



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

METODOLOGÍA PARA ESTIMAR EL IMPACTO DE LA ATENCIÓN DE
VENEDORES SOBRE EL COMPORTAMIENTO DE COMPRA DE LOS CLIENTES
EN UNA TIENDA DE RETAIL

MEMORIA PARA OPTAR AL TÍTULO DE INGENIERO CIVIL INDUSTRIAL

NICOLÁS ARTURO GARCÍAS CHANDÍA

PROFESOR GUÍA

MARCELO OLIVARES ACUÑA

MIEMBROS DE LA COMISIÓN

ANDRÉS MUSALEM SAID

ARIEL SCHILKRUT COGAN

SANTIAGO DE CHILE

ABRIL 2014

MEMORIA PARA OPTAR AL TÍTULO DE INGENIERO CIVIL INDUSTRIAL

POR: NICOLÁS GARCÍAS CHANDÍA

FECHA: 15/04/2014

PROF. GUÍA: SR. MARCELO OLIVARES

Se desarrolló un estudio para analizar cómo la atención por parte de los vendedores afecta el comportamiento de compra de los clientes en una tienda de retail. La metodología combina novedosos datos del proceso de compra del cliente (recolectados mediante una tecnología de reconocimiento de imágenes) y datos del punto de venta. Por medio de test de hipótesis se analizó exploratoriamente el efecto que tienen las acciones por parte del personal de ventas sobre la probabilidad de compra de los clientes. En este sentido las acciones que más se correlacionan positivamente con la probabilidad de compra son: estar disponible en el caso que algún cliente necesite ayuda y facilitar o persuadir a un cliente para que ingrese al probador. Adicionalmente, no existe evidencia estadística para validar que los vendedores escogen a clientes que tienen más tendencia a comprar. Sin embargo, este análisis no permite descartar que la correlación entre mayor incidencia de compra y atención de clientes es inducida por heterogeneidad no observada en la propensión de compra de los clientes. Para corregir este problema de variables omitidas en el modelo, se desarrolló un modelo econométrico basado en variables instrumentales que utiliza datos agregados de ventas y tasas de atención de clientes, y que además corrige por sesgo de selección y error de muestreo. Los resultados de este modelo sugieren que la atención por parte de los vendedores se correlaciona positivamente con la conversión (fracción de clientes que compra algún producto). Específicamente al aumentar la fracción de clientes atendidos en 11 % (una desviación estándar) sobre la media, la conversión aumenta un 4,5 %. Otra medida muy utilizada son las ventas por entrada de clientes, métrica muy poderosa ya que resume dos efectos importantes: porcentaje de clientes que compran y dólares promedio gastado por cada uno de ellos. Se concluye que las ventas por entrada de clientes aumenta en 7,4 dólares aumentando la misma fracción de clientes atendidos. Los resultados del modelo se utilizan para sugerir decisiones de dotación de personal que consideran el *trade-off* entre el costo salarial versus el incremento en venta inducido por mejoras en el nivel de servicio al cliente.

Agradecimientos

A mis padres por su apoyo incondicional, cariño y por todos los momentos de estrés que me soportaron en este período universitario. Han sido un pilar fundamental para mi desarrollo personal.

A mis hermanos por su apoyo y compañía en todo momento.

A mis amigos de la Universidad, por todos esos excelentes momentos vividos en la Universidad, infinitas horas de estudio, salidas, risas, locuras y por sobre todo amistad.

A mis amigos de la vida que siempre estuvieron ahí a pesar de la distancia.

A mis compañeros de balonmano, con los cuales se vivieron muchos momentos felices y tristes tanto fuera como dentro de la cancha.

A los profesores Marcelo y Andrés por su paciencia, compromiso, ayuda y por toda la enseñanza que me brindaron durante el desarrollo de esta tesis, sin su ayuda no lo hubiera logrado.

A mi jefe Ariel, Álvaro y mis compañeros de trabajo que me recibieron con los brazos abiertos y colaboraron conmigo cada vez que necesité de su ayuda.

A mi polola Cata, quién en este último tiempo ha sido fundamental en mi vida y gracias a ella estoy feliz nuevamente. Gracias por todo lo que me entregas y por acompañarme siempre. Te Amo.

Y por sobre todo a mí padre biológico, quién me apoya en todo momento desde el cielo...

Tabla de Contenido

1.	Introducción	8
2.	Descripción de Prácticas de Gestión de la Calidad de Servicio en Retail	11
3.	Revisión de la literatura y formulación de hipótesis	13
4.	Metodología de Recopilación de Datos	16
4.1.	Fuentes para la recopilación de datos	16
4.2.	Descripción de las tiendas en el estudio	18
5.	Definición de métricas para los modelos econométricos	19
5.1.	Definición de variables	19
5.2.	Controles	25
6.	Medición de impacto mediante videos de seguimiento completo	28
7.	Medición de impacto mediante videos parciales	33
7.1.	Especificación modelo econométrico	33
7.1.1.	Derivación modelo lineal	33
7.1.2.	Especificación modelo Logit	35
7.2.	Corrección de sesgo de selección	36
7.3.	Modelo para corregir Endogeneidad	38
7.4.	Resultados	41
8.	Medición de impacto mediante experimentos de campo	46
9.	Análisis de sensibilidad	47
10.	Implicancias para la gestión de la fuerza de venta	48

11. Conclusiones	50
12. Bibliografía	53
13. Anexos	55

Índice de cuadros

1.	Estadísticos Descriptivos vendedores, clientes y ratio	23
2.	Estadísticos descriptivos	24
3.	Estadísticos descriptivos ambas tiendas	25
4.	Estadísticos descriptivos variables obtenidas desde los videos de seguimiento completo	28
5.	<i>Ejemplo instrumento.</i> Fuente: Elaboración propia	39
6.	Modelo principal, utilizando VI	42
7.	Modelo secundario utilizando VI	43
8.	Modelo de conversión por tienda	45
9.	<i>Resultados del experimento realizado.</i> Fuente: Departamento de Marketing de la empresa en estudio y Scopix Solutions.	46
10.	Modelo Logit utilizando controles de día, feriados y clima	55
11.	Modelo Logit utilizando controles de tienda, cliente y número de semana	56
12.	Modelo Logit con y sin controles	57
13.	Modelo Logit cambiando la muestra para la tasa de atención	58
14.	Modelos utilizando variables instrumentales	59
15.	Modelos sin usar variables instrumentales	60

Índice de figuras

1. *Tráfico por cámara. El color azul indica el conteo de entrada de la cámara número uno y el rojo el conteo de la cámara número dos, en tanto el color verde es la suma de los anteriores. Es posible distinguir que desde el mes de Septiembre hasta Diciembre la cámara dos tuvo una des-conexión entregando así conteo cero para ese período. Fuente: Elaboración Propia 21*
2. *Clientes, vendedores y ratio por día de semana para las dos tiendas. Fuente: Datos Scopix Solutions, Gráfico elaboración propia. 24*
3. *Distribución tiempos de espera antes del contacto entre vendedor y cliente. Fuente: Elaboración propia. 26*
4. *Trayectoria de los clientes en el interior de la tienda. Tanto el camino como las fases que aparecen se construyeron por la observación de videos de seguimiento completo. Fuente: Elaboración propia 29*
5. *Sesgo selección. Fuente: Elaboración propia 37*
6. *Error estándar versus número de muestras. Fuente: Elaboración propia. 61*
7. *Modelo de Logit de la atención versus tasa de atención. No hay evidencia para rechazar linealidad. Fuente: Elaboración propia. 62*
8. *Resultados para la estimación de distintos criterios: Efecto de compromiso en la conversión. Fuente: Elaboración propia. 62*
9. *Tabla resumen con los valores que toma el parámetro de la tasa de atención al cambiar los controles en el modelo. Fuente: Elaboración propia. 63*

1. Introducción

En los últimos años los profesionales del retail han desarrollado herramientas cuantitativas que integran datos del punto de venta (POS), tarjetas de fidelización e investigación de mercados para entender mejor cómo los consumidores reaccionan a diferentes niveles de precios o distintas estrategias de promociones, siendo posible obtener abundante información de este tipo. De los datos del punto de venta se obtiene información de lo que compran los clientes, pero lamentablemente no sirve como fuente de información para obtener datos de las intervenciones realizadas por parte del personal de ventas en el proceso de compra del cliente. La información que se maneja sobre los niveles de servicio en tienda¹ a los que son expuestos los clientes en el proceso de compra es escasa, de esta forma se está perdiendo la oportunidad de saber qué impulsó al cliente a realizar o no la compra.

Incluso recopilando información del proceso de compra de los clientes, las empresas de retail no saben quiénes son sus compradores, lo que implica que no se pueden concentrar en la tasa de conversión ni intercepción (tasa de atención), es decir, porcentaje de clientes quienes tienen contacto con algún empleado (Underhill 1999). En esta misma línea las tiendas deben considerar los tres factores más importantes relacionados con sus clientes, en primer lugar, el tiempo que los compradores pasan en una tienda, que sirve como un indicador de cuanto comprarán. En segundo lugar, un aumento en la tasa de atención se relaciona con un aumento en calidad de servicio y disminución en la tasa de mermas. Por último, el tiempo de espera afecta negativamente en la satisfacción de los clientes, es decir, cuando los tiempos de espera en la fila son elevados los clientes retienen este hecho negativo, pudiendo dañar la reputación de la tienda por medio de comentarios perjudiciales sobre esta.

Todos los clientes que llegan a una tienda lo hacen con un deseo común: encontrar un producto que cumpla con sus expectativas. En la búsqueda de este propósito los clientes pueden tener dificultades para encontrar lo que desean, siendo en este momento donde el vendedor toma protagonismo, pudiendo convertirse en un factor importante en la experiencia final de compra del cliente. Con respecto a lo anterior, los vendedores deben ser capaces de entender a cada tipo de cliente y ayudarlos con sus requerimientos, de manera de ser capaces de captar consumidores que sean rentables, es decir, quienes generan más ganancias para la empresa. Las empresas se están centrando cada vez más en la construcción a largo plazo de relaciones con clientes clave, en donde los vendedores son a menudo responsables de esta relación (Bradford et al. 2010).

Existen métodos para incentivar a los vendedores, por ejemplo comisiones de acuerdo al nivel de venta que en ocasiones pueden ser contraproducentes para la tienda. Dado lo anterior, los

¹ Todas las acciones por parte del personal que ayudan al cliente a tener una buena experiencia de compra.

vendedores pueden esforzarse en maximizar sus ventas, dejando de lado la calidad de servicio, medida que se correlaciona positivamente con la satisfacción de sus clientes. En esta línea existen diversos estudios (Graham y Harvey 2005) que demuestran que uno de los elementos que más valoran los clientes es la atención cortés, es decir, la cortesía es esencial para entregar una buena calidad de servicio. Un área importante de estudio es la relevancia y el rol que tienen los vendedores en el proceso de compra del cliente. El manejo del personal no es un tema menor ya que los gastos asociados son altísimos, por ejemplo solo en el 2006 las cadenas de retail gastaron 393 billones de dólares en salarios para empleados², cifra que representó más del 10 % de los ingresos de ese año, por lo tanto, el manejo de la fuerza de ventas es clave en la rentabilidad de la tienda. Con este fin, el objetivo principal de esta investigación es estudiar el rol que tienen los vendedores y cómo su interacción con el cliente influye en las ventas, lo que puede ser utilizado como base para apoyar las decisiones relacionadas con la formación y dotación del personal.

En esta investigación, se desarrolló una novedosa técnica de recolección de datos basada en análisis de videos que permiten hacer seguimiento de las interacciones entre clientes y empleados. Las cámaras de la tienda se utilizan para grabar videos de una muestra de clientes que compran en la tiendas, codificando estos registros mediante una combinación de algoritmos de reconocimiento de imagen y la inspección humana, para así generar datos acerca de la experiencia del cliente en la tienda. Las muestras consideran distintos tipos de datos relacionados a su visita, tales como: contacto del cliente con el personal de ventas, visita al probador, el tiempo que el cliente se demora en comprar o el tiempo de permanencia en el local.

El propósito de este análisis es entender el mecanismo por el cual la participación del vendedor impulsa las ventas, pudiendo ser útil para diseñar estándares de formación de los empleados y la conformación de la dotación de personal. Si bien la información que se puede adquirir sobre la experiencia de compra es valiosa, su recopilación es costosa ya que se requiere inspección humana para analizar los videos y por lo tanto puede ser difícil implementar un estudio a gran escala que involucre múltiples tiendas. Por esta razón, se recolectaron también videos de corta duración, de los cuales se puede obtener una muestra de los clientes durante los primeros minutos de una visita, registrando el tiempo transcurrido hasta que hubo contacto con un agente de ventas. El beneficio de utilizar estos videos es que se puede recolectar una gran cantidad de muestras, haciendo posible escalar el análisis a otros locales de la cadena, y así poder realizar comparaciones. Por el contrario, la desventaja de este material es que no es posible medir directamente si un cliente en la muestra realizó la compra, ya que el tiempo de duración del video es limitado en este caso.

