%	-1%	-1%	0%	0%	0%
No Premium	General	0%	0%	0%	0%	0%	0%	2%	1%	1%	0%	0%	0%
	Vestuario	0%	0%	0%	0%	0%	0%	1%	1%	1%	0%	0%	0%
	Decoración	1%	1%	1%	1%	0%	0%	1%	1%	0%	0%	0%	0%
	Electrónica	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Norte	General	0%	0%	0%	0%	0%	0%	2%	1%	1%	0%	0%	0%
	Vestuario	0%	0%	0%	0%	0%	0%	1%	0%	0%	0%	0%	0%
	Decoración	0%	0%	1%	2%	1%	-1%	0%	0%	0%	0%	0%	0%
	Electrónica	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Oriente	General	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
	Vestuario	0%	0%	0%	0%	0%	0%	-3%	-2%	-2%	1%	0%	0%
	Decoración	-1%	-1%	-1%	2%	1%	0%	-1%	-1%	-1%	0%	0%	0%
	Electrónica	2%	2%	2%	0%	0%	-1%	2%	2%	2%	-1%	-1%	-1%
Poniente	General	0%	0%	0%	0%	0%	0%	1%	0%	0%	0%	0%	0%
	Vestuario	0%	0%	0%	0%	0%	0%	-1%	-1%	-1%	0%	0%	0%
	Decoración	-1%	0%	-1%	4%	2%	0%	0%	0%	0%	0%	0%	0%
	Electrónica	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Sur	General	0%	0%	0%	0%	0%	0%	-2%	-1%	-1%	1%	1%	0%
	Vestuario	0%	0%	0%	0%	0%	0%	-1%	-1%	-1%	1%	1%	0%
	Decoración	0%	0%	-1%	3%	2%	0%	0%	0%	0%	0%	0%	0%
	Electrónica	0%	0%	0%	0%	0%	0%	0%	0%	0%	-1%	-1%	0%
	Promedio	0%	0%	0%	1%	0%	0%	0%	0%	0%	0%	0%	0%

Tabla 21: Diferencia de estimaciones al cambiar el horizonte de tiempo. Fuente: Elaboración propia.

7 Conclusiones

En base a los antecedentes y estudios anteriores acerca del fenómeno de canibalización entre canales, se desprende que es necesario un análisis econométrico de los datos, pues hay ocasiones donde un nuevo canal de venta provoca una sinergia positiva. Con el fin de determinar la real relación entre los canales de la empresa de retail se modela la situación mediante Vectores autoregresivos para series de tiempo, metodología replicable a diferentes escenarios, obteniendo estimaciones para dinámicas complejas. Además, ha sido vinculada a funciones de impulso respuesta para un análisis más específico que incluya el efecto de aumentar la venta en el canal online sobre los ingresos en el offline.

Los modelos utilizados tienen en promedio 5 rezagos y las estimaciones, en general, indican que efectivamente hay relaciones estadísticamente significativas entre las ventas del tiempo t de cada canal con las realizadas en los meses anteriores en ambos canales, en donde predomina una naturaleza de correlación simple por sobre causalidad en el sentido de Granger. En cuanto a la interpretación de los modelos, estos señalan que las correlaciones de ventas en tienda y sitio web son principalmente negativas. Por otra parte, se destaca que en el 100% de los casos analizados las estimaciones de los modelos mejoran, según las métricas consideradas, al incorporar variables exógenas de control. Todos los modelos seleccionados consideran las variables cyber y navidad, en cambio, solo un 46% de los casos los modelos mejoran las métricas asociadas al incluir aquella que indica la cantidad de tiendas abiertas.

Las variables exógenas navidad y cyber presentaron significancia individual en prácticamente todos los modelos. La particularidad del mes de Diciembre es que mantiene una correlación positiva con la venta de ambos canales, siendo de mayor magnitud con las tiendas físicas y específicamente con la venta en vestuario. Por otra parte, son los clientes no premium aquellos que incrementan más sus compras por navidad, y desde una perspectiva por zonas es RM Poniente. Con respecto a la variable cyber, las estimaciones sugieren que existe una relación estadísticamente significativa con ambos canales, pero, en general con signos contrarios. A nivel global de la empresa, en los meses con *cyber* las tiendas disminuyen sus ventas en \$7MM de los cuales ~50% corresponde a productos de electrónica. Por otra parte, los clientes más sensibles a aumentar su consumo durante los eventos promocionales online

son los premium, y desde una perspectiva de zonas geográfica, los que residen en RM Poniente son los que compran más.

Las estimaciones de canibalización/sinergia, según las definiciones planteadas, indican que en un 96% de los casos analizados existe evidencia de que el canal online ha hecho que las ventas en las tiendas físicas disminuyan. La gran excepción es lo que ocurre en la zona norte con la venta de productos electrónicos, donde ambos canales mantienen una sinergia positiva leve (<1%). Importante señalar que el objetivo de las métricas planteadas como canibalización/sinergia es representar una modelación radical del fenómeno, buscando ilustrar cuanto más habría vendido la tienda si no existiera el canal online, monto que luego es normalizado para su comparación entre categorías de productos y años.

A nivel global de las ventas se tiene que el canal offline ha disminuido sus ventas en ~\$166MM anualmente en los últimos 3 años, de los cuales solo un ~16% es explicado por el aumento en las ventas del sitio web. Se concluye que deben existir otras razones por las cuales las tiendas físicas están vendiendo menos, entre las cuales se encuentra que la población esté comprando menos o que estén siendo fidelizados por la competencia. Así, los resultados sugieren la evaluación de un cierre de algunas de las tiendas a lo largo de Chile.

Por otra parte, los modelos indican que, al estimar la canibalización sobre la venta del canal offline, es decir, evaluar qué porcentaje de la venta realizada en las tiendas se ha perdido debido a la incorporación del canal online, el fenómeno sigue un claro patrón. El 64% de los casos mantiene el mismo nivel de canibalización/sinergia entre los años 2017 y 2019, en cambio el resto tiende al aumento en máximo 2 puntos porcentuales. A nivel general de ventas, se tiene que las zonas más afectadas son las de la Región Metropolitana. Mientras tanto en la distinción por tipo de cliente según su comportamiento de compra, ambos poseen prácticamente la misma canibalización. En específico, los clientes de elite han dejado de comprar vestuario en tienda debido a una preferencia por hacerlo por en canal online. Mientras que para todos los otros análisis realizados la categoría que más se ha dejado de vender offline es decoración, lo cual no solo se debe a que los clientes prefieran el sitio web, sino que también a que la empresa ha decidido ir disminuyendo el stock disponible en tienda, forzando en cierta medida la transferencia de ventas de un canal a otro.

Si ahora se considera la métrica que indica el porcentaje de las ventas online que son atribuibles al fenómeno de canibalización, se observa que en general la estimación disminuye a través de los años. En este contexto, en general (96% de los casos) la categoría con mayor canibalización es vestuario siendo la excepción la venta de decoración en la zona Oriente.

Sin lugar a duda, un análisis global de canibalización entrega parámetros que permiten entender la dinámica entre los canales, sin embargo, esconde diferentes efectos, pues los resultados indican que cada categoría de productos, tipo de clientes y zona geográfica presentan el fenómeno en diferente magnitud.

Importante destacar que los resultados propuestos se basan en un horizonte $t=14$ meses para la evaluación de las IRF's. Sin embargo, dada la construcción de la metodología donde la respuesta del canal offline ante las ventas del medio online es normalizada según la cantidad de meses considerados (Expresión 10 en sección 5.3), los resultados son robustos y las variaciones mínimas al cambiar el horizonte de tiempo a $t=6$ o $t=18$.

Otro aspecto a recalcar es que en el análisis preliminar de la relación entre canales (sección 3.4) se planteó la hipótesis de que predominan los casos de canibalización lo cual fue corroborado con los modelos VARX planteados. Sin embargo, se ratifica el hecho de que las dinámicas entre canales son complejas y que no basta con un análisis de evolución en las proporciones de ventas en los medios online y offline, pues el orden de magnitud encontrado en la sección preliminar no corresponde a lo estimado por la metodología propuesta. Además, mediante los modelos utilizados se hallaron también insight sobre la relación de canales en fechas de peak de ventas (navidad y cyber).

8 Trabajo futuro propuesto

A raíz de los hallazgos del trabajo de título se distinguen tres ejes de investigaciones futuras sugeridas. La primera de ellas es replicar la metodología planteada con modelos Vectoriales autorregresivos para series de tiempo para ahondar en el análisis realizado para el periodo 2017-2019. Así como las ventas globales encubren diferentes niveles de canibalización para cada una de las 3 principales categorías de productos, se plantea la hipótesis que estos a su vez también ocultan diferentes dinámicas para conjuntos más específicos de artículos. El realizar este análisis por zona geográfica, o más detallado por comuna, permitirá tomar decisiones sobre qué productos exponer en las tiendas.

Sin lugar a duda la pandemia provocada por el Covid-19 ha cambiado el comportamiento de compra de los clientes, sobre todo durante las cuarentenas. Sin embargo, queda la interrogante de que ocurrirá cuando se reactiven la mayoría de las tiendas físicas. ¿Las personas seguirán comprando online?, en comparación con el 2019 ¿Qué tipo de productos se venderán menos en tiendas? Con esto en mente, el segundo eje plantea que la metodología con modelos VAR se puede replicar para estimar la canibalización del canal online al offline en el año 2020. Dadas las condiciones sanitarias debido al covid 19, se plantea incorporar solo los datos de dicho año a nivel de ventas semanales entre Enero y Diciembre, utilizando un 80% para calibración y el resto para validación. Se propone mantener la variable exógena de navidad y cyber (en caso de que se realice alguno en 2020), además se tiene la hipótesis de que durante el periodo de análisis la variable que indica la cantidad de tiendas abiertas tendrá una mayor relevancia ya que, por ejemplo, durante la primera quincena de Agosto en la Región Metropolitana solo hubo una tienda abierta, en cambio en el resto del país más de 10. Igualmente se recomienda utilizar datos sobre la evolución de la pandemia en Chile como variable exógena, por ejemplo, casos activos y/o porción del territorio con cuarentena. Se sugiere que el análisis se realice por zona geográfica como en el trabajo de título o incluso con particiones del territorio más específicas (regiones, comunas, etc). Según los resultados que se encuentren se recomienda evaluar el cierre de tiendas a lo largo de Chile.

Por otra parte, el 2020 también significó cambios dentro del canal online, hasta el 2019 la aplicación de la empresa representaba un porcentaje de las ventas despreciable, en cambio, durante el nuevo año se fomentó la utilización de la

app mediante mayor publicidad en los medios de comunicación, así como también con promociones exclusivas de despacho gratis sin mínimo de compra. Esta innovación en el canal digital abre la posibilidad de analizar como la aplicación ha canibalizado las ventas del sitio web.

El tercer eje de estudios futuros recomendados es complementario a la estimación de canibalización mediante modelos VAR. Similar a lo que realizó Kollmann, Kuchertz y Kayser en 2012 en su publicación "*Cannibalization or synergy? Consumers' channel selection in online-offline multichannel systems*" (Kollmann et al., 2012), se propone estudiar qué factores influyen en la decisión del canal de compra. Se realizó un experimento de campo para determinar el impacto de los descuentos en el mismo canal y en el otro. Ellos concluyeron que el deseo de servicio en lugar de la aversión al riesgo contribuye a que los clientes prefieran el canal offline, lo cual puede ser una de las hipótesis a validar. En este caso se sugiere realizarlo para comprender el nuevo comportamiento de los clientes luego del paso de la pandemia, incluyendo análisis más específicos por tipo de productos y/o según zonas geográficas.