² Fuente: Bureau of Labor Statistics Quarterly Census Employment and Wages for NAICS codes 44-45 (Retail Trade).

La contribución clave de este estudio es desarrollar una metodología que vincule los datos de video con la información del punto de venta, estimando el efecto de la atención en las ventas. Lo anterior implica estimar la probabilidad de compra del cliente utilizando como métrica la conversión, que combina datos del punto de ventas, en particular el número de transacciones con datos de tráfico en tienda. Este estadístico entrega información sobre el número de personas que compran con respecto a la totalidad de personas que ingresaron a la tienda. Otros datos recogidos con alta frecuencia incluye el tráfico en tiendas, el número de clientes que compran y el número de asociados de ventas. Finalmente se desarrolló un modelo econométrico para estimar la probabilidad de compra por medio de la conversión en función de la tasa de atención, controlando por clientes en tienda, clima, feriados y otros controles de estacionalidad.

Un reto importante en este análisis es dar cuenta de algunos problemas con las variables y la especificación del modelo econométrico utilizado. El primer problema está asociado con la medición de la tasa de atención a nivel desagregado, variable que no dispone de algunos registros, los cuales al no ser considerados se podría estar sesgando la selección de la muestra. El problema más desafiante que se aborda en este trabajo es un posible sesgo de endogeneidad que presenta la tasa de atención o contacto entre el vendedor con el cliente. Cuando los clientes tienen intenciones de comprar, pueden buscar contactar a un vendedor o los vendedores pueden escoger atender a clientes que presentan una alta probabilidad de compra (si fuera posible identificarlos), como consecuencia no existe un efecto claro de esta variable en la probabilidad de compra. El cliente puede tener características no observables que determinan la intención de compra (por ejemplo, un cliente interesado en la compra puede ser más propenso a buscar ayuda como se comentó anteriormente), siendo no necesariamente un efecto causal. Para mitigar este problema de endogeneidad, se propone el uso de variables instrumentales, postulando como candidato el número de clientes sobre el número de vendedores. Este instrumento se correlaciona negativamente con la tasa de atención, específicamente a medida que el tráfico en la tienda aumenta esta medida también lo hace, sin embargo los vendedores se pueden dedicar a una fracción menor de clientes, es decir, el porcentaje de clientes atendidos disminuye (el análisis es el mismo en el sentido contrario). Inclusive esta variación del instrumento es exógena a las características no observables de los clientes (por controles de estacionalidad y otros factores observables), cumpliendo así dos condiciones necesarias para ser un instrumento, ser relevante y exógeno o válido.

El análisis exploratorio arrojó como resultado que las acciones que más se correlacionan (positivamente) con la probabilidad de compra son: estar disponible en el caso que algún cliente necesite ayuda y facilitar o persuadir a un cliente para que ingrese al probador. Adicionalmente, no existe evidencia estadística para validar que los vendedores escogen a clientes que tienen más tendencia a comprar. Sin embargo, este análisis no permite determinar la causalidad del

efecto de la atención sobre la intención de compra. Utilizando un modelo econométrico incluyendo estimaciones con variables instrumentales, se determinó que la tasa de atención tiene un efecto positivo sobre la conversión, específicamente al aumentar la fracción de clientes atendidos en 11 % (una desviación estándar) sobre la media, la conversión aumenta un 4,5 %. Otra medida muy utilizada son las ventas por entrada de clientes, métrica muy poderosa ya que resume dos efectos importantes: porcentaje de clientes que compran y dólares promedio gastado por cada uno de ellos. Se concluye que las ventas por entrada de clientes aumentan en 7,4 dólares al aumentar la misma fracción de clientes atendidos.

A continuación se presentan algunas prácticas de gestión en el retail, para luego hacer una revisión de la literatura y la postulación de algunas hipótesis que relacionan la atención con la probabilidad de compra. Posteriormente, por medio de un análisis descriptivo se buscó una idea previa del efecto de la atención sobre la conversión para más adelante testear las hipótesis de una forma más rigurosa mediante un modelo econométrico, incorporando una solución para los problemas de error de medición, sesgo de selección y de endogeneidad. Por último, se realizó un experimento de campo y un análisis económico para validar los resultados obtenidos.

2. Descripción de Prácticas de Gestión de la Calidad de Servicio en Retail

En la competitiva industria del retail, muchos gerentes de local consideran que la experiencia en la tienda es crítica para convertir tráfico en ventas y futuras visitas. Para tener una experiencia de compra superior, no sólo se requiere inventario en la tienda, sino que también una fuerza de ventas que sea experta en asegurar una eficiente y agradable visita para los clientes. En este sentido, el manejo de la fuerza de ventas tiene características similares a lo largo de la mayoría de las tiendas de retail. Se tienen cuatro similitudes, en primer lugar las tiendas necesitan suficiente personal para asegurar un servicio adecuado al cliente, lo que conlleva tratar directamente con los clientes durante el proceso de compra. En segundo lugar, el personal de ventas debe encargarse de manejar el inventario en la tienda. En tercer lugar, el personal de ventas es requerido para mostrarle al cliente los lugares de interés dentro de la tienda y por último los vendedores son utilizados también como cajeros. Estas cuatro funciones son las principales tareas que cumple el personal de ventas en una tienda de retail, por lo que para efectos de esta tesis se analizarán aquellas que estén relacionadas con la interacción entre clientes y vendedores. En cuanto a los contratos de los empleados de retail, se definen tres categorías: Trabajadores a tiempo completo, trabajadores a tiempo parcial y trabajadores de temporada. Los trabajadores a tiempo completo se encuentran durante todo

el año, quienes tienen horas fijas definidas que se extienden por el horario de atención de la tienda, mientras que los trabajadores a tiempo parcial se diferencian con los anteriores porque trabajan menos horas. Por último se tiene a los trabajadores de temporada, quienes trabajan por un corto período y son utilizados para manejar estacionalidad, principalmente cuando el tráfico aumenta bruscamente.

La planificación de la fuerza de ventas involucra determinar la cantidad correcta de trabajadores a tiempo completo, parcial y de temporada, para asignarlos de acuerdo a las horas predichas en un determinado período. En cuanto a la planificación del personal (tarea desarrollada por el gerente de tienda), existen diferencias considerables entre pequeñas y grandes tiendas. Esto ya que en pequeñas tiendas esta decisión es tomada sin el soporte de alguna herramienta, mientras que las cadenas más grandes basan su planificación en modelos matemáticos, junto a millonarias inversiones en distintas herramientas utilizadas para la gestión de personal. A pesar de utilizar estas herramientas pueden existir limitaciones en la óptima asignación, ya que los departamentos dentro de la cadena tienen distintos propósitos. Por ejemplo, el departamento de finanzas tiene como objetivo controlar los gastos de personal, lo que va en contra de otros departamentos como el de marketing, el cual buscará tener el suficiente personal para cumplir con algún estándar de calidad de servicio.

Por lo general, las grandes tiendas de retail presentan un complejo sistema para planificar la dotación de personal, partiendo por un gerente por departamento quien es el responsable por el manejo de trabajadores en su área, tomando las decisiones basadas en las ventas. Los gerentes de departamento reportan al asistente de gerencia y éste al gerente de local. Este último tiene un equipo liderado por el gerente de recursos humanos, quien tiene un rol vital durante el período más alto de ventas, ya que debe ser capaz de contratar a una gran cantidad de trabajadores para dicha época. Por lo general las asignaciones se planifican tanto en el largo como en corto plazo. La asignación a largo plazo se revisa en el comienzo de cada año fiscal, determinando la cantidad mínima de horas del personal requerido en cada tienda, con lo cual el gerente de local puede determinar la proporción de vendedores a tiempo completo y parcial en cada cargo. La planificación a corto plazo (mensual) involucra decisiones del correcto número de trabajadores a tiempo completo y parcial que se asignarán por medio de modelos matemáticos (principalmente regresiones) y programas de gestión. En los modelos se utiliza como principal variable la predicción de ventas, que permite estimar el personal requerido con buenos resultados, mientras que por medio del programa de gestión se ve la disponibilidad de trabajadores (restringidos muchas veces por el personal mínimo requerido por la corporación). Una limitación que puede tener esta asignación es la poca flexibilidad que tienen los empleados a tiempo completo, teniendo más flexibilidad si se asignan trabajadores de tiempo parcial. La mayor flexibilidad se alcanza al utilizar trabajadores de temporada, quienes son asignados solo en fechas especiales donde la demanda crece significativamente. Los temas de reclutamiento

y capacitación son muy importantes, donde la gerencia de recursos humanos debe describir perfiles idóneos para cada cargo. Luego de esto, se entrevistan a varios candidatos por puesto (quienes deben pasar por numerosas pruebas), donde finalmente pueden pasar meses para una decisión final. Luego de que el trabajador es contratado viene la fase de entrenamiento, en la cual numerosas veces los gerentes de local toman esta responsabilidad (Fisher y Krishnan 2005), mientras que en otros casos se traspasa a los trabajadores de tiempo completo.

Ya desde varios años tecnologías ligadas al punto de ventas y manejo de dotación de personal han sido incorporados por numerosas cadenas de retail, e incluso en los últimos años se ha empezado a combinar con recientes tecnologías muy novedosas, como contadores de tráfico y sensores de movimiento, que capturan las acciones de los clientes dentro de la tienda. Otro tipo de tecnología utilizada en el último tiempo es el monitoreo al interior de la tienda en tiempo real, que a través de una plataforma web se muestra la actividad en la tienda, brindando alertas oportunas para la toma de decisiones. Un ejemplo de aplicación en tiempo real es el cálculo de promedio de largos de cola, tiempos de espera y tráfico en línea, que por medio de un modelo matemático permite a los gerentes de local tener recomendaciones de la cantidad de cajas abiertas que son necesarias para tener un nivel de servicio óptimo (determinado largo de cola o tiempo de espera). Un desafío es seguir buscando patrones entre los datos del punto de venta y las métricas medidas en tiempo real, logrando de esta forma mejorar la conversión y ventas.

3. Revisión de la literatura y formulación de hipótesis

Se han realizado diversos estudios empíricos que examinan los efectos de la dotación de personal en el servicio entregado y el impacto que tiene éste en las ventas. Es así como buscar el aumento en la efectividad del personal de ventas durante una interacción con el cliente es una tarea importante para considerar en la gestión de la tienda (Harish, Barton y Mita 1988). Estos autores van más allá y especifican que tener desarrollado un personal de ventas con conocimiento, buena motivación y una efectiva adaptación, se puede traducir en ganancias a largo plazo para la empresa. La efectiva adaptación conlleva a que los vendedores deban corregir sus errores y aprender de las buenas prácticas.

En los últimos años se ha vuelto muy popular la medida de conversión (en conjunto con el manejo del personal de ventas), la cual se define como la capacidad de convertir tráfico en transacciones (Conroy and Bearse 2006). La forma exacta para calcular esta medida es tomar el total de compras y dividirlo por el número total de clientes que ingresaron a la tienda. La importancia de esta medida radica en su positiva correlación con las ventas, lo que lleva a los

gerentes de local a buscar un constante aumento de ésta. En el área de marketing, la lealtad de marca está asociada a afirmaciones tales como: reducción de niveles de incertidumbre en la compra; reducción de costos en la búsqueda de información asociados a cualquier decisión de consumo; reducción de la sensibilidad en los compradores a los cambios en el precio; aumento en la tolerancia de los consumidores a variaciones en la calidad y resistencia de los clientes fieles a las promociones de los competidores (Assael, 1999; Schiffman y Lazar, 2001; Hoffman y Bateson, 2002). Del mismo modo, la lealtad de marca se asocia con niveles superiores de satisfacción del cliente (Lambin, 2004), y a su vez la satisfacción está asociada con elementos como nivel de desempeño, calidad, precio, reconocimiento e imagen (Aaker, 1996). Un cliente al ser leal a una marca tiende a recomendar un producto o marca que constantemente compra (mostrando un compromiso con la empresa), afectando positivamente en la rentabilidad de la compañía. Un incremento en la conversión está asociado con un aumento en la lealtad de los clientes y mejoras en la calidad de servicio (Conroy and Bearse 2006). Estos últimos también hacen referencia a los distintos niveles de conversión, dependiendo de la industria o sector. Por ejemplo, en la industria de Internet se observa la conversión más baja de todos los sectores tendiendo a 0 %, en contraste con un almacén de barrio donde esta medida llega casi al 100 %. Las tiendas de esta tesis se clasifican como tiendas por departamento y cuentan con un nivel bajo de conversión, llegando a un 27 %, tal como lo especifica el artículo. Siguiendo en esta misma línea (Ton 2009), el aumento de la dotación de personal se asocia a un aumento en márgenes de ganancia, efecto explicado por la conformidad de los clientes con el servicio recibido y no por la diferencia entre las expectativas de la calidad de servicio que esperan recibir y el recibido. Específicamente la autora encuentra que al aumentar en una desviación estándar la fuerza de ventas, aumenta en un 10 % los márgenes de beneficios.