9 Bibliografía

A. Ansari, C.F. Mela, S.A. Neslin, Customer channel migration, *Journal of marketing research* 45, 2008, pp. 60–76.

B. Deleersnyder, I. Geyskens, K. Gielens, M.G. Dekimpe, How cannibalistic is the Internet channel? A study of the newspaper industry in the United Kingdom and the Netherlands, *International Journal of Research in Marketing* 19, 2002, pp. 337–348.

C. Hanck, M. Arnold, A. Gerber, M. Schmelzer, 2019, *Econometrics with R*.

D. Herhausen, J. Binder, M. Schoegel, A. Herrmann, Integrating bricks with clicks: retailer level and channel level outcomes of online offline channel integration, *Journal of Retailing* 91, 2015, pp. 309-325.

J. Avery, T.J. Steenburgh, J. Deighton, M. Caravella, Adding bricks to clicks: predicting the patterns of cross-channel elasticities over time, *Journal of Marketing* 76, 2012, pp. 96–111.

J.E.M. Van Nierop, P.S. Leeflang, M.L. Teerling, K. Huizingh, The impact of the introduction and use of an informational website on offline customer buying behavior, *International Journal of Research in Marketing* 28, 2011, pp. 155–165.

J. Wooldridge, 2010, *Introducción a la econometría: un enfoque moderno*, 4ª edición, México, Cengage Learning.

L. Huang, X. Lu, S. Ba, An empirical study of the cross-channel effects between web and mobile shopping channels. *Information & Management* 53, 2015, pp. 265 – 278.

M. Dekimpe, D. Hanssens, Persistence Modeling for Assessing Marketing Strategy Performance, 2003.

M. Goic, R. Álvarez, R. Montoya, The Effect of House Ads on Multichannel Sales, Journal of Interactive Marketing 42, 2018, pp. 23-45.

M.M Montoya-Weiss, G. Voss, D. Grewal, Determinants of Online Channel Use and Overall Satisfaction With a Relational, Multichannel Service Provider, Journal of The Academy of Marketing Science 31, pp. 448-458.

O. Maimon, L. Rokach, 2010, Data Mining and Knowledge Discovery Handbook, Boston, Springer, pp. 1-15.

S.A. Neslin, D. Grewal, R. Leghorn, V. Shankar, M.L. Teerling, J.S. Thomas, et al, Challenges and Opportunities in Multichannel Customer Management, Journal of Service Research 9. 2006, pp. 95-112.

R. Suárez, Nelson (2012). La revolución en la toma de decisiones estadísticas: el p-valor. Telos, 14(3),439-446. [Fecha de Consulta: Enero de 2020]. Disponible en: <https://www.redalyc.org/articulo.oa?id=993/99324907004>

R. Tsay, 2014, Multivariate time series analysis with R and Financial Application, Chicago, Wiley.

T. Chai, R.R. Draxler, Root mean square error (RMSE) or mean absolute error (MAE)? Arguments against avoiding RMSE in the literatura, Geoscientific Model Development 7, 2014, pp.1247-1250.

T. Kollmann, A. Kuckertz, I. Kayser, Cannibalization or synergy? Consumers' channel selection in online-offline multichannel systems, Journal of Retailing and Consumer Services 19, 2009, pp. 186-194.

10 Anexos

10.1 Puntos críticos IRF

En el Gráfico 6 se encuentran los detalles de los puntos críticos para cada modelo y cuyo promedio utilizado para los cálculos de canibalización es $t=14$.

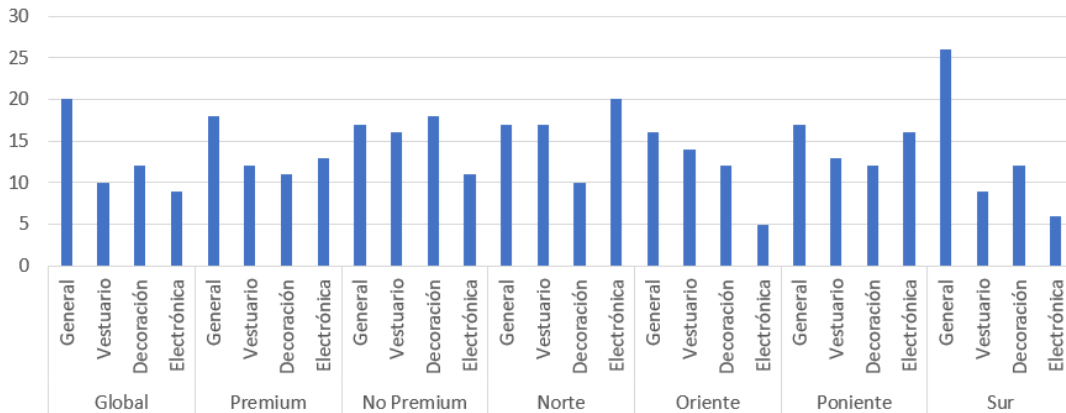


Gráfico 6: Puntos de corte IRF's de los modelos. Fuente: Elaboración propia.

10.2 Causalidad de Granger

En la Tabla 22, se encuentra en detalle en que análisis se encontró evidencia de causalidad en el sentido de Granger con 95% de confianza.

		¿Venta online -¿ Venta offline?		¿Venta online -¿ Venta offline?	
		Correlación	Causalidad de Granger	Correlación	Causalidad de Granger
Global	General		✓	✓	
	Vestuario	✓		✓	
	Electrónica	✓		✓	
	Decoración		✓	✓	
Premium	General	✓		✓	
	Vestuario		✓	✓	
	Electrónica	✓		✓	
	Decoración	✓		✓	
No premium	General	✓		✓	
	Vestuario	✓		✓	
	Electrónica	✓		✓	
	Decoración		✓		✓
Zona norte	General	✓		✓	
	Vestuario	✓		✓	
	Electrónica	✓		✓	
	Decoración		✓	✓	
Zona RM Oriente	General		✓	✓	
	Vestuario		✓	✓	
	Electrónica	✓			✓
	Decoración		✓		✓
Zona RM Poniente	General		✓	✓	
	Vestuario		✓	✓	
	Electrónica	✓		✓	
	Decoración		✓	✓	
Zona sur	General	✓			✓
	Vestuario	✓		✓	
	Electrónica	✓		✓	
	Decoración	✓		✓	

Tabla 22: Identificación de casos donde hay causalidad en el sentido de Granger. Fuente: Elaboración propia.

10.3 Elecciones de modelos y funciones de impulso respuesta

Se elaboran modelos VAR y VARX que difieren entre sí en la cantidad de rezagos, así como también en las variables exógenas que contemplan las variables binarias Navidad y Cyber, además de una numérica que indica la cantidad de tiendas abiertas en promedio en el periodo. Se selecciona el modelo que sobresalga en las métricas.

10.3.1 Ventas globales

La Tabla 23 muestra los modelos considerados para las ventas globales generales, de los cuales se elige el número 4.

		Modelo 1	Modelo 2	Modelo 3	Modelo 4	Modelo 5
Modelo	Variable exógenas cyber	Si	Si	Si	Si	No
	Variable exógena navidad	Si	Si	Si	Si	No
	Variable exógena # tiendas	Si	Si	No	No	No
	Cantidad rezagos	7	9	5	7	5
Supuestos	Raíces unitarias	Si	Si	Si	Si	Si
	Autocorrelacion residuos	No	Si	Si	No	Si
	Normalidad residuos	Si	Si	Si	Si	No
	Homocedasticidad residuos	Si	Si	Si	Si	Si
Data: train	LL	-545	-487	-620	-545	-725
	AIC	1,162	1,063	1,291	1,158	1,495
	BIC	1,213	1,123	1,330	1,207	1,528
	RMSE offline	1,583	896	3,288	1,590	28,452
	RMSE online	1,708	1,550	2,575	1,708	7,248
	MAE offline	1,133	742	2,542	1,132	17,707
	MAE online	1,397	1,315	1,827	1,408	5,238
Data: test	RMSE offline	9,808	9,958	10,559	9,838	36,837
	RMSE online	8,607	8,669	6,619	8,078	18,349
	MAE offline	7,379	7,856	8,136	7,606	29,260
	MAE online	5,743	6,326	4,296	5,260	13,328

Tabla 23: Comparación modelos ventas globales generales. Fuente: Elaboración propia.

En la Tabla 24 se resume la comparación entre los diferentes modelos para las ventas globales en vestuario. Se elige el Modelo 1 debido que cumple con todos los supuestos de un modelo VAR y sobresale en la mayoría de las métricas.

		Modelo 1	Modelo 2	Modelo 3	Modelo 4	Modelo 5
Modelo	Variable exógenas cyber	Si	Si	Si	Si	No
	Variable exógena navidad	Si	Si	Si	Si	No
	Variable exógena # tiendas	Si	Si	No	No	No
	Cantidad rezagos	3	8	2	8	7
Supuestos	Raíces unitarias	Si	No	Si	No	No
	Autocorrelacion residuos	No	No	No	No	No
	Normalidad residuos	Si	Si	Si	Si	Si
	Homocedasticidad residuos	Si	Si	Si	Si	Si
Data: train	LL	-630	-485	-657	-485	-596
	AIC	1,299	1,015	1,342	1,045	1,252
	BIC	1,330	1,071	1,364	1,099	1,295
	RMSE offline	3,171	697	3,980	1,175	11,026
	RMSE online	1,216	507	1,281	518	1,392
	MAE offline	2,617	540	3,178	974	7,497
	MAE online	816	425	874	437	1,095
Data: test	RMSE offline	11,293	14,574	10,912	7,590	49,200
	RMSE online	2,774	3,084	3,686	2,875	3,158
	MAE offline	7,154	9,934	7,713	5,624	32,874
	MAE online	1,953	2,773	2,694	2,194	2,618

Tabla 24: Modelos VAR y VARX de ventas globales: vestuario. Fuente: elaboración propia

Para el caso de las ventas globales en electrónica, Tabal 25, se ha elegido el Modelo 1.

		Modelo 1	Modelo 2	Modelo 3	Modelo 4	Modelo 5
Modelo	Variable exógenas cyber	Si	Si	Si	Si	No
	Variable exógena navidad	Si	Si	Si	Si	No
	Variable exógena # tiendas	Si	Si	No	No	No
	Cantidad rezagos	5	6	5	4	5
Supuestos	Raíces unitarias	Si	Si	Si	Si	Si
	Autocorrelacion residuos	No	Si	No	No	No
	Normalidad residuos	Si	Si	Si	Si	No
	Homocedasticidad residuos	Si	Si	Si	Si	Si
Data: train	LL	-565	-536	-567	-615	-654
	AIC	1,186	1,134	1,187	1,373	1,354
	BIC	1,227	1,184	1,226	1,307	1,387
	RMSE offline	1,338	1,208	1,362	1,818	7,159
	RMSE online	1,186	941	1,260	2,328	5,487
	MAE offline	1,114	1,027	1,146	1,492	4,821
	MAE online	927	767	971	1,463	3,985
Data: test	RMSE offline	2,732	3,056	2,967	3,477	8,082
	RMSE online	4,210	3,620	3,700	4,874	5,891
	MAE offline	2,380	2,803	2,682	2,835	6,789
	MAE online	2,981	1,971	2,335	3,069	4,097

Tabla 25: Modelos VAR y VARX para categoría Electrónica. Fuente: Elaboración propia

Finalmente, para la venta en decoración a nivel global, se modela con la opción 4 de la Tabla 26.