Otros estudios (concentrados en la dotación de personal) demuestran que un aumento de personal en tienda se relaciona con un mejor servicio prestado, lo que repercute finalmente en las ventas. En particular existe una gran oportunidad para aumentar los ingresos con el solo hecho de planificar la dotación de personal en función de los niveles de tráfico, asignando la cantidad adecuada de personal en el lugar y tiempo correcto (Fisher, Krishnan y Netessine 2009). Del mismo modo, es más probable tener ventas más altas cuando los productos se encuentran bien ubicados (Ton y Raman 2010) y cuando el personal de ventas está disponible (Fisher et al. 2006). En general, lo que suelen hacer los encargados de local es planificar la dotación de personal por períodos extensos, lo que trae consigo una falta de personal en los días y horas de mayor tráfico y con ello un impacto negativo en el desempeño de la tienda (Mani, Ksavan, Swaminathan y Jayashankar 2011). La tasa de atención (porcentaje de clientes atendidos por parte del personal de ventas) y conversión se utilizan para plantear la siguiente hipótesis a validar:

Hipótesis 1.a: *El aumento de la tasa de atención implica un aumento en la conversión. Se*

postula también: 1.b El aumento en la tasa de atención implica un aumento en las ventas por entrada de clientes (Ventas totales dividido por entrada total).

De esta forma, se buscará analizar el impacto que tendría un aumento de las personas atendidas en la tienda. Lo anterior presenta una oportunidad para identificar los periodos claves donde se pueda mejorar el nivel de servicio, aumentando así las ventas. Es así como en un trabajo muy reciente (Perdikaki, Kesavan y Swaminathan 2012) demuestra que ante un alza en el tráfico de la tienda existe un aumento promedio de las ventas, inclusive este aumento es mayor cuando se aumenta al personal de ventas en conjunto con el tráfico. De esta relación los autores encuentran que el volumen de ventas es una función cóncava creciente del tráfico, específicamente teniendo el personal de ventas en el promedio y aumentando en una desviación estándar el tráfico sobre la media, los volumen de ventas aumenta en 8,14 dólares. Finalmente los autores encuentran que el aumento en la tasa de conversión se asocia con un incremento en el tráfico futuro. En otro trabajo (Mani et al. 2013) usan detallados datos de tráfico para investigar si las tiendas tienen falta de personal, y el posible impacto de lo anterior en las ventas. Específicamente, para una muestra de días entre lunes y viernes encuentran que las tiendas presentan falta de personal el 40,21 % del tiempo, llegando a un 64,98 % en las horas puntas. Finalmente determinaron que en promedio la pérdida de ventas llega a un 8,56 % en las horas puntas (muestra de 41 tiendas). Una métrica de desempeño que permite medir la congestión en tienda es el ratio entre clientes y vendedores, medida que al tener un valor alto, implica que cada vendedor debe ser capaz de atender un mayor número de clientes, disminuyendo así el porcentaje de clientes atendidos. Se postula la siguiente hipótesis a verificar:

Hipótesis 2: El aumento de la dotación de personal impacta positivamente en la conversión.

(Netessine et al. 2010) y (Ton 2009) concluyen que la mayoría de las tiendas tienden a tener falta de personal. Incluso (Fisher y Raman 2010) concluyen que la mayoría de los gerentes ven el manejo de personal como un costo y no como un activo. Incluso argumentan que existe dificultad para cuantificar los beneficios en el futuro por tener una mejor dotación de personal. Se quiere estudiar si existen diferencias entre las tiendas, específicamente:

Hipótesis 3: Existen diferencias en la efectividad de la tasa de atención sobre el nivel de conversión entre las tiendas.

Por último se medirá cuantitativamente el impacto económico que tendría para la tienda agregar un vendedor adicional de tiempo completo o parcial, buscando validar la tercera hipótesis:

Hipótesis 4: *Existe oportunidad para asignar personal adicional, aumentando las utilidades de la tienda.*

4. Metodología de Recopilación de Datos

Entender correctamente el proceso de recolección de los datos y el significado que tiene cada una de las variables (contexto en el cual se midieron), es fundamental para poder realizar conclusiones que sean válidas y correctas. En el presente trabajo se utilizaron videos provenientes de cámaras instaladas en el interior de la tienda. De las cámaras es posible obtener registros de videos con una duración definida previamente, dependiendo del tipo de métrica.³

A continuación se describen las fuentes para la recolección y el tipo de datos que se pueden obtener, para luego mostrar dos ejemplos en distintas cadenas de retail.

4.1. Fuentes para la recopilación de datos

Se obtuvieron registros de videos cada 15 minutos durante el horario de atención de la tienda, siendo posible obtener una o más métricas por cada registro de video. Se definen cuatro tipos de fuentes para la recopilación de los datos:

- videos de corta duración. Estos poseen una duración de veinte segundos. Este material permite la medición de métricas del tipo conteo de personas, específicamente se utilizó para contar el número de clientes y vendedores presentes en la sala de ventas (área de la tienda donde pueden interactuar clientes y vendedores). Para llevar a cabo la medición, se fijó la imagen en los diez segundos del video, permitiendo un desfase de algunos segundos en el caso de que la imagen fuera difusa en el instante de medición. Para que el conteo se considere válido se deben identificar de manera clara a clientes o vendedores en el caso que corresponda. A diferencia de (Perdikaki et al. 2012), los datos asociados al número de vendedores disponibles difieren, ya que en esta tesis se utilizó la cantidad de vendedores que efectivamente ejercían sus actividades en la sala

³ Se define métrica como la variable que se quiere medir. Estas pueden ser tiempo de espera, conteo de clientes o vendedores, servicio al cliente o quiebre de stock. Desde que se recopila el registro de video hasta que se obtiene el valor numérico de la métrica es un proceso complejo en donde se utilizó un software para obtener el valor final de la métrica más fácilmente. Estos valores pasan por un proceso de calidad, por el cual se revisó el 10 % de la muestra para detectar errores en la evaluación, corrigiendo en caso de ocurrencia. Para minimizar estos errores se entrenó a las personas que asignan valores a las métricas por medio de reglas de evaluación asociadas a cada una de estas.

de ventas en un horario determinado, es decir, se deja fuera del conteo a vendedores que se encontraban en colación, bodega o dejaron sus labores por algún motivo. En el caso que existiera un grupo de personas, se contabilizó como un cliente si se pudo identificar de manera clara que las demás personas eran acompañantes (sin intención de compra), en caso de existir dudas se contabilizó por separado. Con respecto a los vendedores, se contabilizó de manera individual aunque estuvieran en grupos, ya que el trabajo de vendedor es independiente al resto. Es importante indicar que en la tienda de estudio los vendedores cumplen también el rol de cajeros, por lo tanto no se contabilizó como vendedor si éste se encuentra en el área delimitada de las cajas.

- videos de seguimiento parcial. Estos registros permiten medir métricas que dependan del tiempo, como por ejemplo si existe contacto entre cliente y vendedor, teniendo como máximo 8 minutos para evaluar esta situación. Un cliente se consideró como atendido si existió contacto visual con el vendedor (a una distancia prudente), pudiéndose observar claramente una conversación entre ambos. Recordar que aunque el vendedor atienda, los clientes no siempre quieren comprar algo, pero en este caso el vendedor ya cumplió su función. El criterio para seleccionar a una persona en el video fue: cliente permanece fijo en algún sector de la tienda al menos 30 segundos o es atendido antes de este tiempo. Se evaluó al primer cliente que cumpliera con alguno de los requisitos recién mencionados, teniendo en cuenta que el cliente seleccionado no podía estar en el interior de la tienda al principio del video, ya que no es posible saber si fue atendido antes o no. Se distinguirán tres posibles resultados al evaluar esta métrica:

1. Cliente es atendido. Existe contacto entre el vendedor y cliente antes de que finalice el video.
2. Cliente no es atendido. Cliente se retira de la tienda antes de que finalice el video sin haber tenido contacto con algún vendedor.
3. Finaliza video y no ocurre ninguna de las dos condiciones anteriores, es decir, al finalizar el video el cliente permanece en la tienda sin haber tenido contacto. En esta situación la métrica tomará el valor de no observado. En el caso de que a la métrica se le asigne esta condición, no es posible en un principio agregar esta evaluación a la muestra, por el solo hecho de no tener un resultado válido. El problema de no incluirlos implica un sesgo de selección en la muestra, ya que se estaría seleccionando a los clientes dependiendo de la evaluación. Para ejemplificar, supongamos que dos clientes (en videos distintos) aparecen en el mismo instante del video. El primero es atendido, mientras que en el segundo cliente permanece en la tienda al finalizar el video, sin saber qué ocurrió con él (no fue atendido ni se retiró de la tienda). En la muestra se tendría que considerar

solo al cliente del primer caso, y como consecuencia de lo anterior se tendría un sesgo de selección en la muestra. Para corregir este problema se aplicó una metodología (se explicará más adelante) para incluir estos registros en la muestra final.

- Sensor para Conteo instantáneo. Estos sensores van integrados en las cámaras, las cuales a su vez van instaladas en cada entrada de la tienda. Estos sensores detectan cada vez que un cliente ingresa a la tienda (movimiento y dirección de las personas), registrando el total de clientes que visitan la tienda. Mas específicamente se obtiene el número de personas que ingresaron a la tienda, con intervalos de 15 minutos entre cada dato, es decir, si el registro obtenido corresponde por ejemplo a las 10:15, esto equivale a la entrada de personas entre las 10:00:00 y las 10:14:59.
- video de seguimiento completo. Estos registros permiten medir métricas que dependan del tiempo, pero con un límite de 60 minutos. Este material se utilizó para hacer un seguimiento completo del cliente, siendo posible observar desde el momento que ingresa a la tienda hasta que se retira, y así poder medir las interacciones que ocurren en el proceso de compra, poniendo especial énfasis en la intervención realizada por parte de los vendedores. Obtener estos videos es muy costoso, por ello se limitará el uso de esta fuente de recopilación de datos.

4.2. Descripción de las tiendas en el estudio

A continuación se mostrarán dos ejemplos que aplican para este tipo de datos, los cuales se utilizan en los análisis de esta tesis. En primer lugar se midió en dos tiendas de retail pertenecientes a la misma cadena, con especialización en vestimenta femenina casual (incluyendo todo tipo de vestidos, faldas, joyas y accesorios), teniendo como segmento objetivo a mujeres entre 20 y 40 años. Esta cadena de retail opera con cientos de tiendas en 37 estados en los Estados Unidos, Puerto Rico, the U.S. Virgin Islands y Canadá. Las tiendas tienen el mismo sistema de atención, es decir, utilizan entre 2 a 4 vendedoras, quienes visten con prendas adquiridas en el mismo local y cumplen tanto el rol de vendedoras como cajeras a la vez. La tienda se encuentra dividida por zonas claramente identificables, tales como: sector de cajas, sector de ventas, mesón de promociones y sector de probadores. En esta tienda se utilizaron videos de corta duración, seguimiento parcial, seguimiento completo y sensor de conteo instantáneo. Con este material se midió la cantidad de clientes y vendedores en la sala de ventas, largos de cola y tiempo de espera en el sector de cajas, conteo de entrada de clientes y si un cliente fue atendido o no. En segundo lugar, se midió en una tienda de retail con más de 400 locales en Canadá, con operaciones en el rubro de automóviles, herramientas y

ferretería, artículos deportivos y artículos para el hogar, entre otros. Esta tienda tiene mayor tamaño que la anterior, pero el servicio está basado en atención vía vendedores para asistir a los clientes al igual que la otra cadena. En esta tienda se midieron las mismas métricas y se utilizaron todas las fuentes de medición, salvo videos de seguimiento completo. Toda esta información es mostrada por medio de una plataforma web, en donde el cliente puede ingresar para tener una visión de lo que está ocurriendo en la tienda (a nivel de hora, día, mes y año), siendo posible tomar decisiones de gestión de acuerdo a los objetivos de la cadena.

5. Definición de métricas para los modelos econométricos

En este capítulo se definirán las variables que fueron utilizadas en el modelo, analizando los estadísticos descriptivos más importantes, para luego determinar implicancias previas relacionadas a las hipótesis por medio de un análisis exploratorio.

5.1. Definición de variables

Los datos recopilados abarcan el año 2012 entre los meses de Marzo y Octubre, teniendo dos fuentes principales : (i) Datos POS y (ii) Datos de las cámaras de video.

(i) Los datos del punto de venta contienen el detalle de cada transacción realizada, tal como la identificación de cada de producto (SKU), precio pagado por el cliente con el descuento asociado, tipo de transacción, método de pago, fecha transacción, tienda e identificación del cliente en el caso de que exista tarjeta de lealtad asociada. Las variables creadas a partir de estos datos son:

Ventas ($Sales_{jt}$): Se define como la suma de todas las ventas en dólares agrupadas por tienda j y día t . En el valor total se incluyen descuentos asociados y únicamente se consideraron transacciones del tipo venta.