		Modelo 1	Modelo 2	Modelo 3	Modelo 4	Modelo 5
Modelo	Variable exógenas cyber	Si	Si	Si	Si	No
	Variable exógena navidad	Si	Si	Si	Si	No
	Variable exógena # tiendas	Si	Si	No	No	No
	Cantidad rezagos	5	6	6	7	7
Supuestos	Raíces unitarias	Si	Si	Si	Si	Si
	Autocorrelacion residuos	No	No	No	No	No
	Normalidad residuos	Si	Si	Si	Si	No
	Homocedasticidad residuos	Si	Si	Si	Si	Si
Data: train	LL	-504	-480	-483	-462	-506
	AIC	1,065	1,026	1,027	992	1,074
	BIC	1,107	1,073	1,070	1,041	1,117
	RMSE offline	350	336	359	299	626
	RMSE online	725	586	589	586	1,205
	MAE offline	271	271	283	241	467
	MAE online	582	465	465	465	990
Data: test	RMSE offline	1,233	1,270	885	678	1,890
	RMSE online	1,577	1,602	1,611	1,505	2,350
	MAE offline	937	902	710	535	1,450
	MAE online	1,050	871	904	848	1,438

Tabla 26: Modelos VAR y VARX para la categoría Decoración, análisis global. Fuente: Elaboración propia.

10.3.2 Clientes premium

Para la venta general de los clientes se modela con la opción 1 de la Tabla 27. La IRF asociada a la estimación se encuentra en el Gráfico 7.

		Modelo 1	Modelo 2	Modelo 3	Modelo 4	Modelo 5
Características modelo	Variable exógenas cyber	Si	Si	Si	Si	No
	Variable exógena navidad	Si	Si	Si	Si	No
	Variable exógena # tiendas	Si	Si	No	No	No
	Cantidad rezagos	5	6	5	4	5
Supuestos modelo	Raíces unitarias	Si	No	Si	Si	No
	Autocorrelacion residuos	No	No	No	No	No
	Normalidad residuos	Si	Si	No	Si	Si
	Homocedasticidad residuos	Si	Si	Si	Si	Si
Métricas data train	LL	-592	-553	-634	-594	-670
	AIC	1,241	1,171	1,314	1,242	1,383
	BIC	1,283	1,218	1,347	1,280	1,416
	RMSE offline	2,209	1,871	2,858	2,355	4,755
	RMSE online	1,708	1,021	2,768	1,730	3,822
	MAE offline	1,753	1,507	2,351	1,881	3,423
	MAE online	1,165	848	1,830	1,181	2,848
Métricas data test	RMSE offline	7,556	8,990	7,594	5,522	13,710
	RMSE online	2,866	4,935	4,234	2,684	4,730
	MAE offline	5,128	5,728	5,845	4,150	11,009
	MAE online	2,060	3,577	2,656	1,818	3,507

Tabla 27: Modelos VAR y VARX para ventas generales en clientes Premium, análisis general. Fuente: Elaboración propia.

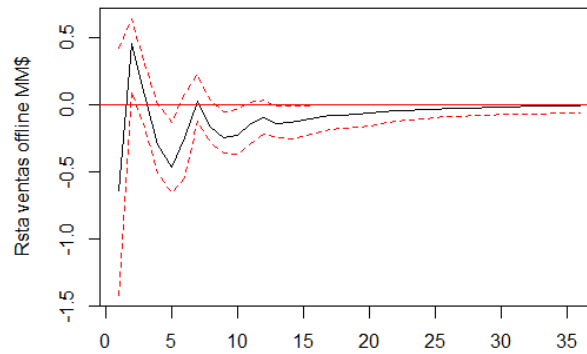


Gráfico 7: IRF con 95% de confianza, modelo de venta general clientes premium. Fuente: Elaboración propia

Según la comparación de la Tabla 28, se elige el cuarto, debido a que cumple con todos los supuestos y sus métricas en general son mejores. La IRF asociada al modelo está representada en el Gráfico 8.

		Modelo 1	Modelo 2	Modelo 3	Modelo 4	Modelo 5
Modelo	Variable exógenas cyber	Si	Si	Si	Si	No
	Variable exógena navidad	Si	Si	Si	Si	No
	Variable exógena # tiendas	Si	Si	No	No	No
	Cantidad rezagos	4	5	3	5	5
Supuestos	Raíces unitarias	Si	Si	Si	Si	Si
	Autocorrelacion residuos	No	No	No	No	No
	Normalidad residuos	Si	Si	Si	Si	No
	Homocedasticidad residuos	Si	Si	Si	Si	Si
Data: train	LL	-556	-526	-582	-531	-602
	AIC	1,160	1,109	1,201	1,115	1,248
	BIC	1,196	1,151	1,229	1,154	1,280
	RMSE offline	1,240	1,081	1,545	1,259	5,186
	RMSE online	631	476	644	476	938
	MAE offline	1,061	884	1,284	1,009	3,069
	MAE online	438	319	448	319	633
Data: test	RMSE offline	5,200	5,012	3,436	2,527	7,818
	RMSE online	1,133	969	1,113	868	1,260
	MAE offline	3,036	2,996	2,578	1,841	6,700
	MAE online	818	862	871	749	1,017

Tabla 28: Modelos VAR y VARX para categoría Vestuario, análisis clientes Premium. Fuente: Elaboración propia.

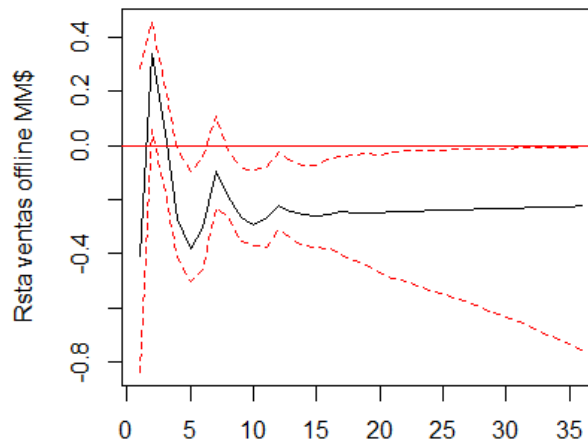


Gráfico 8: IRF con 95% de confianza, modelo de venta vestuario en clientes premium. Fuente: Elaboración propia

En cuanto a la venta de electrónica para los clientes Premium se decide que el “mejor” modelo es el cuarto de la Tabla 29. Su IRF se encuentra en el Gráfico 9.

		Modelo 1	Modelo 2	Modelo 3	Modelo 4	Modelo 5
Modelo	Variable exógenas cyber	Si	Si	Si	Si	No
	Variable exógena navidad	Si	Si	Si	Si	No
	Variable exógena # tiendas	Si	Si	No	No	No
	Cantidad rezagos	4	5	4	5	5
Supuestos	Raíces unitarias	Si	Si	Si	Si	Si
	Autocorrelacion residuos	No	No	No	No	No
	Normalidad residuos	No	Si	No	Si	No
	Homocedasticidad residuos	Si	Si	Si	Si	Si
Data: train	LL	-634	-552	-594	-553	-618
	AIC	1,303	1,161	1,232	1,158	1,279
	BIC	1,328	1,202	1,266	1,197	1,312
	RMSE offline	1,586	1,201	1,464	1,202	3,758
	RMSE online	1,780	905	1,602	928	3,879
	MAE offline	1,338	1,004	1,221	1,006	2,590
	MAE online	1,163	649	989	683	2,694
Data: test	RMSE offline	4,547	2,210	3,518	2,114	5,019
	RMSE online	2,319	1,260	2,393	1,438	2,750
	MAE offline	3,266	1,769	2,509	1,683	4,325
	MAE online	1,711	1,035	1,666	1,090	2,255

Tabla 29: Modelos VAR y VARX para categoría electrónica, análisis clientes Premium. Fuente: Elaboración propia.

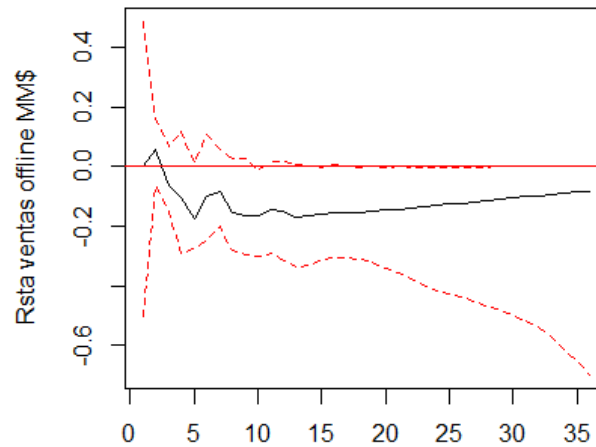


Gráfico 9: IRF con 95% de confianza, modelo de venta en electrónica clientes premium. Fuente: Elaboración propia

Finalmente, para la categoría Decoración, Tabla 30, se elige el número 4 pues es el que posee mejores métricas en relación con los otros. LA IRF asociada se encuentra en el Gráfico 10.

		Modelo 1	Modelo 2	Modelo 3	Modelo 4	Modelo 5
Modelo	Variable exógenas cyber	Si	Si	Si	Si	No
	Variable exógena navidad	Si	Si	Si	Si	No
	Variable exógena # tiendas	Si	Si	No	No	No
	Cantidad rezagos	6	7	4	7	5
Supuestos	Raíces unitarias	Si	Si	Si	Si	Si
	Autocorrelacion residuos	No	No	No	No	No
	Normalidad residuos	Si	Si	Si	Si	Si
	Homocedasticidad residuos	Si	Si	Si	Si	Si
Data: train	LL	-451	-424	-591	-433	-517
	AIC	966	920	1,045	934	1,078
	BIC	1,013	972	1,079	983	1,111
	RMSE offline	215	159	270	185	388
	RMSE online	383	328	537	370	974
	MAE offline	176	123	219	155	374
	MAE online	304	276	454	292	752
Data: test	RMSE offline	705	612	304	283	554
	RMSE online	1,089	1,392	986	1,073	1,678
	MAE offline	424	412	219	225	427
	MAE online	726	1,111	725	688	1,232

Tabla 30: Modelos VAR y VARX para la categoría decoración en clientes Premium, análisis venta Decoración. Fuente: Elaboración propia.

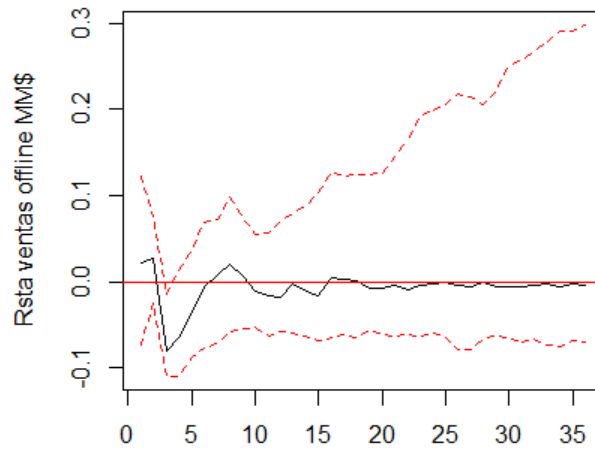


Gráfico 10: IRF con 95% de confianza, modelo de venta en decoración para clientes premium. Fuente: Elaboración propia

10.4 Clientes no Premium

Para la venta general de los clientes no Premium se modela con las características de la opción 3 de la Tabla 31. Su IRF asociada se encuentra en el Gráfico 11.

		Modelo 1	Modelo 2	Modelo 3	Modelo 4	Modelo 5
Modelo	Variable exógenas cyber	Si	Si	Si	Si	No
	Variable exógena navidad	Si	Si	Si	Si	No
	Variable exógena # tiendas	Si	Si	No	No	No
	Cantidad rezagos	5	6	5	6	6
Supuestos	Raíces unitarias	Si	Si	Si	No	No
	Autocorrelacion residuos	No	No	No	No	No
	Normalidad residuos	Si	Si	Si	Si	Si
	Homocedasticidad residuos	Si	Si	Si	Si	Si
Data: train	LL	-561	-524	-563	-525	-607
	AIC	1,178	1,112	1,178	1,112	1,266
	BIC	1,220	1,159	1,217	1,156	1,304
	RMSE offline	1,142	1,149	1,628	1,526	5,827
	RMSE online	996	544	1,021	544	2,305
	MAE offline	1,190	1,173	1,236	1,188	4,388
	MAE online	679	440	672	439	1,719
Data: test	RMSE offline	5,453	6,055	3,823	3,945	21,885
	RMSE online	3,544	3,946	3,146	3,936	4,515
	MAE offline	3,532	3,555	2,815	2,778	16,410
	MAE online	2,644	2,762	2,281	2,764	3,525

Tabla 31: Comparación modelos VAR y VARX para venta general, cliente no premium. Fuente: Elaboración propia

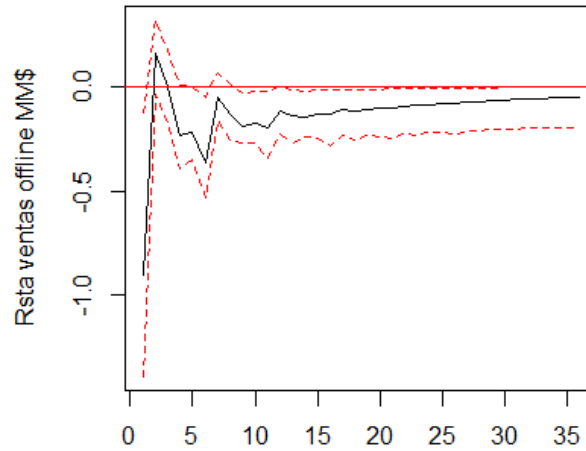


Gráfico 11: IRF con 95% de confianza, modelo de venta general clientes no premium. Fuente: Elaboración propia.

En cuanto a la venta en vestuario para los clientes no premium se decide modelar con la opción 4 de la Tabla 32, pues, aunque una de las raíces unitarias del polinomio característico es levemente mayor a 1, es el modelo con mejores métricas. La IRF del modelo se aprecia en el Gráfico 12.

		Modelo 1	Modelo 2	Modelo 3	Modelo 4	Modelo 5
Modelo	Variable exógenas cyber	Si	Si	Si	Si	No
	Variable exógena navidad	Si	Si	Si	Si	No
	Variable exógena # tiendas	Si	Si	No	No	No
	Cantidad rezagos	5	6	2	5	6
Supuestos	Raíces unitarias	Si	No	Si	No	No
	Autocorrelacion residuos	No	No	No	No	No
	Normalidad residuos	Si	Si	Si	Si	Si
	Homocedasticidad residuos	Si	Si	Si	Si	Si
Data: train	LL	-523	-488	-598	-527	-566
	AIC	1,203	1,041	1,225	1,107	1,184
	BIC	1,145	1,088	1,247	1,145	1,122
	RMSE offline	1,236	1,207	1,556	1,385	4,785
	RMSE online	383	218	586	384	718
	MAE offline	1,026	979	1,494	1,096	3,317
	MAE online	250	167	371	250	524
Data: test	RMSE offline	5,008	5,931	3,558	2,680	20,441
	RMSE online	1,062	1,924	1,847	962	2,081
	MAE offline	2,876	3,211	2,277	1,974	14,827
	MAE online	938	1,787	1,340	819	1,704

Tabla 32: Modelos VAR y VARX para vestuario, clientes no Premium. Fuente: Elaboración propia.

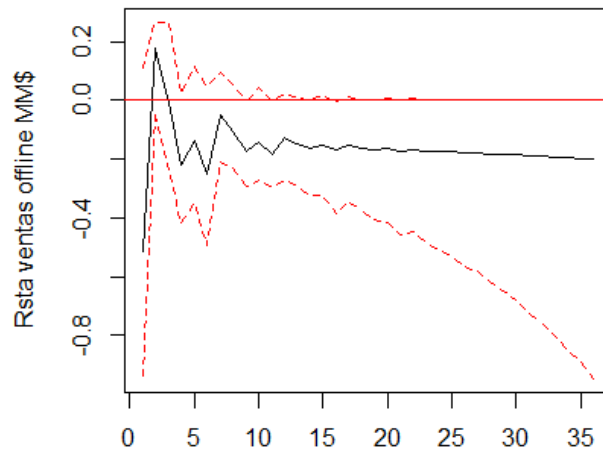


Gráfico 12: IRF con 95% de confianza, modelo de venta vestuario en clientes no premium. Fuente: Elaboración propia

Los modelos de la Tabla 33 indican las opciones para representar las ventas de clientes no premium para la categoría electrónica, de ellos se elige la opción 1. La IRF asociada se encuentra en el Gráfico 13.

		Modelo 1	Modelo 2	Modelo 3	Modelo 4	Modelo 5
Modelo	Variable exógenas cyber	Si	Si	Si	Si	No
	Variable exógena navidad	Si	Si	Si	Si	No
	Variable exógena # tiendas	Si	Si	No	No	No
	Cantidad rezagos	5	6	5	6	
Supuestos	Raíces unitarias	Si	Si	Si	Si	Si
	Autocorrelacion residuos	No	Si	No	Si	No
	Normalidad residuos	Si	Si	Si	Si	No
	Homocedasticidad residuos	Si	Si	Si	Si	Si
Data: train	LL	-500	-466	-502	-466	-672
	AIC	1,055	995	1,056	992	1,357
	BIC	1,097	1,042	1,095	1,036	1,367
	RMSE offline	560	523	569	527	2,479
	RMSE online	404	234	424	234	1,849
	MAE offline	436	411	437	409	1,616
	MAE online	309	184	312	182	1,319
Data: test	RMSE offline	886	1,107	1,087	1,258	2,829
	RMSE online	1,725	1,944	1,660	1,939	3,937
	MAE offline	728	942	877	1,022	2,009
	MAE online	1,148	1,294	1,159	1,284	2,410

Tabla 33: Modelos VAR y VARX para electrónica en clientes no premium. Fuente: Elaboración propia.

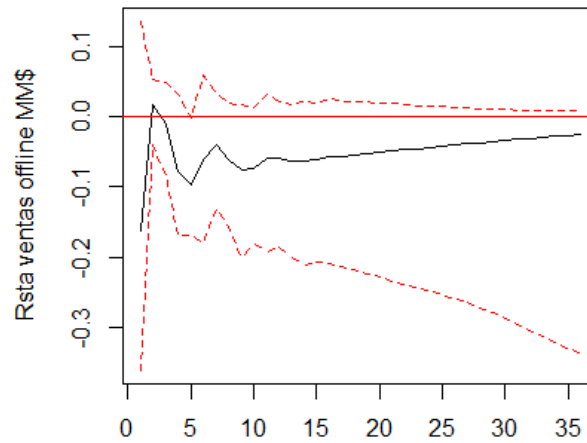


Gráfico 13: IRF con 95% de confianza, modelo de venta electrónica en clientes no premium. Fuente: Elaboración propia

Finalmente, para los clientes no premium en la categoría de decoración se considera el modelo 4 de la Tabla 34. El Gráfico 14 presenta la IRF asociada.

		Modelo 1	Modelo 2	Modelo 3	Modelo 4	Modelo 5
Modelo	Variable exógenas cyber	Si	Si	Si	Si	No
	Variable exógena navidad	Si	Si	Si	Si	No
	Variable exógena # tiendas	Si	Si	No	No	No
	Cantidad rezagos	4	5	5	6	6
Supuestos	Raíces unitarias	Si	Si	Si	Si	No
	Autocorrelacion residuos	No	No	No	No	No
	Normalidad residuos	Si	Si	Si	Si	Si
	Homocedasticidad residuos	Si	Si	Si	Si	Si
Data: train	LL	-464	-441	-442	-416	
	AIC	977	938	935	892	975
	BIC	1,013	970	974	936	1,013
	RMSE offline	144	130	131	130	238
	RMSE online	255	285	289	201	451
	MAE offline	108	105	109	106	197
	MAE online	291	218	214	162	356
Data: test	RMSE offline	376	281	239	240	627
	RMSE online	792	677	644	648	1,008
	MAE offline	199	194	172	204	516
	MAE online	624	517	464	402	711

Tabla 34: Modelos VAR y VARX para decoración en clientes no premium. Fuente: Elaboración propia.

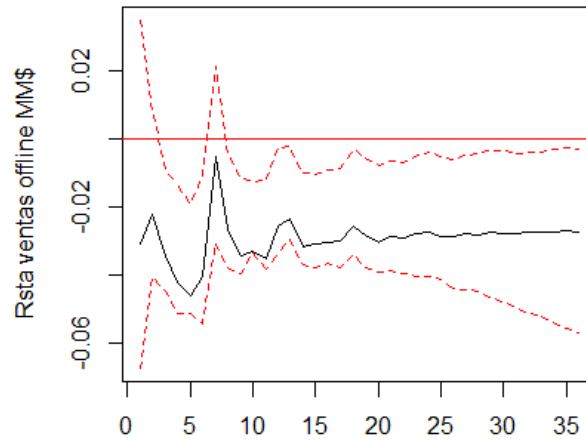


Gráfico 14: IRF con 95% de confianza, modelo de venta decoración en clientes no premium. Fuente: Elaboración propia

10.5 Análisis por zona

10.5.1 Zona norte

Para las ventas generales de la zona norte se utiliza el modelo 4 de la Tabla 35, cuya IRF asociada se encuentra en el Gráfico 15.

		Modelo 1	Modelo 2	Modelo 3	Modelo 4	Modelo 5
Características modelo	Variable exógenas cyber	Si	Si	Si	Si	No
	Variable exógena navidad	Si	Si	Si	Si	No
	Variable exógena # tiendas	Si	Si	No	No	No
	Cantidad rezagos	2	5	3	5	5
Supuestos modelo	Raíces unitarias	Si	Si	Si	Si	Si
	Autocorrelacion residuos	No	No	No	No	No
	Normalidad residuos	Si	Si	No	Si	No
	Homocedasticidad residuos	Si	Si	Si	Si	Si
Métricas data train	LL	-586	-504	-568	-507	-601
	AIC	1,203	1,063	1,172	1,065	1,247
	BIC	1,229	1,195	1,199	1,104	1,280
	RMSE offline	957	638	934	695	4,029
	RMSE online	826	420	772	425	1,198
	MAE offline	782	492	771	485	2,415
Métricas data test	MAE online	575	272	520	279	835
	RMSE offline	3,189	2,440	3,012	2,467	5,980
	RMSE online	2,124	1,026	1,665	1,060	1,667
	MAE offline	2,411	1,940	2,245	1,869	4,873
	MAE online	1,591	766	1,185	750	1,285

Tabla 35: Modelos VAR y VARX para clientes zona norte en venta general. Fuente: Elaboración propia.