Transacciones ($Trans_{jt}$): Se define como el total de transacciones agrupadas por tienda j y día t . Se incluyen únicamente transacciones del tipo venta, dejando fuera transacciones del tipo cambio de producto y devoluciones.

Ticket ($Ticket_{jt}$): Se define como el monto promedio de todas las boletas del tipo venta que se registraron en un día t y tienda j . Este se calcula como se muestra en la ecuación (1).

$$Ticket_{jt} = \frac{Sales_{jt}}{Trans_{jt}} \quad (1)$$

Una observación importante es que estos datos solo aportan información referente a clientes que compraron algún producto, lo que motiva a definir las variables provenientes de las cámaras, los cuales contienen información de las distintas interacciones que tienen los clientes en la tienda (independiente si compró o no). Las variables son:

Tráfico ($Traf_{jt}$): Se define como la suma de todas las entradas de personas para la tienda j en el día t . Como se comentó anteriormente esta variable entrega información de la cantidad real de clientes que ingresaron a la tienda mediante el conteo automático por medio de un sensor, siendo esta métrica un indicador de la congestión en la tienda.

Este sistema de conteo puede tener fallas y por lo tanto es importante detectar cuándo ocurre esta situación, si no, el valor obtenido puede no ser representativo. El principal error ocurre cuando las cámaras pierden conexión, como consecuencia de lo anterior se registran valores cero de conteo, siendo que la tienda se encuentra en un día de normal funcionamiento. Ocurrido esto, no hay forma de recuperar el conteo exacto de personas, solo se podría realizar una estimación por medio de un pronóstico para el día en cuestión. Para tener más claridad de lo anterior, en la figura 1 se muestra el tráfico para una de las dos tiendas. El color azul representa el conteo de la cámara 1 perteneciente a la entrada principal, el rojo muestra el conteo de la cámara 2 y el verde es la suma de los dos anteriores. Los criterios utilizados para el manejo de estos datos fueron:

1. Eliminar de la muestra los días en donde alguna cámara perdió conexión total.
2. Eliminar de la muestra tanto el día previo como el día siguiente de la pérdida de conexión. La razón de esto se debe a que la conexión o desconexión de la cámara no necesariamente ocurre al principio de la jornada laboral, es decir, la ocurrencia se puede efectuar en cualquier hora del día, no teniendo registro del momento exacto en que ocurrió. En la figura 1 se puede apreciar que para los días contiguos a una falla, la media es notablemente menor al tráfico de los otros días.
3. Dejar fuera del análisis todos los registros desde Septiembre en adelante debido a que la cámara número dos perdió conexión total en ese período (solo para una de las dos tiendas).

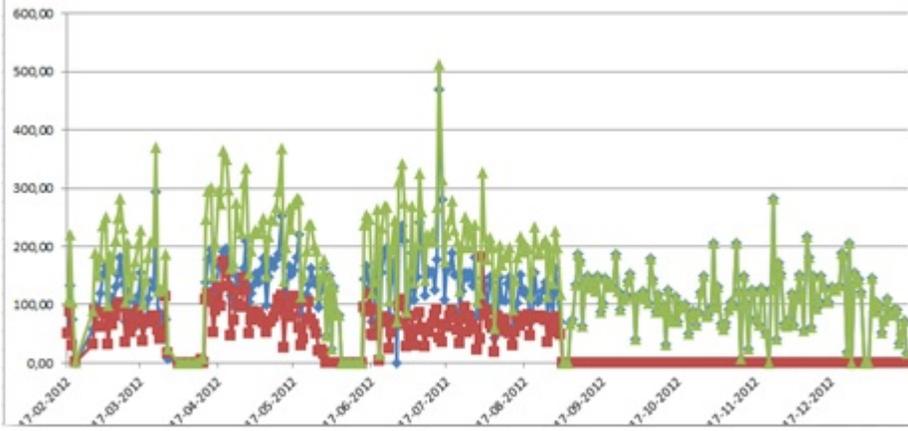


Fig. 1: Tráfico por cámara. El color azul indica el conteo de entrada de la cámara número uno y el rojo el conteo de la cámara número dos, en tanto el color verde es la suma de los anteriores. Es posible distinguir que desde el mes de Septiembre hasta Diciembre la cámara dos tuvo una des-conexión entregando así conteo cero para ese período. Fuente: Elaboración Propia

De la muestra final se eliminó el 8,9% del total de los días considerando ambas tiendas, lo que equivale a 53 días de un total de 595, disminuyendo el error o sesgo de los resultados finales.⁴

Una vez definido la variable tráfico, es posible incorporar dos nuevas variables (estas se utilizarán como variables dependientes en los modelos) que relacionan datos transaccionales con datos de cámara, como sigue a continuación:

Conversión ($Conv_{jt}$): Se define como el número de transacciones sobre la entrada total de clientes en la tienda, ver ecuación (2). Esta variable mide el desempeño de la tienda en términos de la capacidad para convertir un potencial comprador en comprador y se utilizó como variable dependiente en el modelo principal.

$$Conv_{jt} = \frac{Trans_{jt}}{Traj_{jt}^*} \quad (2)$$

Ventas por entrada ($Salpw_{jt}$): Se define como el total de ventas sobre la entrada total de clientes en la tienda, ver ecuación (3). Esta métrica combina dos efectos importantes que se impulsa por el compromiso de conversión: personas que compran productos y promedio de dinero gastado por cliente que visita la tienda.

⁴ Adicionalmente se detectaron errores en los datos que venían de forma desagregada. En algunos casos para un día y tienda en particular se duplicaban lo registros, como consecuencia de esto el conteo para esos días era mayor que el real. Los datos duplicados fueron eliminados también.

$$Salpw_{jt} = \frac{Sales_{jt}}{Traf_{jt}} \quad (3)$$

A continuación se definen otras métricas obtenidas de las cámaras, tomando una muestra cada quince minutos dentro del horario de funcionamiento de la tienda. Las variables son las siguientes:

Vendedores ($Assoc_{jt}$): Se define como la cantidad de vendedores presentes en la sala de ventas para una tienda j y día t . Dado la definición, se excluye del conteo a todo trabajador que se encuentre fuera de sus labores, por ejemplo en hora colación o desempeñando funciones como cajero.

Clientes ($Cust_{jt}$): Se define como la cantidad de clientes presentes en las zonas visibles de la tienda para la tienda j y el día t . Se deja fuera del conteo a los clientes en el interior del probador. Esta es una variable alternativa al tráfico, con la diferencia que esta permite saber cuantos clientes habían, prácticamente en el mismo instante del contacto entre vendedor y cliente (si es que existe).

Atención (Att_{ijt}): Variable binaria que toma el valor 1 si existe contacto entre vendedor y cliente i para la tienda j y el día t , 0 en caso contrario. Un cliente es atendido cuando existe contacto visual a una distancia prudente distinguiendo comunicación con el vendedor, la cual permita establecer que existe interacción entre las partes sin importar quien tomó la iniciativa.

Tasa de atención ($RAtt_{jt}$): Se define como la proporción de clientes atendidos para una tienda j en un día t , ver ecuación (4). Esta es la principal variable del estudio, buscando determinar el efecto de esta variable en la conversión.

$$RAtt_{jt} = \sum_i Att_{ijt}, \forall j,t. i \text{ representa la muestra recopilada} \quad (4)$$

Esta variable se utiliza como un estimador del nivel atención, teniendo asociado un error de muestreo, que se acentúa aún más para días con pocos registros (llegando a ser un estimador poco representativo). En la muestra se consideraron solo los días con 10 o más muestras. El criterio de selección se basó en el cálculo del máximo error estándar por día (también se calculo el error estándar muestral), seleccionando el número de muestras que cumplen con las siguientes condiciones:

1. Mínimo error estándar.

2. Error estándar es marginal al agregar una muestra adicional.⁵ Por último se define:

Ratio (SR_{jt}): Definido como la cantidad de clientes sobre el número de vendedores para una tienda j y día t , ver ecuación (5). Esta variable, a diferencia del tráfico, indica qué tan congestionada está la tienda condicional al personal disponible.

$$SR_{jt} = \frac{Cust_{jt}}{Assoc_{jt}} \quad (5)$$

En la tabla 1 se muestran los estadísticos descriptivos para las variables Clientes, Vendedores y Ratio que se encuentran agrupadas por día, teniendo en la muestra individual (sin agrupar) 7137 registros. En promedio hay casi 2 vendedores activos (sin contar vendedores en la zona de cajas), es decir cumpliendo el rol de atender a los clientes que circulan en la tienda. Al día en promedio hay 3 clientes activos en la tienda y en el período de mayor tráfico este valor puede llegar hasta aproximadamente 7 clientes en el interior de la tienda.

Tab. 1: Estadísticos Descriptivos vendedores, clientes y ratio

Variable	Mean	Std. Dev.	Min.	Max.
Vendedores	1.908	0.353	1.036	3.107
Clientes	3.293	0.945	0.857	6.5
Ratio	1.739	0.444	0.706	3.491
N		391		

Con respecto al ratio de clientes y vendedores, a medida que aumenta esta variable los vendedores deben ser capaces de atender a más clientes, afectando negativamente la tasa de atención en la tienda. En la figura 2 se observa que el día sábado presenta la mayor cantidad de clientes (para las dos tiendas), teniendo un ratio mayor para este día (casi 2 clientes por vendedor), ya que la cantidad de vendedores no aumenta en la misma proporción que lo hace el tráfico. En la misma figura no se aprecia una gran variabilidad de personal en la semana, teniendo un promedio de 2 vendedores cada día por tienda. Por lo anterior se asume que no existe un manejo de dotación del personal para mantener el mismo nivel de atención durante la semana.

En la tabla 2 se muestran los estadísticos descriptivos para el tráfico, conversión, *ticket* promedio y tasa de atención, todas estas variables agrupadas por día y tienda. El tráfico promedio por día es de 212, de los cuales el 27% en promedio realiza una transacción (conversión). Adicionalmente el ticket promedio es de 135 dólares, resultando en promedio 7727 dólares en

⁵ De todas formas en los modelos finales se realizó un análisis de sensibilidad, variando el número de muestras en la tasa de atención por día. Ver apéndice número 1.

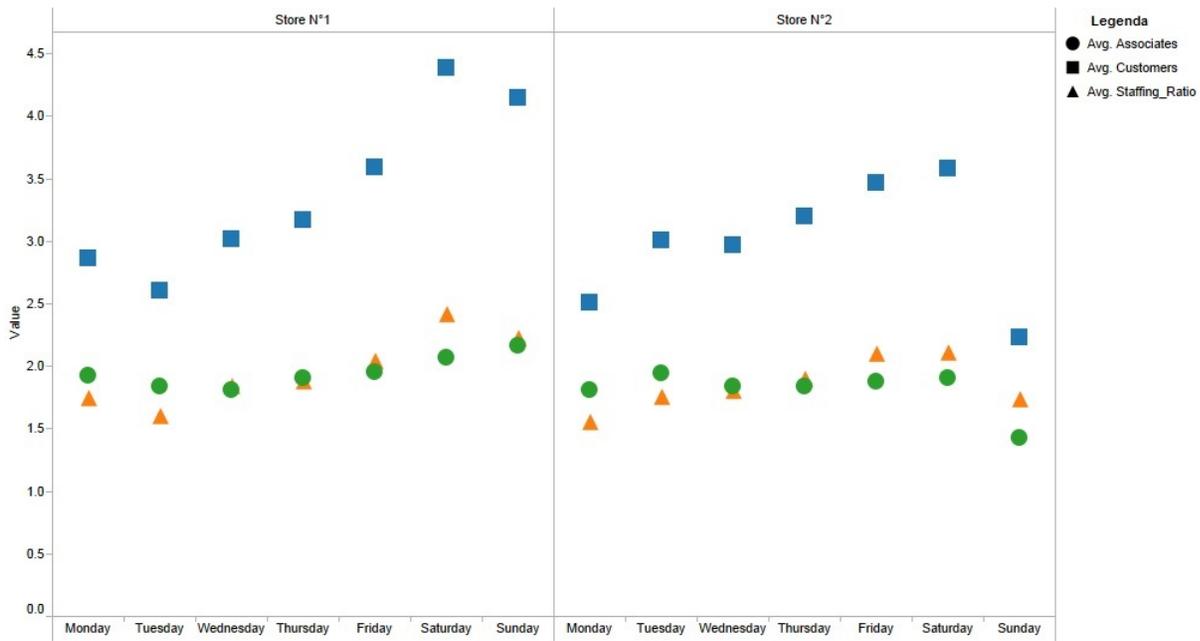


Fig. 2: *Cientes, vendedores y ratio por día de semana para las dos tiendas.* Fuente: Datos Scopix Solutions, Gráfico elaboración propia.

Tab. 2: Estadísticos descriptivos

Variable	Mean	Std. Dev.	Min.	Max.	N
tráfico	212.638	65.635	60	513	298
conversión	0.272	0.09	0.114	0.765	254
ticket	135.783	30.135	66.682	329.359	333
tasa_de_atención	0.585	0.123	0.211	1	392
tasa_de_atención*	0.505	0.111	0.186	0.838	392

ventas diarias. Finalmente la mitad de los clientes tiene algún contacto con un vendedor sin importar quién tomó la iniciativa. Se muestran 2 variables que reflejan la tasa de atención, la que se muestra en (*) incluye los registros no observados (en el desarrollo de la metodología se explica cómo se agregaron los registros categorizados como no observados en la muestra). Al agregar a la muestra los registros no observados, la media disminuye de un 58 % a un 50 %, tendencia razonable ya que la probabilidad de que un cliente sea atendido condicional a que lleva cierto tiempo en la tienda sin ser atendido es menor a la media de la tasa de atención.

De la muestra obtenida con los videos de seguimiento completo se generó la distribución del tiempo de espera hasta que se produce el contacto entre vendedor y cliente, ver figura 3. Del total de la muestra, el 84,5 % del total de los contactos se produjo antes de los 8 minutos. Esto implica que la tasa de atención debería tener al menos 16,5 % de evidencias etiquetadas como no observado. Considerando que no necesariamente los clientes llegan al principio del video, se espera que esta cifra sea mayor. Tomando los datos individuales (videos de corta duración), se recopilaron 4576 registros válidos (casos que no se etiquetan como no observado) y en caso de los registros *no observados* son 1228, representando un 21 % de la muestra total. De todos modos sin considerar los registros no observados se captura la mayor parte de la muestra de clientes.

En la tabla 3 se muestran los estadísticos descriptivos para las variables ticket, tráfico y tasa de atención por tienda. Nuevamente ambas presentan estadísticos similares, condición necesaria para aplicar una parte de la metodología.

Tab. 3: Estadísticos descriptivos ambas tiendas

Variable	Mean	Std. Dev.	N
ticket 1	138.293	33.329	172
tráfico 1	215.871	68.371	132
tasa de atención 1	0.485	0.111	203
ticket 2	133.103	26.138	161
tráfico 2	210.066	63.467	166
tasa de atención 2	0.526	0.107	189

5.2. Controles

Las dos tiendas en estudio se encuentran en la misma ciudad (a menos de 100 kilómetros de distancia), por ende no existen diferencias por ubicación. La estacionalidad podría jugar un rol importante, afectando tanto al tráfico como a las ventas. Se controlará por :

Semana (η_t^w): Se define como una variable binaria, tomando el valor 1 si el día corresponde

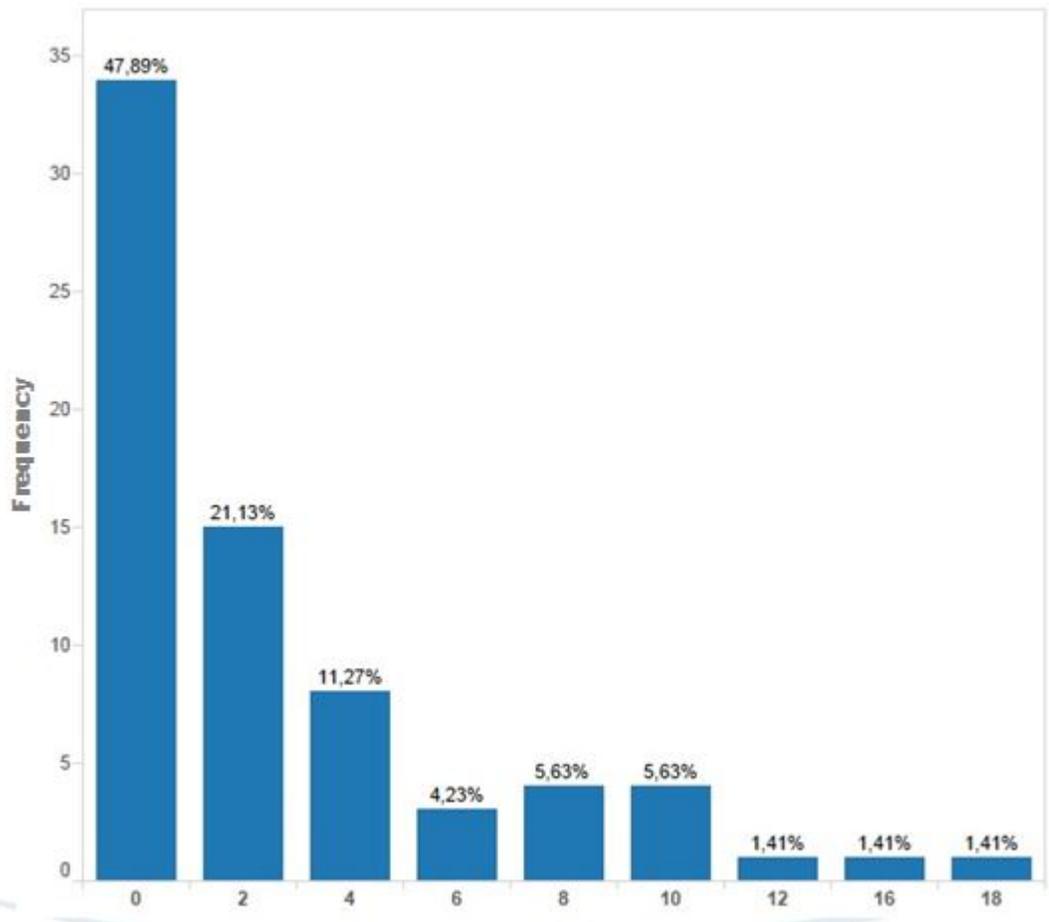


Fig. 3: *Distribución tiempos de espera antes del contacto entre vendedor y cliente.* Fuente: Elaboración propia.

a la semana w . La cantidad de variables binarias incluidas en el modelo es igual al número de semanas incluidas en la muestra -1.⁶

Día (δ_{dt}): Se define como una variable binaria, tomando el valor de 1 si el día t corresponde al día de la semana d , $d \in \{Lunes, Martes, Miercoles, Jueves, Viernes, Sabado\}$.

Tienda (s_j): Se define como una variable binaria, tomando el valor 1 si la tienda j corresponde a la tienda 1 y 0 en caso contrario.

Feriado (γ_t): Se define como una variable binaria, tomando el valor 1 si el día t es feriado y 0 en caso contrario. El efecto que tiene esta variable se debe a que cercano a los días feriados las tiendas de retail suelen tener promociones en sus productos, consideraron también los 4 días previos. Los feriados que coinciden con la muestra son: *President day, Memorial day, Independence day, Labor day, Columbus day, Mother day* y *Women day*.

El clima es otro efecto importante, siendo las condiciones climáticas influyentes en el tráfico. Los datos para construir este control se recopilaron desde el sitio web de *Weather Underground*,⁷ en el cual se localizó el aeropuerto más cercano a la tienda, obteniendo de esta forma las condiciones climáticas tales como, temperatura, velocidad del viento, estado visible del clima, entre otros. Las variables binarias que se utilizaron para controlar por condiciones climáticas fueron:

Despejado (α_{jt}^c): Variable toma el valor de 1 si el día t y tienda j se encuentra despejado y 0 en caso contrario.

Lluvia (α_{jt}^r): Variable toma el valor de 1 si el día t y tienda j se encuentra con lluvia y 0 en caso contrario.

Nublado (α_{jt}^f): Variable toma el valor de 1 si el día t y tienda j se encuentra nublado y 0 en caso contrario.

Las variables que se utilizaron para controlar por temperatura son la siguientes:

Frío (β_{jt}^c): Variable toma el valor de 1 si la temperatura promedio para el día t y tienda j es menor a 10° celsius y 0 en caso contrario.

⁶ Sabiendo el valor de $w-1$ variables, se puede determinar la semana de cada registro, siendo las semanas excluyentes. Lo mismo se aplica para las otras variables dummy que sean de la misma naturaleza.

⁷ Para más información visitar :<http://www.wunderground.com/history/>

Calor (β_{jt}^h): Variable toma el valor de 1 si la temperatura promedio para el día t y tienda j es mayor a 20° celsius y 0 en caso contrario.

Notar que estas variables no son excluyentes. De los resultados se destaca que para los días de lluvia menos clientes van a la tienda, reflejado en la correlación negativa entre estas variables.

6. Medición de impacto mediante videos de seguimiento completo

Se utilizaron videos de seguimiento completo para observar las trayectorias de los clientes al interior de la tienda, recopilando datos de las interacciones que se producen en el proceso de compra, poniendo especial énfasis en la intervención por parte de los vendedores. En la figura 4 se muestra el proceso completo del cliente, desde que ingresa a la tienda hasta que se retira. Según los resultados, las fases o acciones que más afectan en la probabilidad de compra son el contacto entre cliente y vendedor y el uso del probador. Algunos elementos del proceso mostrado en la figura 4 se pueden repetir, es decir, un cliente puede ser atendido o entrar en el probador más de una vez en el proceso de la misma compra.

Tab. 4: Estadísticos descriptivos variables obtenidas desde los videos de seguimiento completo

Variable	Mean	Std. Dev.	Min.	Max.	N
Tomo_prenda	0.696	0.461	0	1	191
tasa de atención	0.372	0.485	0	1	191
vendedor_se_acerca	0.493	0.504	0	1	71
Tiempo_espera_vendedor	220.93	235.773	7	1110	71
tiempo_de_servicio	85.254	115.088	2	708	71
Probador	0.319	0.467	0	1	191
Tiempo_probador	388.049	429.85	2	2004	61
compra	0.236	0.425	0	1	191
Tiempo_tienda	591.675	582.168	22	2808	191

Se recopiló una muestra de 191 clientes, midiendo las etapas más importantes del proceso de compra. Para esto se definen algunas variables adicionales:

Toma prenda: Tiempo (segundos) del video en que el cliente en cuestión toma una prenda, sacándola completamente del estante o perchero donde se encuentra. Al sacarlo de un perchero, la prenda debe quedar completamente fuera de este, en el caso de los mesones, se debe sacar la prenda de los planos verticales virtuales que producen los bordes de la mesa.

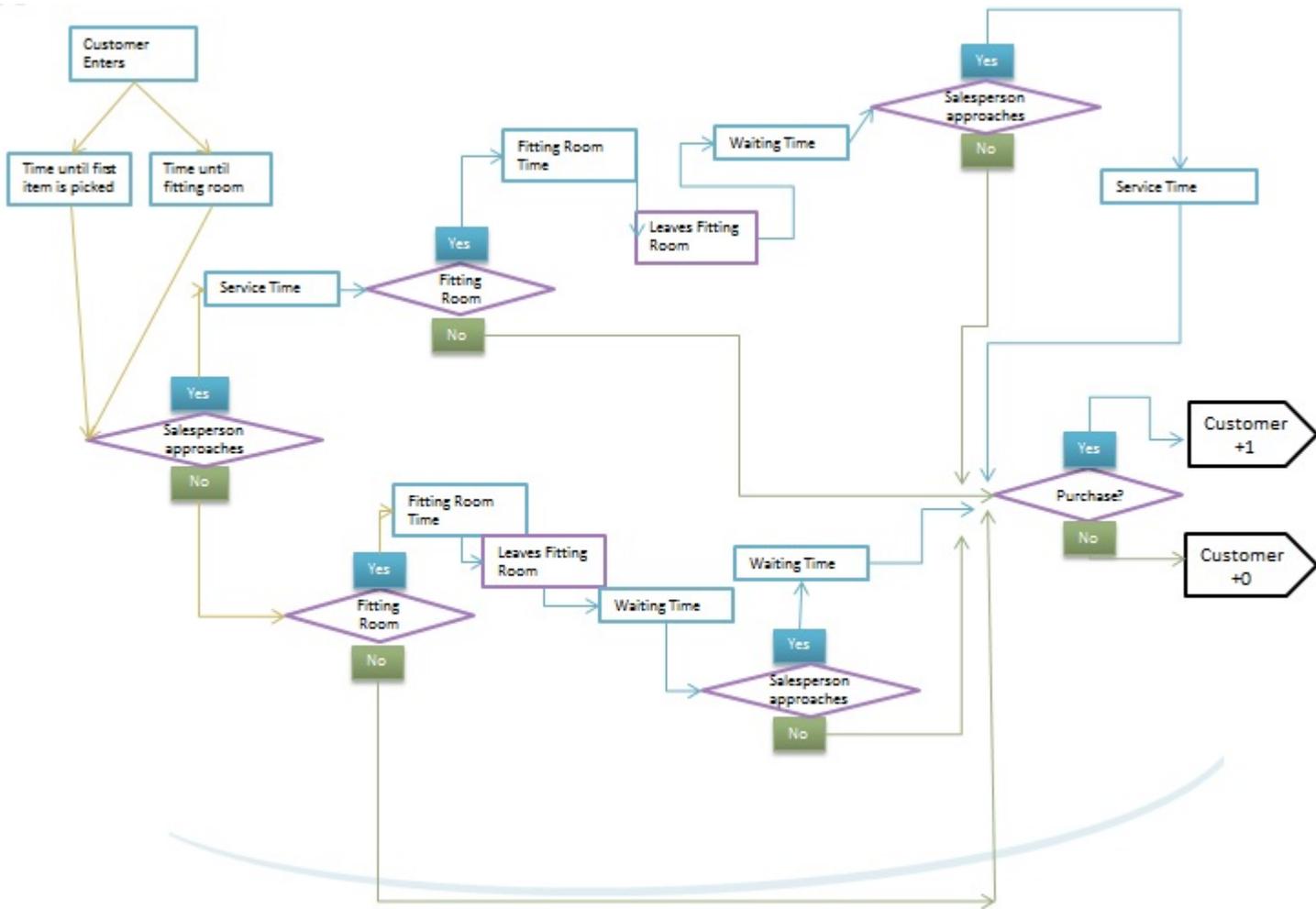


Fig. 4: Trayectoria de los clientes en el interior de la tienda. Tanto el camino como las fases que aparecen se construyeron por la observación de videos de seguimiento completo. Fuente: Elaboración propia

Vendedor se acerca: Variable binaria que indica si el vendedor se acerca al cliente (valor 1) o si no lo hace (valor 0), siendo el vendedor quien toma la iniciativa.