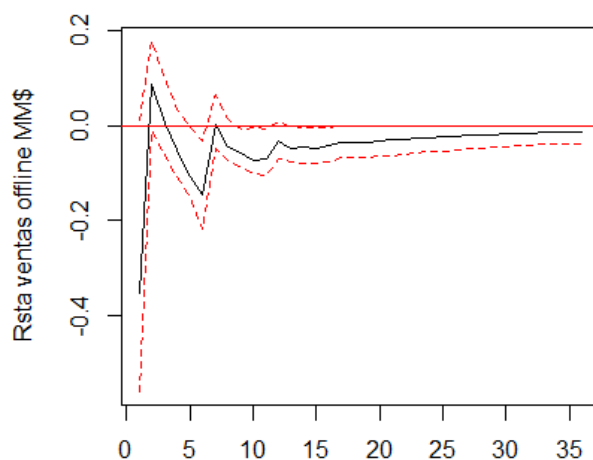


Gráfico 15: IRF con 95% de confianza, modelo de venta general en zona norte. Fuente: Elaboración propia

Se modelan las ventas de la zona norte en la categoría vestuario con el modelo 2 de la Tabla 36 y la IRF asociada en el Gráfico 16.

		Modelo 1	Modelo 2	Modelo 3	Modelo 4	Modelo 5
Modelo	Variable exógenas cyber	Si	Si	Si	Si	No
	Variable exógena navidad	Si	Si	Si	Si	No
	Variable exógena # tiendas	Si	Si	No	No	No
	Cantidad rezagos	2	5	2	5	5
Supuestos	Raíces unitarias	Si	Si	Si	Si	Si
	Autocorrelacion residuos	No	No	No	No	No
	Normalidad residuos	Si	Si	Si	No	No
	Homocedasticidad residuos	Si	Si	Si	Si	Si
Data: train	LL	-518	-452	-526	-461	-542
	AIC	1,069	960	1,080	975	1,129
	BIC	1,094	1,002	1,102	1,013	1,162
	RMSE offline	597	396	660	516	1,706
	RMSE online	205	140	210	140	298
	MAE offline	503	298	534	384	1,627
	MAE online	144	91	144	91	201
Data: test	RMSE offline	1,906	1,451	1,905	1,388	4,178
	RMSE online	668	326	656	328	509
	MAE offline	1,415	1,098	1,434	1,045	3,502
	MAE online	502	275	495	270	393

Tabla 36: Modelos VAR y VARX para clientes zona norte en vestuario. Fuente: Elaboración propia.

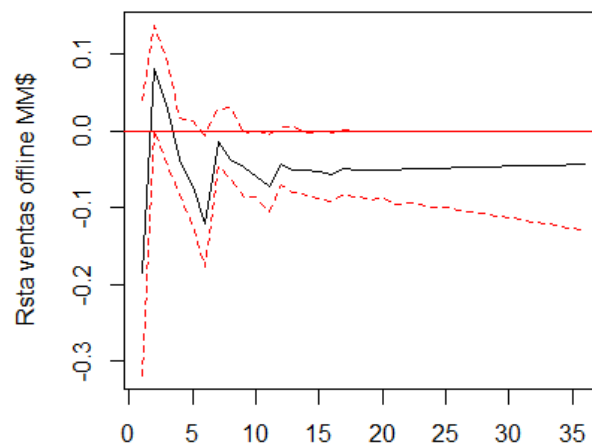


Gráfico 16: IRF con 95% de confianza, modelo de venta vestuario en zona norte. Fuente: Elaboración propia

En cuanto a las ventas en electrónica en la zona norte, estas se modelan con la segunda opción de la Tabla 37. La IRF asociada se encuentra en el Gráfico 17.

		Modelo 1	Modelo 2	Modelo 3	Modelo 4	Modelo 5
Modelo	Variable exógenas cyber	Si	Si	Si	Si	No
	Variable exógena navidad	Si	Si	Si	Si	No
	Variable exógena # tiendas	Si	Si	No	No	No
	Cantidad rezagos	4	5	4	5	5
Supuestos	Raíces unitarias	Si	Si	Si	Si	Si
	Autocorrelacion residuos	Si	Si	No	No	Si
	Normalidad residuos	No	Si	No	Si	No
	Residuos Homocedasticos	Si	Si	Si	Si	Si
Data: train	LL	-495	-453	-497	-453	-539
	AIC	1,039	961	1,039	959	1,123
	BIC	1,075	1,003	1,072	998	1,156
	RMSE offline	358	269	367	272	1,170
	RMSE online	375	202	383	203	634
	MAE offline	276	212	292	221	720
	MAE online	240	161	240	159	483
Data: test	RMSE offline	2,195	2,052	2,112	2,067	1,947
	RMSE online	2,117	2,365	2,351	2,348	1,497
	MAE offline	1,946	1,822	1,863	1,835	1,518
	MAE online	2,196	2,204	2,149	2,193	2,211

Tabla 37: Modelos VAR y VARX para clientes zona norte en electrónica. Fuente: Elaboración propia.

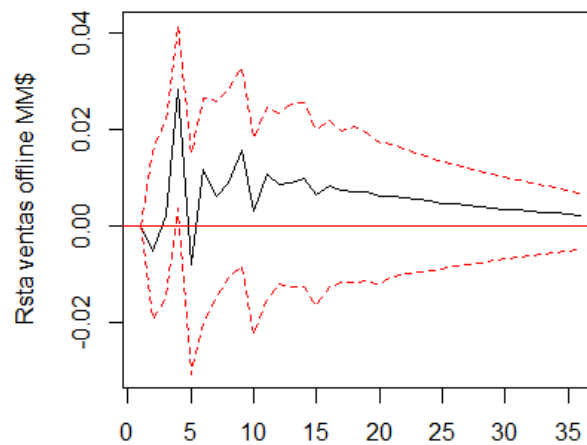


Gráfico 17: IRF con 95% de confianza, modelo de venta electrónica en zona norte. Fuente: Elaboración propia

Finalmente, para las ventas de decoración de la zona norte se utiliza el modelo 2 de la Tabla 38 y su función de IRF se encuentra en el Gráfico 18.

		Modelo 1	Modelo 2	Modelo 3	Modelo 4	Modelo 5
Modelo	Variable exógenas cyber	Si	Si	Si	Si	No
	Variable exógena navidad	Si	Si	Si	Si	No
	Variable exógena # tiendas	Si	Si	No	No	No
	Cantidad rezagos	4	5	4	5	
Supuestos	Raíces unitarias	Si	Si	Si	Si	Si
	Autocorrelacion residuos	No	No	No	No	No
	Normalidad residuos	Si	Si	Si	Si	Si
	Homocedasticidad residuos	Si	Si	Si	Si	Si
Data: train	LL	-419	-398	-420	-400	-460
	AIC	886	853	883	852	955
	BIC	922	895	917	891	982
	RMSE offline	87	73	88	77	133
	RMSE online	159	140	160	140	327
	MAE offline	69	60	69	64	105
	MAE online	133	112	135	112	241
Data: test	RMSE offline	146	162	150	167	216
	RMSE online	301	298	294	299	594
	MAE offline	88	117	92	125	174
	MAE online	188	170	179	171	372

Tabla 38: Modelos VAR y VARX para clientes en zona norte en decoración. Fuente: Elaboración propia.

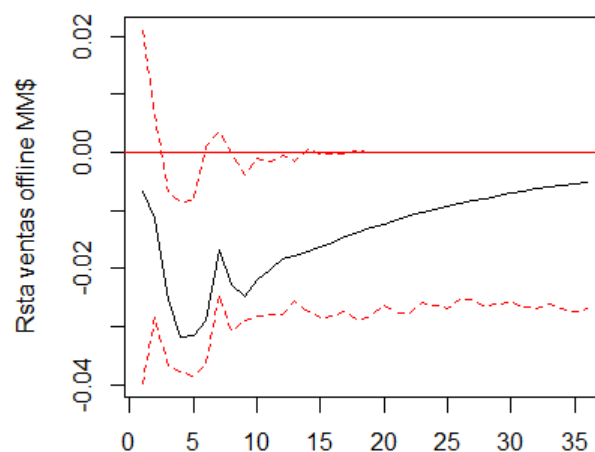


Gráfico 18: IRF con 95% de confianza, modelo de venta decoración en zona norte. Fuente: Elaboración propia

10.5.2 Zona RM Oriente

Desde la Tabla 39 se elige el modelo 2 para representar las ventas generales de la zona RM Oriente. El Gráfico 19, muestra la función de IRF asociada al modelo.

		Modelo 1	Modelo 2	Modelo 3	Modelo 4	Modelo 5
Características modelo	Variable exógenas cyber	Si	Si	Si	Si	No
	Variable exógena navidad	Si	Si	Si	Si	No
	Variable exógena # tiendas	Si	Si	No	No	No
	Cantidad rezagos	3	5	2	3	5
Supuestos modelo	Raíces unitarias	Si	Si	Si	Si	Si
	Autocorrelacion residuos	No	No	No	No	No
	Normalidad residuos	Si	Si	Si	Si	No
	Homocedasticidad residuos	Si	Si	Si	Si	Si
Métricas data train	LL	-538	-489	-559	-539	-563
	AIC	1,116	1,034	1,146	1,114	1,169
	BIC	1,147	1,076	1,168	1,142	1,202
	RMSE offline	533	392	571	534	1,446
	RMSE online	544	417	616	554	1,031
	MAE offline	438	329	479	443	925
	MAE online	419	278	471	401	715
Métricas data test	RMSE offline	1,268	962	1,456	1,246	2,124
	RMSE online	870	744	997	853	1,087
	MAE offline	923	688	1,164	917	1,721
	MAE online	714	618	829	672	874

Tabla 39: Modelos VAR y VARX para la venta general en clientes de zona RM Oriente. Fuente: Elaboración propia.

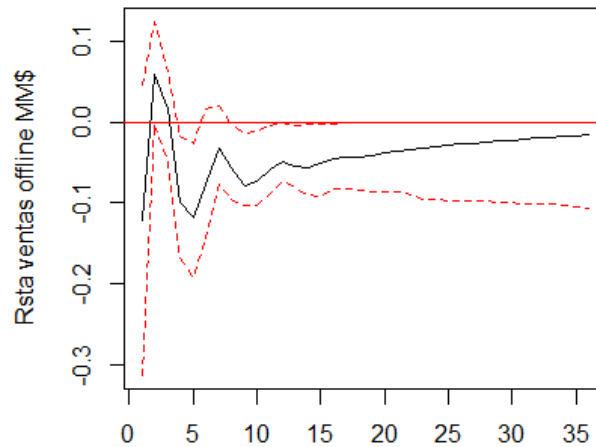


Gráfico 19: IRF con 95% de confianza, modelo de venta general en zona oriente. Fuente: Elaboración propia

En cuanto a las ventas de vestuario en la zona RM Oriente, estas se modelan con la opción 2 y su IRF se encuentra en el Gráfico 20.

		Modelo 1	Modelo 2	Modelo 3	Modelo 4	Modelo 5
Modelo	Variable exógenas cyber	Si	Si	Si	Si	No
	Variable exógena navidad	Si	Si	Si	Si	No
	Variable exógena # tiendas	Si	Si	No	No	No
	Cantidad rezagos	4	5	3	5	5
Supuestos	Raíces unitarias	Si	Si	Si	Si	Si
	Autocorrelacion residuos	No	No	Si	Si	No
	Normalidad residuos	Si	Si	Si	Si	No
	Homocedasticidad residuos	Si	Si	Si	Si	Si
Data: train	LL	-448	-424	-474	-431	-501
	AIC	943	903	985	914	1,047
	BIC	980	945	1,013	953	1,080
	RMSE offline	254	231	348	287	1,109
	RMSE online	123	95	130	96	220
	MAE offline	207	185	275	229	702
	MAE online	89	58	89	58	153
Data: test	RMSE offline	722	660	734	679	1,594
	RMSE online	263	201	260	200	289
	MAE offline	469	471	544	506	1,349
	MAE online	181	146	171	142	235

Tabla 40: Modelos VAR y VARX para la categoría vestuario en clientes de zona RM Oriente. Fuente: Elaboración propia.