Tiempo de espera vendedor: Tiempo (segundos) del video en que el cliente tiene algún tipo de contacto con un vendedor, considerándose como contacto, para este efecto, cuando un vendedor ofrece ayuda al cliente o bien el cliente pide ayuda a un vendedor, registrándose el tiempo desde que el cliente ingresa hasta que se produce el contacto.

Tiempo de servicio: Tiempo (segundos) del video en que termina el tiempo de servicio del vendedor hacia el cliente (el tiempo se comienza a registrar cuando se produce el contacto). Sólo se considera el primer servicio.

Probador: Variable binaria que toma el valor 1 si cliente ingresa al probador y 0 en caso contrario.

Tiempo probador: Tiempo (segundos) desde que el cliente ingresa al probador hasta que se retira.

Compra: Variable binaria que toma el valor 1 si cliente realiza compra y 0 en caso contrario.

Tiempo tienda: Tiempo (segundos) total de permanencia en la tienda por parte del cliente.

Los estadísticos descriptivos se muestran en la tabla 4. En promedio casi el 70% de los clientes tomó alguna prenda, sacándola completamente del estante o perchero. En base a la definición anterior, la mayoría de los clientes se interesa por al menos algún producto, limitando superiormente la conversión en un 70%. En cuanto al nivel de servicio, en promedio el 37% de los clientes tuvo contacto con algún vendedor. Llama la atención que este valor sea bajo si se toma en cuenta que en promedio entran aproximadamente 20 clientes por hora.⁸

Por la baja cantidad de clientes que ingresan al local se identificó que los vendedores no tienden a ser activos en términos de tomar la iniciativa, pasando la mayor parte del tiempo ordenando productos y prestando ayuda si un cliente así lo solicita (comportamiento ratificado por medio de los videos). En cuanto a la iniciativa, la mitad es por parte de los vendedores y la otra mitad por parte de los clientes, es decir, el bajo nivel de contacto también es el resultado del poco requerimiento por parte de los clientes hacia los vendedores. En promedio el contacto se efectúa en los 220 segundos con un tiempo de servicio promedio de 85 segundos. Otro dato interesante es que en promedio el 31,9% de los clientes ingresan al probador. Ya entrando al probador el cliente demuestra más intención de compra (luego de tomar alguna

⁸ Este valor se estimó tomando en cuenta que el tráfico promedio por día es 212, con un total de 10 horas de atención.

prenda). Finalmente, en promedio el 23,6 % de los clientes termina comprando, valor de magnitud similar a la entrada al probador, lo que ya puede entregar una idea de alguna relación entre estas dos variables. La muestra recopilada permitió reducir el nivel de incertidumbre, desarrollando de esta manera implicancias previas por medio de test de hipótesis⁹, resultados mostrados a continuación:

1. Efecto contacto vendedor-cliente: Si un cliente es atendido por un vendedor: ¿Es más probable que compre? En donde H_0 : conversión media de clientes atendidos es igual a la conversión media de los clientes que no son atendidos, obteniendo un 41 % versus 13 % (p-valor $< 0,01$) respectivamente, rechazando así la hipótesis nula de que la atención no tiene efecto en la conversión. Lo anterior implica que al existir contacto entre vendedor y cliente (sin importar quién toma la iniciativa) impacta positivamente en la ventas por medio de la conversión.
2. Si el vendedor tomó la iniciativa y se acercó al cliente: ¿Es más probable que el cliente compre? en contraste al caso donde no existe atención. De la muestra de 191 clientes, el grupo perteneciente a los clientes no atendidos son 120 casos, teniendo una media en la conversión del 13,3 %, mientras que el grupo de clientes que son atendidos por iniciativa del vendedor son 35 casos, con una media en la conversión del 20 %. El estadístico asociado al test resulta : $z = 0,98$ con un $p - valor = 0,32$. Por lo tanto, no existe evidencia estadística para rechazar la hipótesis nula con un nivel de confiabilidad del 95 %, es decir, ser atendido en caso de que el vendedor sea quien toma la iniciativa no tiene efectos significativos en la conversión con respecto a la situación donde no existe atención. De esta forma, no existe evidencia estadística para concluir que los vendedores escogen a clientes con una mayor probabilidad de compra.
3. Si el cliente tomó la iniciativa, ¿Es más probable que compre? En donde H_0 : conversión media de clientes atendidos cuando cliente toma la iniciativa es igual a la conversión media de los clientes que no son atendidos, obteniendo un 61 % versus un 13 % (p-valor $< 0,01$) respectivamente, rechazando la hipótesis nula de que cuando el cliente toma la iniciativa y se acerca al vendedor se tiene la misma conversión que al no recibir atención. Este resultado es claro y nos permite concluir previamente que los clientes que soliciten ayuda a un vendedor tienen una mayor probabilidad de comprar algún producto.

⁹ Específicamente se aplicó test de proporciones para dos muestras independientes, utilizando el siguiente estadístico: $Z = (p_i - p_j) / S_{p_i - p_j} \sim \mathbb{Z}(1 - \alpha/2)$, donde $i \neq j$. En donde: $S_{p_i - p_j} = \text{raiz}(p^*(1 - p^*) [1/n_i + 1/n_j])$

Donde $p^* = (n_i * p_i + n_j * p_j) / (n_i + n_j)$,

Con p_i : proporción de la muestra i que presenta una característica determinada

n_i = número de observaciones de la muestra i

4. Si el cliente ingresa al probador, ¿Es más probable que compre? H_0 : conversión media de clientes que ingresan al probador es igual a la conversión media de los clientes que no ingresan, obteniendo un 52,4 % versus 10 % (p-valor < 0,01) respectivamente. Lo anterior deja en manifiesto que si un cliente ingresa al probador presenta una probabilidad mayor de compra, es decir, si se toma el tiempo de probarse una prenda es porque posiblemente la llevará en el caso de que se cumplan sus expectativas.
5. Si cliente es atendido, ¿Es más probable que ingrese al probador? H_0 : la tasa media de entrada al probador de clientes que son atendidos es igual a la tasa media de entrada al probador de los clientes que no son atendidos, obteniendo un 56 % versus 17 % (p-valor < 0,01) respectivamente, rechazando la hipótesis nula de que la atención no genera una mayor entrada al probador.

Tanto los estadísticos descriptivos como los test de hipótesis permiten concluir preliminarmente qué rol deberían tener los vendedores con el fin de aumentar la conversión en la tienda, pudiendo ser los siguientes:

- Estar disponible en el caso que algún cliente necesite ayuda (ya sea para encontrar algún producto, preguntar por un precio, solicitar un producto u otra talla).
- Facilitar o persuadir a un cliente para que ingrese al probador.

Las acciones anteriores por parte de los vendedores ayudan a aumentar la conversión en la tienda y un posible entrenamiento para que ellos ejecuten estas tareas con más dedicación podría favorecer positivamente las ventas. Como resultado previo, la probabilidad de compra para un cliente que recibe atención es más alta que en el caso de los clientes no atendidos, siendo este resultado muy importante ya que se tiene una idea de la dirección de este efecto. Ante esto, las posibles explicaciones son:

1. La atención aumenta la probabilidad de compra
2. Los clientes que tienen la intención de comprar algún producto a menudo contactan a un vendedor para solicitar ayuda, sin embargo, podrían comprar de igual forma sin recibir ayuda por parte del vendedor.

Con las dos explicaciones anteriores no es posible cuantificar el efecto de la atención sobre la conversión, e incluso podrá existir un problema de endogeneidad en la variable de atención. Hasta el momento se tienen tres limitaciones:

1. Problema de causalidad: no se tiene claridad si el nivel de atención influye en la intención de compra o el cliente al tomar la decisión de comprar induce la atención, por lo que se requiere una fuente exógena de variación en la tasa de atención para solucionar este problema.
2. Elevado costo para replicar el análisis a gran escala con variadas tiendas y períodos.
3. Dificultad para medir impacto en la canasta de compra y otros aspectos relacionados a la experiencia de compra.

Ante las anteriores limitaciones se utilizaron datos asociados a videos de seguimiento parcial, de los cuales se puede tomar una elevada frecuencia de muestras al día (cada 15 minutos), logrando realizar un seguimiento del cliente de hasta 8 minutos.

7. Medición de impacto mediante videos parciales

En este capítulo se mostrará el modelo econométrico utilizado, explicando los problemas asociados a su especificación, para finalmente describir la metodología que permite solucionar estos problemas.

7.1. Especificación modelo econométrico

7.1.1. Derivación modelo lineal

La principal motivación de este trabajo es estudiar el impacto que tiene la atención de los vendedores en el comportamiento de compra de los clientes, por ello la tasa de atención definida como la proporción de clientes atendidos para un cierto período (día) es la principal variable en estudio. No solo es importante entender el efecto que tiene esta variable sobre conversión, sino que también los posibles problemas ligados a la especificación del modelo. A continuación se muestra cómo se deriva el modelo en el caso que fuera lineal, que a pesar de no ser el modelo utilizado, permitirá un mejor entendimiento (definición y estimación) de la tasa de atención y conversión. Se define:

- $y_i^* \in \{0, 1\}$, toma el valor de 1 cuando el cliente i realiza una compra y 0 si el cliente i no realiza compra.
- $w_i^* \in \{0, 1\}$, toma el valor 1 si el cliente i es atendido, y 0 en caso contrario.

- $g(w_i)$ función de la tasa de atención.

Las variables w_i^* , y_i^* son variables que entregan información para cada cliente que ingresa a la tienda, datos que no se disponen debido a que se maneja una muestra del total de clientes. El valor esperado para la compra de un cliente en función de la atención recibida se modela de la siguiente forma:

$$E(y_i^*) = Pr(y_i^* = 1) = g(w_i^*) + e_i \quad (6)$$

Adicionalmente, otro problema asociado a esta muestra es que no se tiene la información para el mismo cliente (i), es decir, si fue atendido y si compró o no. Esto motiva la utilización de datos transaccionales agregados por tienda (j) y día (t), los cuales permiten estimar la conversión utilizando esta información en conjunto al tráfico. Se define el número de períodos (T) como el total de días que se tiene en la muestra. Se define la tasa de conversión como sigue:

$$Conv_{jt} = \frac{\sum_{i \in I_t} y_{ijt}^*}{Traf_{jt}} = \frac{Trans_{jt}}{Traf_{jt}}, \text{ donde } t \in \{1, 2, 3..T\} \quad (7)$$

Utilizando (6) y (7) se calcula el valor esperado de conversión:

$$E(Conv_{jt}) = \frac{\sum_{i \in I_t} E(y_{ijt}^*)}{Traf_{jt}} = \frac{\sum_{i \in I_t} g(w_{ijt}^*)}{Traf_{jt}} + \varepsilon_{jt} \quad (8)$$

Como se tiene una muestra de N_{jt} observaciones de la atención por día y tienda, la ecuación (8) se aproxima utilizando:

$$\frac{\sum_{i \in I_t} g(w_{ijt}^*)}{Traf_{jt}} \approx \frac{\sum_{i \in N_{jt}} g(w_{ijt})}{N_{jt}} + \xi_{jt}, \text{ donde } N_{jt} \subseteq I_{jt} \quad (9)$$

Se define:

$$\frac{\sum_{i \in N_{jt}} g(w_{ijt})}{N_{jt}} = Ratt_{jt}$$

Finalmente agregando los controles, se tiene lo siguiente:

$$Conv_{jt} \approx \alpha_1 + \alpha_2 Ratt_{jt} + \alpha_3 Cust_{jt} + \text{controles} + e_{jt} \quad (10)$$

Donde : $\text{controles} = \alpha_4 \gamma_t + \alpha_5 \delta_{dt} + \alpha_6 \sum_w \eta_t^w + \alpha_7 \alpha_{jt}^c + \alpha_8 \alpha_{jt}^r + \alpha_9 \alpha_{jt}^f + \alpha_{10} \beta_{jt}^c + \alpha_{11} \beta_{jt}^h + \alpha_{12} s_j$

Sin embargo, esta especificación podría presentar algunos problemas con respecto a la tasa de atención y la conversión. En el caso de la conversión, al estimarse por medio de una regresión lineal, las estimaciones podrían caer fuera del rango para el cual está definida la variable $[0,1]$.