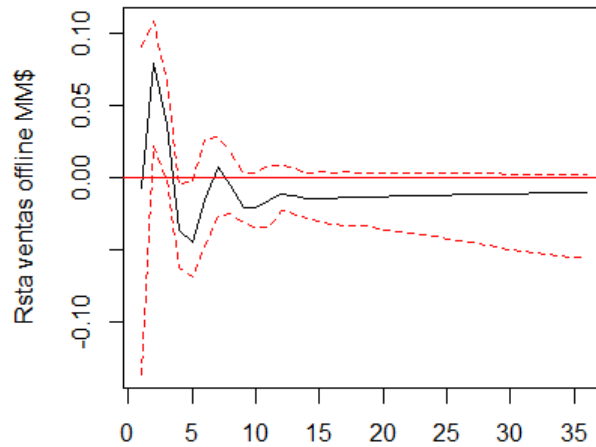


Gráfico 20: IRF con 95% de confianza, modelo de venta vestuario en zona RM Oriente. Fuente: Elaboración propia

Por otra parte, el primer modelo de la Tabla 41 se ha utilizado para representar las ventas en electrónica de la zona. La función de IRF asociada se encuentra en el Gráfico 21.

		Modelo 1	Modelo 2	Modelo 3	Modelo 4	Modelo 5
Modelo	Variable exógenas cyber	Si	Si	Si	Si	No
	Variable exógena navidad	Si	Si	Si	Si	No
	Variable exógena # tiendas	Si	Si	No	No	No
	Cantidad rezagos	3	6	3	6	6
Supuestos	Raíces unitarias	Si	Si	Si	Si	Si
	Autocorrelacion residuos	No	Si	No	Si	No
	Normalidad residuos	No	Si	No	Si	Si
	Homocedasticidad residuos	Si	Si	Si	Si	Si
Data: train	LL	-487	-421	-487	-421	-469
	AIC	1,013	905	1,011	903	990
	BIC	1,045	952	1,038	947	1,028
	RMSE offline	233	199	236	203	293
	RMSE online	287	162	290	164	461
	MAE offline	194	167	195	172	245
	MAE online	212	134	206	136	374
Data: test	RMSE offline	400	325	358	383	459
	RMSE online	483	666	530	549	652
	MAE offline	308	284	270	294	342
	MAE online	405	620	481	518	564

Tabla 41: Modelos VAR y VARX para la categoría electrónica en clientes RM Oriente. Fuente: Elaboración propia.

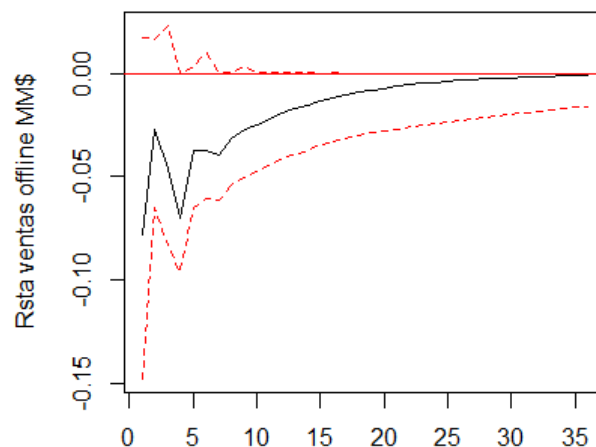


Gráfico 21: IRF con 95% de confianza, modelo de venta electrónica en zona RM Oriente. Fuente: Elaboración propia

Finalmente, los modelos de la Tabla 42 representan las opciones para modelar las ventas en RM Oriente para la categoría decoración, de ellos se elige el número cuatro, la IRF asociada se encuentra en la Gráfica 22.

		Modelo 1	Modelo 2	Modelo 3	Modelo 4	Modelo 5
Modelo	Variable exógenas cyber	Si	Si	Si	Si	No
	Variable exógena navidad	Si	Si	Si	Si	No
	Variable exógena # tiendas	Si	Si	No	No	No
	Cantidad rezagos	3	4	3	4	4
Supuestos	Raíces unitarias	Si	Si	Si	Si	Si
	Autocorrelacion residuos	No	No	No	No	No
	Normalidad residuos	Si	Si	Si	Si	Si
	Homocedasticidad residuos	Si	Si	Si	Si	Si
Data: train	LL	-412	-398	-422	-402	-452
	AIC	882	844	879	848	940
	BIC	914	881	907	881	967
	RMSE offline	68	50	69	55	118
	RMSE online	146	144	147	146	297
	MAE offline	51	40	53	42	97
	MAE online	116	112	116	112	240
Data: test	RMSE offline	100	93	92	81	157
	RMSE online	239	245	238	237	428
	MAE offline	86	80	76	63	113
	MAE online	217	227	215	215	340

Tabla 42: Modelos VAR y VARX para la categoría decoración en clientes de zona RM Oriente. Fuente: Elaboración propia.

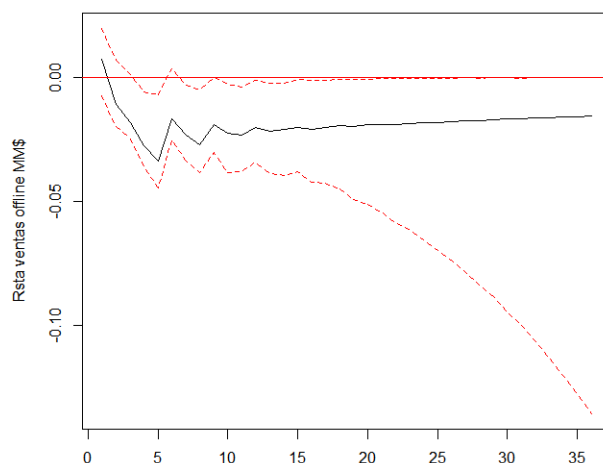


Gráfico 22: IRF con 95% de confianza, modelo de venta decoración en zona RM Oriente. Fuente: Elaboración propia

10.5.3 Zona RM Poniente

Las ventas generales de la zona a son modeladas con la segunda opción presentada en la Tabla 43. El gráfico 23 representa la IRF asociada.

		Modelo 1	Modelo 2	Modelo 3	Modelo 4	Modelo 5
Características modelo	Variable exógenas cyber	Si	Si	Si	Si	No
	Variable exógena navidad	Si	Si	Si	Si	No
	Variable exógena # tiendas	Si	Si	No	No	No
	Cantidad rezagos	3	5	2	3	5
Supuestos modelo	Raíces unitarias	Si	Si	Si	Si	Si
	Autocorrelacion residuos	No	No	No	No	No
	Normalidad residuos	Si	Si	Si	Si	No
	Homocedasticidad residuos	Si	Si	Si	Si	Si
Métricas data train	LL	-538	-489	-559	-539	-563
	AIC	1,116	1,034	1,146	1,114	1,169
	BIC	1,147	1,076	1,168	1,142	1,202
	RMSE offline	533	392	571	534	1,446
	RMSE online	544	417	616	554	1,031
	MAE offline	438	329	479	443	925
Métricas data test	MAE online	419	278	471	401	715
	RMSE offline	1,268	962	1,456	1,246	2,124
	RMSE online	870	744	997	853	1,087
	MAE offline	923	688	1,164	917	1,721
	MAE online	714	618	829	672	874

Tabla 43: Modelos VAR y VARX para la venta general en clientes de zona RM Poniente. Fuente: Elaboración propia.

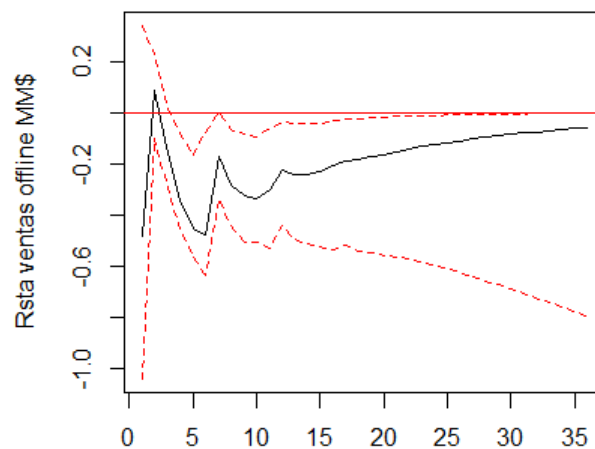


Gráfico 23: IRF con 95% de confianza, modelo de venta general en zona RM Poniente. Fuente: Elaboración propia

Para las ventas de la zona en productos de vestuario, se utiliza el modelo 2 de la Tabla 44. Además, en el Gráfico 24 se tiene la función IRF asociada.

		Modelo 1	Modelo 2	Modelo 3	Modelo 4	Modelo 5
Modelo	Variable exógenas cyber	Si	Si	Si	Si	No
	Variable exógena navidad	Si	Si	Si	Si	No
	Variable exógena # tiendas	Si	Si	No	No	No
	Cantidad rezagos	3	5	3	5	5
Supuestos	Raíces unitarias	Si	No	Si	No	Si
	Autocorrelacion residuos	No	No	No	No	No
	Normalidad residuos	Si	Si	No	Si	No
	Homoedasticidad residuos	Si	Si	Si	Si	Si
Data: train	LL	-552	-500	-559	-509	-593
	AIC	1,145	1,057	1,153	1,069	1,230
	BIC	1,176	1,098	1,181	1,108	1,263
	RMSE offline	936	663	1,025	834	4,843
	RMSE online	488	361	504	361	775
	MAE offline	779	522	840	662	2,830
	MAE online	342	240	331	240	530
Data: test	RMSE offline	2,554	1,919	2,444	1,838	7,182
	RMSE online	926	671	866	659	1,018
	MAE offline	1,644	1,590	1,451	1,392	5,903
	MAE online	660	566	625	552	818

Tabla 44: Modelos VAR y VARX para la categoría vestuario en clientes RM Poniente. Fuente: Elaboración propia.

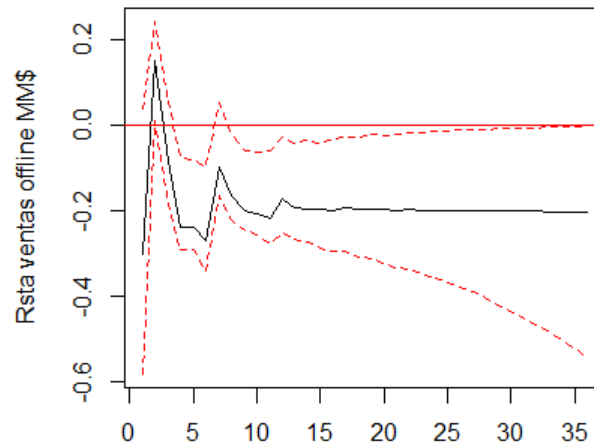


Gráfico 24: IRF con 95% de confianza, modelo de venta vestuario en zona RM Poniente. Fuente: Elaboración propia

Considerando las opciones de la Tabla 45, para las ventas en electrónica se ha elegido la opción 4. La función de impulso respuesta asociada se encuentra en el Gráfico 25.