Con respecto a la tasa de atención, al ser una estimación de la tasa de atención real, tendrá asociado un error de muestreo. También esta variable presenta un problema de sesgo de selección asociado a la recopilación de la muestra, como se ilustra en el siguiente ejemplo: supongamos que se tenemos 2 clientes (A y B), quienes llegan en el minuto 6 del video. Cliente A es atendido en el minuto 7, por lo tanto la variable atención toma el valor de 1. Supongamos que para B pasaron los 2 minutos y no fue atendido pero tampoco se fue de la tienda, resultado de lo anterior la variable se considera como *no observado*, desechando este registro de la muestra. Lo anterior se considera un problema en el sesgo de selección, donde se incluye al cliente A, pero B no se considera por no tener certeza de lo que ocurrido con él, siendo que tanto A como B llegaron en el mismo instante del video, es decir, un registro se toma en cuenta dependiendo de lo ocurrido con el cliente.

Por último, se presenta un problema desafiante, la tasa de atención podría presentar problemas de endogeneidad por sesgo de causalidad simultánea, ya que no es claro si la atención influye en la compra del cliente o al revés. Si efectivamente es una variable endógena y no se corrige, los estimadores podrían llegar a ser sesgados e inconsistentes, consecuencia de esto las conclusiones finales podrían ser erróneas.

7.1.2. Especificación modelo Logit

Entendiendo que los problemas anteriores son muy importantes de corregir con el fin de tener conclusiones que sean válidas, se desarrolló una metodología aplicada que se hace cargo de todos estos problemas asociados a la especificación del modelo. En particular, la ecuación (10) se estimará por medio de un modelo logit, logrando de esta manera valores predichos en un rango entre $[0, 1]$, mismo recorrido que tiene la variable dependiente.

En el caso de los registros no observados se estimó la probabilidad de que esos clientes fueran atendidos, agregando un valor numérico entre 0 y 1 para la tasa de atención. Los problemas de

endogeneidad y error de muestreo se resolvieron mediante el uso de variables instrumentales. Se postula el siguiente modelo (datos agrupados por día y tienda) :

$$\text{Logit}(\text{Conv}_{jt}) = \alpha_1 + \alpha_2 \text{Ratt}_{jt} + \alpha_3 \text{Cust}_{jt} + \text{controles} + e_{jt} \quad (11)$$

Donde:

$$\text{Logit}(\text{Conv}_{jt}) = \log\left(\frac{\text{Conv}_{jt}}{1 - \text{Conv}_{jt}}\right) \quad (12)$$

En la ecuación (11) se agregaron como variables independientes tasa de atención (Ratt_{jt} , principal variable en estudio) y clientes en tienda (Cust_{jt}), variable que permite controlar posibles cambios en la conversión por efecto del aumento de congestión en la tienda. Las demás variables agregadas en el modelo corresponden a controles de tienda, estacionalidad y clima, las cuales fueron definidos en el capítulo anterior. A continuación se detallan estos problemas y la solución planteada para llegar al modelo final.

7.2. Corrección de sesgo de selección

Como se comentó anteriormente la variable atención puede tener 3 posibles resultados: Cliente fue atendido, no fue atendido y no observado. En este último caso no se tiene conocimiento de lo ocurrido con el cliente. En la figura 5 se muestra un ejemplo con los tres casos:

1. Se observa contacto entre vendedor y cliente antes del minuto 8.
2. No existe contacto entre vendedor y cliente, retirándose este último antes del minuto 8.
3. No existe contacto entre vendedor y cliente, e incluso video finaliza con el cliente en su interior. El 21 % de los casos corresponden a esta categoría.

Para solucionar este problema se busca estimar el valor de la tasa de atención en todos los registros asociados a la tercera categoría. Específicamente, se estimó la probabilidad de que el cliente sea atendido en los minutos no cubiertos por el video hasta completar los 8 minutos, condicional a que el cliente no ha sido atendido hasta ese momento ni tampoco se ha retirado del local. Esta probabilidad se calculó utilizando una estimación no paramétrica tomando los datos provenientes de los videos de seguimiento completo.¹⁰

¹⁰ Solo fue posible conseguir los datos de una sola tienda, sin embargo ambas tiendas presentan estadísticos descriptivos parecidos, por lo tanto se asumirá que ambas tiendas son de iguales características. Usando estos datos se estimó la probabilidad para ambas tiendas.

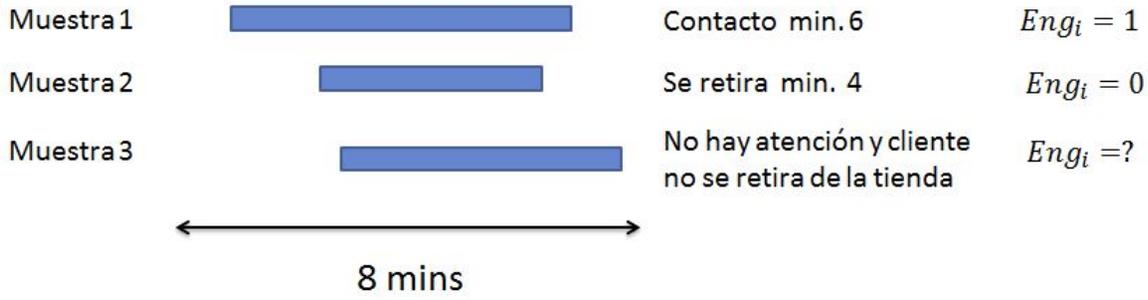


Fig. 5: *Sesgo selección*. Fuente: Elaboración propia

Se define lo siguiente:

- x : Tiempo máximo de observación para un cliente.
- t_a : tiempo del video donde el cliente llega a la tienda.
- t_{end} : tiempo del video en donde se retira.
- Δ : $t_{end} - t_a$, tiempo que el cliente permanece en la tienda ($\Delta < 480$)
- t_l : tiempo del video donde el cliente se retira de la tienda.

Los datos no observados cumplen con lo siguiente : $t_{eng} > \Delta \cap t_l > \Delta$. En los videos etiquetados como no observados, se sabe en el instante del video en el cual el cliente ingresó a la tienda, ya que se registra el tiempo desde el momento que ingresó y se sabe que los videos tienen una duración de 8 minutos. Para los otros videos donde los clientes salieron de la tienda o fueron atendidos, no es posible determinar en qué momento del video ingresaron a la tienda, ya que se registra el tiempo hasta que fueron atendidos o se retiraron del local. Se define el valor esperado de la atención como:

$$E(y_i^*) = Pr(t_{eng} < x \cap t_l > t_{eng} \mid t_{eng} > \Delta, t_l > \Delta) = Pr(A \cap B \mid C) \quad (13)$$

Luego aplicando el teorema de Bayes, se estima la probabilidad de la siguiente forma:¹¹

$$= Pr(t_{eng} < x \mid t_l > t_{eng}, t_{eng} > \Delta, t_l > \Delta) * Pr(t_l > t_{eng} \mid t_{eng} > \Delta, t_l > \Delta) = Pr(A \mid B, C) * P(B \mid C) \quad (14)$$

¹¹ Teorema de Bayes define la probabilidad condicional de A_i dado B, para cualquier i como:

$$Pr(A_i \mid B) = \frac{Pr(A_i \cap B)}{P(B)}$$

Utilizando los datos de los videos de seguimiento completo, la probabilidad se calcula de la siguiente forma:

$$Pr(A | B, C) * Pr(B|C) = \frac{\text{nro. de clientes atendidos } [\Delta, x]}{\text{nro. de clientes que no se van y no son atendidos antes de } \Delta}$$

El tiempo que el cliente lleva en tienda se obtiene de los registros no observados, siendo $x=8$ minutos en este caso. Utilizando (14) se estimó esta probabilidad para cada cliente, usando este resultado como el valor de la atención. La tasa de atención para la muestra agregada presenta una media del 14%, con un intervalo de confianza entre [0,134, 0,145] al 95% de confiabilidad, asumiendo que la media se distribuye normalmente en la muestra.

7.3. Modelo para corregir Endogeneidad

Como se mencionó anteriormente la tasa de atención puede presentar un problema de endogeneidad por sesgo de causalidad simultánea. No se sabe si la atención influye en que el cliente compre o al revés, dado que el cliente se decide previamente a comprar, toma la iniciativa y busca contactar a un vendedor. Si efectivamente la tasa de atención es una variable endógena y no se corrige, los estimadores pueden ser sesgados e inconsistentes, llevando posiblemente a conclusiones erróneas.

Para corregir el problema de la endogeneidad y error de medición de la tasa de atención se utilizó el método de mínimo cuadrados por dos etapas (MC2E) para estimar el modelo mediante variables instrumentales. El instrumento $ZR\hat{Att}_{jt}$ utilizado debe cumplir con las siguientes 2 condiciones:

1. $ZR\hat{Att}_{jt}$ debe ser relevante: $Corr(ZR\hat{Att}_{jt}, Eng_{jt}) \neq 0$
2. $ZR\hat{Att}_{jt}$ debe ser exógeno o valido: $Corr(ZR\hat{Att}_{jt}, \xi_{jt}) = 0$

Un candidato para ser instrumento es el ratio entre vendedores y clientes (SR_{jt}). Cuando la tienda se encuentra congestionada (sin variar el número de vendedores) el ratio toma un valor elevado y la tasa de atención disminuye, ya que baja la capacidad por parte de los vendedores para asistir a los clientes adicionales que se encuentran en la tienda. Resultado de esto, clientes que podrían haber sido atendidos bajo condiciones normales no pueden ser atendidos cuando la tienda está congestionada.

En caso contrario, si existe poca congestión el ratio toma un valor bajo y la tasa de atención aumenta, tal que existe la capacidad por parte de los vendedores para asistir a los clientes

Muestra Video	Atención	#Clientes	#Vendedores	Ratio	Pred. Atención
1	1.0	3.0	1.0	3.00	0.72
2	0.0	6.0	3.0	2.00	0.57
3	1.0	5.0	3.0	1.67	0.51
4	0.0	1.0	2.0	0.50	0.33
5	1.0	4.0	2.0	2.00	0.57
Promedio	0.6	3.8	2.2	1.83	0.54

Tab. 5: *Ejemplo instrumento*. Fuente: Elaboración propia

en tienda, siendo el instrumento relevante. El instrumento es válido porque su valor es independiente a la intención de compra que tenga el cliente (que puede llevar a contactar a un vendedor).

Para construir el instrumento se modeló lo siguiente (datos a nivel individual):

$$\text{logit}(Att_{ijt}) = \gamma_1 + \gamma_2 SR_i + \gamma_3 \gamma_t + \gamma_4 \delta_{dt} + \gamma_5 \sum_w \eta_t^w + \gamma_6 \alpha_{jt}^c + \gamma_7 \alpha_{jt}^r + \gamma_8 \alpha_{jt}^f + \gamma_9 \beta_{jt}^c + \gamma_{10} \beta_{jt}^h + \gamma_{11} s_j + \varepsilon_i \quad (15)$$

Sea $n_{ijt} = \gamma_1 + \gamma_2 SR_i + \gamma_3 \gamma_t + \gamma_4 \delta_{dt} + \gamma_5 \sum_w \eta_t^w + \gamma_6 \alpha_{jt}^c + \gamma_7 \alpha_{jt}^r + \gamma_8 \alpha_{jt}^f + \gamma_9 \beta_{jt}^c + \gamma_{10} \beta_{jt}^h + \gamma_{11} s_j$

Del modelo anterior se estiman los parámetros asociados y se predice la atención de la siguiente forma:

$$\text{Prob}(\text{Atención} = 1) = \frac{1}{1 + \exp[-(n_{ijt} + \varepsilon_i)]} \quad (16)$$

Mediante (16) se obtiene la predicción de la atención para cada cliente de la muestra. Se utilizará como instrumento el promedio de todas las predicciones para una tienda j y un día t en particular ($ZRAtt_{jt}$). En la tabla 5 se presenta un ejemplo para 5 muestras recopiladas para un día específico. La última columna es el resultado de la estimación mediante la ecuación (16), resultando como promedio de estas predicciones 0,54, valor cercano a 0,6 de la tasa de atención estimada con esta muestra para un día en particular. En este ejemplo se puede ver que el instrumento es un buen estimador de la atención y los valores que puede tomar se encuentran entre 0 y 1, misma escala que la atención. La crítica que puede existir es que $\text{logit}(Att_{ijt})$ puede ser una función no lineal con respecto al ratio. Se probaron distintas

especificaciones para este modelo, pero en todos los casos los ajustes eran similares (ver apéndice número 2).