		Modelo 1	Modelo 2	Modelo 3	Modelo 4	Modelo 5
Modelo	Variable exógenas cyber	Si	Si	Si	Si	No
	Variable exógena navidad	Si	Si	Si	Si	No
	Variable exógena # tiendas	Si	Si	No	No	No
	Cantidad rezagos	5	6	2	5	5
Supuestos	Raíces unitarias	Si	Si	Si	Si	Si
	Autocorrelacion residuos	No	Si	No	No	No
	Normalidad residuos	Si	Si	No	Si	No
	Homocedasticidad residuos	Si	Si	Si	Si	Si
Data: train	LL	-526	-492	-604	-527	-596
	AIC	1,109	1,049	1,236	1,195	1,236
	BIC	1,151	1,096	1,258	1,144	1,269
	RMSE offline	792	708	1,089	794	2,363
	RMSE online	627	414	1,151	631	1,729
	MAE offline	664	563	934	664	1,471
	MAE online	438	309	723	441	1,226
Data: test	RMSE offline	1,679	1,861	2,426	1,685	3,435
	RMSE online	909	1,747	1,331	908	1,731
	MAE offline	1,271	1,529	1,788	1,296	2,964
	MAE online	737	1,293	953	714	1,437

Tabla 45: Modelos VAR y VARX para la categoría electrónica en clientes zona RM Poniente. Fuente: Elaboración propia.

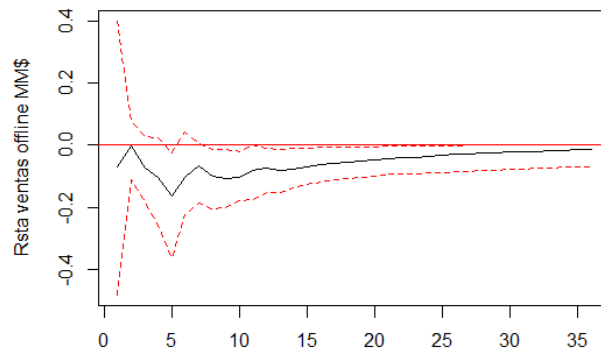


Gráfico 25: IRF con 95% de confianza, modelo de venta electrónica en zona RM Poniente. Fuente: Elaboración propia

Para las ventas de decoración en la zona RM Poniente se elige modelar con la primera opción de la Tabla 46. La función IRF asociada al modelo se encuentra en el Gráfico 26.

		Modelo 1	Modelo 2	Modelo 3	Modelo 4	Modelo 5
Modelo	Variable exógenas cyber	Si	Si	Si	Si	No
	Variable exógena navidad	Si	Si	Si	Si	No
	Variable exógena # tiendas	Si	Si	No	No	No
	Cantidad rezagos	5	6	3	4	
Supuestos	Raíces unitarias	Si	Si	Si	Si	Si
	Autocorrelacion residuos	No	No	No	No	No
	Normalidad residuos	Si	Si	Si	Si	Si
	Homocedasticidad residuos	Si	Si	Si	Si	Si
Data: train	LL	-439	-419	-489	-465	-535
	AIC	934	902	1,014	974	1,098
	BIC	976	949	1,042	1,008	1,120
	RMSE offline	132	126	217	168	361
	RMSE online	268	230	317	304	706
	MAE offline	111	103	187	128	284
	MAE online	219	180	252	251	516
Data: test	RMSE offline	221	269	249	223	557
	RMSE online	497	571	569	568	1,285
	MAE offline	177	213	139	146	496
	MAE online	326	426	380	413	790

Tabla 46: Modelos VAR y VARX para la categoría decoración en clientes zona RM Poniente. Fuente: Elaboración propia.

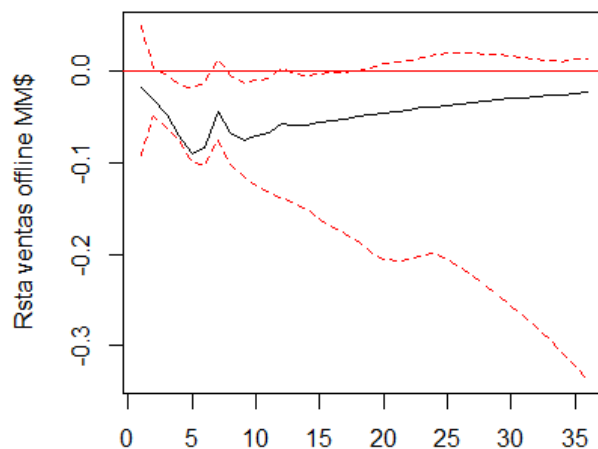


Gráfico 26: IRF con 95% de confianza, modelo de venta decoración en zona RM Poniente. Fuente: Elaboración propia

10.5.4 Zona Sur

Las ventas generales de la zona se modelan a través de la cuarta opción de la Tabla 47. La función de impulso respuesta vinculada al modelo está en el Gráfico 27.

		Modelo 1	Modelo 2	Modelo 3	Modelo 4	Modelo 5
Características modelo	Variable exógenas cyber	Si	Si	Si	Si	No
	Variable exógena navidad	Si	Si	Si	Si	No
	Variable exógena # tiendas	Si	Si	No	No	No
	Cantidad rezagos	5	6	5	6	5
Supuestos modelo	Raíces unitarias	Si	Si	Si	Si	Si
	Autocorrelacion residuos	No	No	No	No	No
	Normalidad residuos	Si	Si	Si	Si	No
	Homoceasticidad residuos	Si	Si	Si	Si	Si
Métricas data train	LL	-537	-501	-539	-505	-618
	AIC	1,131	1,065	1,130	1,070	1,281
	BIC	1,173	1,112	1,169	1,114	1,313
	RMSE offline	1,062	915	1,123	1,042	4,878
	RMSE online	676	398	681	403	1,649
	MAE offline	845	714	814	793	2,900
	MAE online	451	339	470	344	1,201
Métricas data test	RMSE offline	1,804	1,979	1,712	1,844	6,347
	RMSE online	1,702	2,137	1,734	2,162	2,449
	MAE offline	1,697	1,633	1,480	1,451	5,065
	MAE online	1,198	1,458	1,177	1,413	1,675

Tabla 47: Modelos VAR y VARX para venta general en clientes zona Sur. Fuente: Elaboración propia.

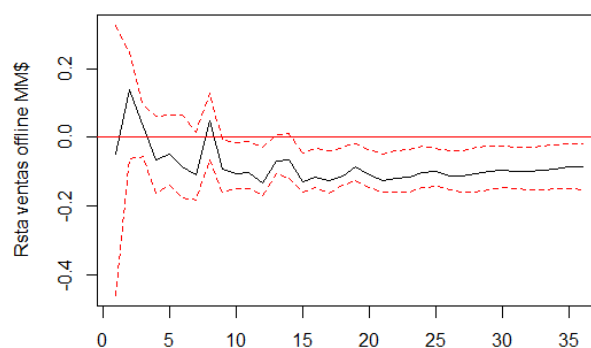


Gráfico 27: IRF con 95% de confianza, modelo de venta general en zona RM Sur. Fuente: Elaboración propia

Considerando las opciones de modelos presentados en la Tabla 48, para las ventas en vestuario se elige el número 4. La IRF vinculada se encuentra en el Gráfico 28.

		Modelo 1	Modelo 2	Modelo 3	Modelo 4	Modelo 5
Modelo	Variable exógenas cyber	Si	Si	Si	Si	No
	Variable exógena navidad	Si	Si	Si	Si	No
	Variable exógena # tiendas	Si	Si	No	No	No
	Cantidad rezagos	3	5	3	5	5
Supuestos	Raíces unitarias	Si	No	Si	No	Si
	Autocorrelacion residuos	No	No	No	No	No
	Normalidad residuos	Si	Si	Si	Si	No
	Homoceasticidad residuos	Si	Si	Si	Si	Si
Data: train	LL	-533	-484	-537	-489	-560
	AIC	1,106	1,024	1,109	1,030	1,165
	BIC	1,137	1,066	1,137	1,069	1,197
	RMSE offline	888	685	922	782	3,152
	RMSE online	290	217	299	216	439
	MAE offline	779	554	778	611	1,863
	MAE online	209	147	208	147	317
Data: test	RMSE offline	1,345	1,154	1,268	1,042	4,098
	RMSE online	654	497	613	491	742
	MAE offline	1,004	1,023	831	817	3,243
	MAE online	515	453	503	447	605

Tabla 48: Modelos VAR y VARX para la categoría vestuario en clientes zona Sur. Fuente: Elaboración propia.

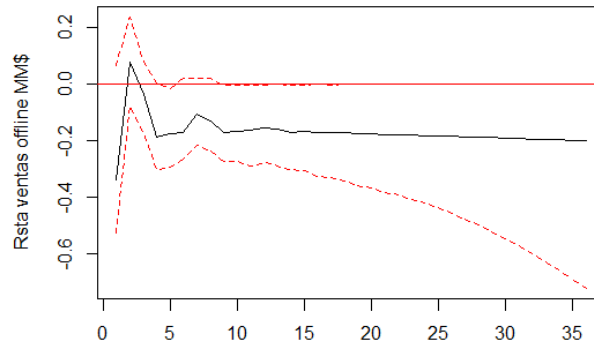


Gráfico 28: IRF con 95% de confianza, modelo de venta vestuario en zona Sur. Fuente: Elaboración propia

Los productos de electrónica vendidos en la zona sur se modelan a través del modelo 1. La IRF asociada al modelo se encuentra en el Gráfico 29.

		Modelo 1	Modelo 2	Modelo 3	Modelo 4	Modelo 5
Modelo	Variable exógenas cyber	Si	Si	Si	Si	No
	Variable exógena navidad	Si	Si	Si	Si	No
	Variable exógena # tiendas	Si	Si	No	No	No
	Cantidad rezagos	3	4	5	6	5
Supuestos	Raíces unitarias	Si	Si	No	Si	Si
	Autocorrelacion residuos	No	No	No	No	No
	Normalidad residuos	Si	Si	Si	Si	Si
	Homoceadasticidad residuos	Si	Si	Si	Si	Si
Data: train	LL	-546	-523	-482	-443	-556
	AIC	1,135	1,095	1,017	947	1,157
	BIC	1,166	1,132	1,056	991	1,189
	RMSE offline	600	510	393	310	1,461
	RMSE online	609	564	344	212	839
	MAE offline	495	404	311	235	952
Data: test	MAE online	395	389	265	167	637
	RMSE offline	1,136	926	1,519	819	2,007
	RMSE online	1,201	1,496	1,162	1,093	1,239
	MAE offline	898	834	1,450	697	1,729
	MAE online	780	1,188	940	706	873

Tabla 49: Modelos VAR y VARX, categoría electrónica clientes zona Sur. Fuente: Elaboración propia.

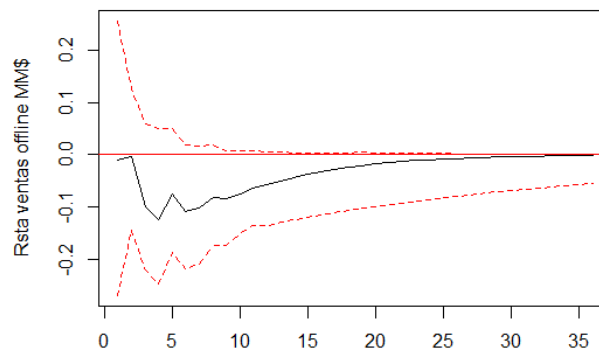


Gráfico 29: IRF 95% de confianza, venta electrónica en zona RM Poniente. Fuente: Elaboración propia

Para finalizar, se tiene el modelo 4 presentado en la Tabla 50 para las ventas en decoración en la zona sur. La función de impulso respuesta vinculada al modelo se encuentra en el Gráfico 30.

		Modelo 1	Modelo 2	Modelo 3	Modelo 4	Modelo 5
Modelo	Variable exógenas cyber	Si	Si	Si	Si	No
	Variable exógena navidad	Si	Si	Si	Si	No
	Variable exógena # tiendas	Si	Si	No	No	No
	Cantidad rezagos	4	5	4	5	5
Supuestos	Raíces unitarias	Si	Si	Si	Si	Si
	Autocorrelacion residuos	No	No	No	No	No
	Normalidad residuos	Si	Si	Si	Si	No
	Homocedasticidad residuos	Si	Si	Si	Si	Si
Data: train	LL	-451	-430	-454	-433	-475
	AIC	949	917	951	918	992
	BIC	986	958	985	957	1,025
	RMSE offline	138	123	145	133	242
	RMSE online	245	220	252	220	434
	MAE offline	107	97	115	108	168
	MAE online	198	177	204	178	348
Data: test	RMSE offline	187	183	148	145	284
	RMSE online	472	436	446	433	806
	MAE offline	160	155	119	131	216
	MAE online	338	251	316	249	498

Tabla 50: Modelos VAR y VARX para la categoría decoración en clientes zona Sur. Fuente: Elaboración propia.

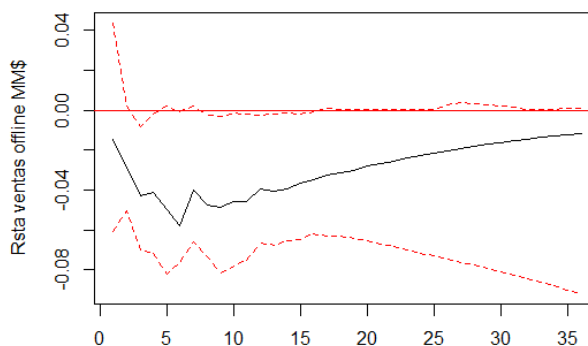


Gráfico 30: IRF con 95% de confianza, modelo de venta decoración en zona RM Sur. Fuente: Elaboración propia

10.6 Comparaciones estimaciones de modelos

A continuación, se presentan en detalle las estimaciones de los modelos por categoría de producto.

10.6.1 Clientes premium

Ventas clientes premium por categoria

	<i>Variable dependiente:</i>					
	Modelo vestuario		Clientes premium Modelo electrónica		Modelo decoración	
	Online	Offline	Online	Offline	Online	Offline
Online.l1	0.25**	0.51*	0.11	0.05	0.03	0.04
Offline.l1	-0.04	-0.06	-0.05	-0.02	-0.30	0.31*
Online.l2	0.30***	-0.005	0.16**	-0.06	0.17	-0.16**
Offline.l2	-0.03	0.05	-0.06	0.05	-0.07	-0.07
Online.l3	0.12	-0.61**	0.07	-0.09	0.11	-0.08
Offline.l3	0.001	0.07	0.04	0.30**	0.53*	0.07
Online.l4	-0.03	-0.60**	0.15**	-0.14*	-0.13	-0.03
Offline.l4	-0.01	0.11	-0.05	0.19*	-0.28	0.13
Online.l5	0.24*	-0.19	0.20**	-0.05	0.15	0.04
Offline.l5	0.003	0.15**	0.08	0.25**	0.40	0.28*
Online.l6					0.19	0.04
Offline.l6					-0.40	-0.32
Online.l7					-0.11	0.10
Offline.l7					-0.44	0.54***
const	-2.58	-6.09	-11.41*	-5.31	0.47	-0.69
navidad	0.83*	22.08***	3.50***	12.84***	0.28	0.95**
cyber	3.35***	0.01	9.99***	-2.09*	3.28***	-0.15
Adjusted R ²	0.85	0.95	0.92	0.87	0.88	0.84
F Statistic	16.14***	50.52***	29.94***	18.62***	14.84***	11.09***

Note:

*p<0.1; **p<0.05; ***p<0.01

Tabla 51: Estimación modelos por categorías de producto para clientes premium. Fuente: Elaboración propia

10.6.2 Clientes no premium

Ventas clientes no premium por categoria

	Variable dependiente:					
	Modelo vestuario		Clientes premium Modelo electrónica		Modelo decoración	
	Online	Offline	Online	Offline	Online	Offline
Online.11	0.26***	0.36	0.08	0.06	0.04	-0.08
Offline.11	-0.02*	0.01	-0.01	0.08	0.06	0.002
Online.12	0.28***	-0.002	0.12*	0.01	0.06	-0.13**
Offline.12	-0.01	0.12**	-0.01	0.11	-0.01	-0.08
Online.13	0.10	-0.55	0.08	-0.06	0.13	-0.14**
Offline.13	0.0003	0.06	0.08	0.30***	0.31**	0.04
Online.14	-0.01	-0.19	0.08	-0.13	-0.01	-0.15**
Offline.14	-0.02	0.04	-0.04	0.15**	-0.27*	0.05
Online.15	0.33**	-0.16	0.22***	-0.003	0.33***	-0.12*
Offline.15	-0.02	0.12**	0.04	0.23***	0.08	0.03
Online.16					0.32*	0.03
Offline.16					-0.32	-0.03
const	-2.18*	-14.81***	-15.39**	-14.38	-1.40	1.72
navidad	0.87**	32.05***	2.18***	9.79***	0.18	1.54***
cyber	3.04***	-0.32	4.88***	-1.07*	1.68***	-0.05
tiendas			0.20	0.15		
Adjusted R ²	0.86	0.97	0.93	0.93	0.89	0.88
F Statistic	17.30***	76.03***	31.76***	36.32***	19.37***	17.60***

Note:

*p<0.1; **p<0.05; ***p<0.01

Tabla 52: Estimación modelos por categorías de producto para clientes no premium. Fuente: Elaboración propia

10.6.3 Categoría vestuario por zonas

Ventas vestuario: comparación zonas

	Variable dependiente:							
	Zona norte		Zona oriente		Zona poniente		Zona sur	
	Online	Offline	Online	Offline	Online	Offline	Online	Offline
Online.11	0.26**	0.34	0.29***	0.64***	0.27**	0.29	0.22**	0.30
Offline.11	-0.02	-0.10**	-0.03	0.0002	-0.03	-0.04	-0.02	0.01
Online.12	0.28**	0.17	0.29***	0.10	0.29**	-0.22	0.28**	-0.03
Offline.12	-0.01	0.06	-0.03	-0.06	-0.02	0.04	-0.02	0.15**
Online.13	0.10	-0.37	0.12	-0.50**	0.09	-0.58***	0.13	-0.67*
Offline.13	0.01	0.02	0.004	0.04	0.002	-0.02	-0.004	0.12*
Online.14	-0.005	-0.40	0.01	-0.42*	-0.02	-0.31	-0.02	-0.45
Offline.14	-0.01	0.07	-0.001	0.10	-0.02	0.04	-0.02	0.08
Online.15	0.30**	-0.46	0.23*	-0.10	0.27*	-0.21	0.27**	-0.05
Offline.15	-0.01	0.12***	0.01	0.05	-0.02	0.11***	-0.01	0.14**
const	-1.06	5.87*	-1.13	3.13*	-2.28	8.50	-0.72	-6.14**
navidad	0.29*	11.57***	0.30***	4.46***	0.77*	21.20***	0.35	13.56***
cyber	1.04***	-0.39	0.79***	0.40*	2.76***	-0.44	1.54***	-0.29
tiendas	0.01	-0.69***	0.01	-0.34***	0.01	-1.05***		
Adjusted R ²	0.85	0.98	0.86	0.96	0.84	0.98	0.85	0.94
F Statistic	14.78***	115.15***	16.70***	56.88***	14.08***	137.41***	16.11***	42.99***

Note:

*p<0.1; **p<0.05; ***p<0.01

Tabla 53: Estimación modelos para categoría vestuario por zonas. Fuente: Elaboración propia

10.6.4 Categoría electrónica por zonas

Ventas electrónica: comparación zonas

	<i>Variable dependiente:</i>							
	Zona norte		Zona oriente		Zona poniente		Zona sur	
	Online	Offline	Online	Offline	Online	Offline	Online	Offline
Offline.l1	0.05	0.13**	-0.33	0.11	-0.07	-0.03	-0.07	0.18*
Online.l1	-0.05	-0.02	0.14	-0.06	0.08	-0.01	0.13	-0.002
Offline.l2	-0.03	0.19***	-0.40***	-0.03	-0.09	-0.02	-0.04	0.16*
Online.l2	0.01	0.01	0.18*	-0.12*	0.12*	-0.09	0.23**	-0.14
Offline.l3	-0.005	0.08	0.09	0.17	-0.01	0.18*	0.06	0.31***
Online.l3	0.21***	0.08	0.13	-0.11	0.06	-0.11	0.16	-0.12
Offline.l4	-0.11	0.12**			-0.07	0.08		
Online.l4	0.15**	-0.03			0.15**	-0.18**		
Offline.l5	-0.03	0.28***			0.04	0.19*		
Online.l5	0.20***	0.02			0.15*	-0.07		
const	-1.96	-3.91***	1.02	-0.13	-5.51	1.12	0.08	-5.54
navidad	4.91***	1.30***	0.37	1.37***	2.48***	9.38***	0.78	6.35***
cyber	-0.61**	2.38***	1.80***	-0.06	6.96***	-1.56*	3.09***	-0.15
tiendas			-0.08	0.07			-0.20	0.18
Adjusted R ²	0.95	0.93	0.79	0.64	0.91	0.89	0.70	0.84
F Statistic	48.41***	36.19***	15.04***	7.86***	29.55***	22.92***	9.66***	20.18***

Note:

*p<0.1; **p<0.05; ***p<0.01

Tabla 54: Estimación modelos para categoría electrónica por zonas. Fuente: Elaboración propia

10.6.5 Categoría decoración por zonas

Ventas decoración: comparación zonas

	<i>Variable dependiente:</i>							
	Zona norte		Zona oriente		Zona poniente		Zona sur	
	Online	Offline	Online	Offline	Online	Offline	Online	Offline
Online.l1	0.01	-0.06	-0.01	-0.07	-0.06	-0.09	-0.02	-0.10
Offline.l1	-0.25	-0.04	-0.20	0.11	-0.39*	-0.04	-0.11	0.04
Online.l2	0.18	-0.15**	0.14	-0.09**	0.10	-0.14*	0.14	-0.15*
Offline.l2	-0.32	-0.19	-0.62***	-0.07	-0.34	-0.08	-0.21	-0.06
Online.l3	0.24**	-0.18***	0.27**	-0.15***	0.17	-0.20***	0.22*	-0.13
Offline.l3	0.15	-0.05	0.21	0.29***	0.03	0.05	0.08	0.11
Online.l4	0.06	-0.15**	-0.03	-0.13***	-0.06	-0.22***	0.04	-0.11
Offline.l4	-0.27	0.05	-0.21	0.03	-0.24	0.15	-0.19	0.15
Online.l5	0.18*	-0.11*			0.11	-0.17***	0.16	-0.10
Offline.l5	0.33	-0.09			0.38	0.09	0.18	0.15
const	-0.40	2.14***	0.19	0.37*	1.71	4.31**	-0.54	1.05
navidad	-0.01	0.37***	0.16	0.50***	-0.18	1.06***	-0.01	0.79***
cyber	1.04***	-0.01	0.81***	0.03	1.83***	-0.12	1.47***	-0.10
tiendas	-0.004	-0.05			-0.07	-0.16*		
Adjusted R ²	0.84	0.70	0.74	0.85	0.84	0.84	0.81	0.79
F Statistic	13.63***	6.69***	10.16***	19.37***	13.68***	13.82***	12.07***	11.14***

Note:

*p<0.1; **p<0.05; ***p<0.01

Tabla 55: Estimación modelos para categoría decoración por zonas. Fuente: Elaboración propia