Ya teniendo un instrumento válido, se procedió a estimar el modelo mediante el método de mínimo cuadrados por 2 etapas (MC2E), procedimiento el cual se explica a continuación:

Se realizó una regresión lineal de $ZRAtt_{jt}$ sobre $Ratt_{jt}$ (primera etapa) con los datos agrupados por día (t) y tienda (j):

$$Ratt_{jt} = \Pi_0 + \Pi_{1*}ZRAtt_{jt} + \text{controles}_2 + \nu_{jt} \quad (17)$$

En donde : $\text{controles}_2 = \pi_2Cust_{jt} + \pi_3\gamma + \pi_4\delta_{dt} + \pi_5 \sum_w \eta_t^w + \pi_6\alpha_{jt}^c + \pi_7\alpha_{jt}^r + \pi_8\alpha_{jt}^f + \pi_9\beta_{jt}^c + \pi_{10}\beta_{jt}^h + \pi_{11}s_j$

Luego se estiman los parámetros anteriores:

$$\hat{Ratt}_{jt} = \hat{\Pi}_0 + \hat{\Pi}_{1*}ZRAtt_{jt} + \text{controles}_2, \text{ donde } i = 1, \dots, n \quad (18)$$

Para finalmente reemplazar \hat{Ratt}_{jt} en conjunto con el residuo de la estimación en la regresión de interés (segunda etapa):

$$\text{Logit}(\text{Conv}_{jt}) = \beta_1 + \beta_2\hat{Ratt}_{jt} + \beta_3Cust_{jt} + \text{controles}_3 + \xi_{jt} \quad (19)$$

Donde : $\text{controles}_3 = \beta_4\gamma_t + \beta_5\delta_{dt} + \beta_6 \sum_w \eta_t^w + \beta_7\alpha_{jt}^c + \beta_8\alpha_{jt}^r + \beta_9\alpha_{jt}^f + \beta_{10}\beta_{jt}^c + \beta_{11}\beta_{jt}^h + \beta_{12}s_j$

Luego de haber estimado el modelo mediante el método de mínimos cuadrados por dos etapas, es posible estimar la magnitud del efecto de la atención sobre la probabilidad de ocurrencia de conversión, para un día y tienda específico, como sigue a continuación:

$$\text{Prob}(\text{Conv}_{jt} = 1) = \frac{1}{1 + \exp[-(\epsilon(\hat{w}_{jt}) + \xi_{jt})]} \quad (20)$$

En donde $\epsilon(\hat{w}_{jt}) = \beta_1 + \beta_2\hat{Ratt}_{jt} + \beta_3Cust_{jt} + \beta_4\gamma_t + \beta_5\delta_{dt} + \beta_6 \sum_w \eta_t^w + \beta_7\alpha_{jt}^c + \beta_8\alpha_{jt}^r + \beta_9\alpha_{jt}^f + \beta_{10}\beta_{jt}^c + \beta_{11}\beta_{jt}^h + \beta_{12}s_j$

La ecuación (20) permite estimar el cambio que se produce en la conversión al variar la tasa de atención, manteniendo constantes los demás parámetros. Cuantificando este efecto, se puede realizar un posterior análisis del impacto económico que conlleva este efecto.

7.4. Resultados

Los datos utilizados en esta sección corresponden a la tienda de retail de vestuario femenino. Los errores estimados que se reportan se calcularon bajo un estimador robusto (White), válidos en caso de heterocedasticidad. En la tabla 6, se muestra el resultado del modelo principal (ecuación 19). En el modelo (1) se estimó utilizando como variable instrumental $ZR\hat{Att}_{jt}$ mientras que el instrumento para el modelo (2) se utilizó SR_{jt} , siendo el modelo (1) el que presenta un coeficiente de determinación (estadístico R^2) marginalmente más alto, el cual explica un 27 % de la variación de la variable dependiente por medio de las variables independientes (versus un 24 %). El hecho de que este estadístico no sea elevado se debe a la complejidad de explicar o predecir la conversión.

El parámetro asociado a la tasa de atención es positivo y estadísticamente significativo con un 95 % de confiabilidad, es decir, el aumento en la tasa de atención genera un aumento en la conversión, validando así la primera hipótesis. Como se transformó la conversión, no se puede inferir directamente el efecto en términos cuantitativos, por ello para estimarlo se calculó el efecto marginal al variar la tasa de atención sobre la conversión. La estimación implica que al aumentar la fracción de clientes atendidos promedios desde la media (0,505) en una desviación estándar (0,11) incrementa la conversión un 4,5 %¹². La mayoría de los controles no son significativos, pero sí varios de estos son marginalmente significativos (p-valor esta entre 5 % y 10 %), por ejemplo para los días martes se tiene un conversión promedio más baja y una de las tiendas tiene una conversión promedio mayor que la otra.

En la tabla 7, se utiliza el mismo modelo, salvo que se cambia la variable dependiente por el *ticket* promedio (1) y ventas por entrada (2). En este caso como la variable dependiente no se transforma por un modelo *logit* es más directo estimar el efecto cuantitativo. En (1) el parámetro asociado a *Ticket* no es significativo con un 95 % de confiabilidad, concluyendo que la tasa de atención no afecta al ticket promedio. Al no variar el ticket promedio, las ventas deberían aumentar el mismo porcentaje que la conversión.

En el modelo (2) usando como variable dependiente ventas por entrada de clientes, se obtiene el parámetro asociado positivo y significativo a un 95 % de confiabilidad. Específicamente, aumentando la fracción de clientes atendidos promedio desde la media, en una desviación estándar, el aumento de ventas por entrada es de 7,4 dólares,¹³ validando la segunda parte de la primera hipótesis. Se desprende que para los días de lluvia las ventas por entrada son 2,1 dólares superior al promedio, en cambio para los días en que la temperatura es muy baja el promedio de ventas por entrada es 1,67 dólares más bajo que el promedio. Como se

¹² Este cálculo es $\Delta Conversion = d(Prob(y_i = 1))/dx * \Delta REng$

¹³ Este se calcula de la siguiente forma : $\beta * \Delta RAtt = 67,45 * 0,11$

Tab. 6: Modelo principal, utilizando VI

	(1)		(2)	
	logitcr		logitcr	
(mean) RAtt	1.998*	(0.873)	2.217*	(0.897)
(mean) Cust	0.056	(0.044)	0.062	(0.046)
day==Monday	-0.123	(0.082)	-0.124	(0.084)
day==Saturday	-0.073	(0.084)	-0.073	(0.086)
day==Sunday	0.175	(0.133)	0.183	(0.133)
day==Thursday	-0.021	(0.087)	-0.028	(0.088)
day==Tuesday	-0.166	(0.093)	-0.171	(0.094)
day==Wednesday	0.143	(0.110)	0.145	(0.113)
isHoliday	-0.127	(0.108)	-0.138	(0.113)
clima==cloudy	-0.215	(0.133)	-0.227	(0.140)
clima==rain	0.221	(0.233)	0.233	(0.234)
cool	0.025	(0.297)	0.011	(0.305)
warm	-0.000	(0.111)	-0.011	(0.114)
store==506	-0.139	(0.0747)	-0.124	(0.077)
nweek==20	-0.347*	(0.174)	-0.357*	(0.182)
nweek==23	-0.528*	(0.220)	-0.560*	(0.233)
nweek==24	-0.524**	(0.181)	-0.521**	(0.189)
nweek==30	-0.359*	(0.179)	-0.365*	(0.185)
nweek==36	-0.441*	(0.198)	-0.449*	(0.205)
Constant	-1.753***	(0.521)	-1.865***	(0.539)
Observations	248		248	
R^2	0.270		0.244	
Adjusted R^2	0.112		0.080	

Standard errors in parentheses

* $p < 0,05$, ** $p < 0,01$, *** $p < 0,001$

obtuvo que el ticket promedio no aumenta ante un cambio en la tasa de atención, el aumento de las ventas por entrada se ve reflejado porque más clientes compran, ya que aumenta la conversión.

Tab. 7: Modelo secundario utilizando VI

	(1)		(2)	
	ticket		salesw	
(mean) RAtt	-1433.3	(1047.7)	67.45*	(30.53)
(mean) Cust	-35.71	(40.25)	3.265*	(1.599)
day==Monday	28.81	(73.81)	-0.813	(2.813)
day==Saturday	-5.933	(36.25)	-0.784	(2.689)
day==Sunday	-175.0	(133.5)	5.857	(4.566)
day==Thursday	6.402	(31.85)	-1.416	(2.979)
day==Tuesday	11.75	(35.84)	-2.807	(3.051)
day==Wednesday	-12.86	(38.17)	6.348*	(3.236)
isHoliday	19.82	(37.54)	-6.047	(3.551)
clima==cloudy	134.6	(95.84)	-3.816	(4.428)
clima==rain	-61.95	(111.3)	19.14**	(7.093)
cool	124.0	(140.8)	-15.19*	(7.622)
warm	76.14	(72.51)	-1.671	(4.021)
store==506	-177.3	(124.8)	-3.165	(2.450)
Constant	1179.9	(773.9)	-1.766	(19.04)
Observations	328		248	
R^2	0.355		0.250	

Standard errors in parentheses

* $p < 0,05$, ** $p < 0,01$, *** $p < 0,001$

La segunda hipótesis postula que un aumento en la dotación de personal por medio del ratio entre clientes y vendedores impacta positivamente en la conversión. Esta especificación tiene como variable independiente el ratio entre clientes y vendedores junto a los respectivos controles y como variable dependiente la atención transformada por medio de un logit (tabla 7.1). En este modelo se tienen 12041 observaciones (datos no están agrupados) y el parámetro del ratio es estadísticamente significativo incluso con un 99,9% de confiabilidad, teniendo signo negativo, lo que implica que al agregar un vendedor la tasa de atención aumenta, resultado acorde a la intuición.

Utilizando el mismo mecanismo para estimar el efecto como se hizo en la segunda hipótesis, se tiene que al agregar un vendedor el porcentaje de clientes atendidos aumenta un 2,3% y una vez más, calculando el efecto marginal, el aumento en conversión es de un 0,9% (se utilizan los mismos parámetros del modelo (1) de la tabla), validando así esta hipótesis.

Ahora bien, ¿existirán diferencias entre las tiendas? En la tabla 8 se muestra el modelo de conversión, estimado para cada tienda. Para la primera tienda el parámetro asociado a

la atención no es significativo, mientras que para la segunda tienda sí lo es con 99% de confiabilidad, es decir, existen diferencias en el efecto de atención sobre la conversión entre las tiendas, confirmando la tercera hipótesis.

(1)	
Tab.7.1 Atención	
eq1	
sr	-0.119*** (0.0120)
day==Monday	0.0594 (0.0724)
day==Saturday	0.00113 (0.0719)
day==Sunday	0.0344 (0.0809)
day==Thursday	0.149* (0.0701)
day==Tuesday	0.168* (0.0727)
day==Wednesday	0.0140 (0.0718)
isHoliday	0.0805 (0.0787)
cool	0.136 (0.359)
warm	0.238* (0.0953)
clima==cloudy	0.132 (0.0892)
clima==rain	-0.234 (0.222)
hourofday==11	-0.479*** (0.0972)
hourofday==12	-0.497*** (0.0960)
hourofday==13	-0.492*** (0.0948)
hourofday==14	-0.609*** (0.0956)
hourofday==15	-0.733*** (0.0957)
hourofday==16	-0.627*** (0.0964)
hourofday==17	-0.567*** (0.0996)
hourofday==18	-0.397*** (0.108)
hourofday==19	-0.183 (0.156)
hourofday==20	13.63 (794.9)
store==506	-0.234*** (0.0408)
Constant	4.870 (9.948)
Observations	12041
<i>AIC</i>	16440.2
<i>BIC</i>	16891.4
ll	-8159.1

Standard errors in parentheses

* $p < 0,05$, ** $p < 0,01$, *** $p < 0,001$

Tab. 8: Modelo de conversión por tienda

	(1)		(2)	
	logitcr		logitcr	
(mean) RAtt	2.250	(1.842)	2.529**	(0.932)
(mean) Cust	0.126	(0.144)	0.0236	(0.0669)
day==Monday	-0.0296	(0.116)	-0.194	(0.117)
day==Saturday	-0.0143	(0.0939)	-0.110	(0.125)
day==Sunday	0.187	(0.272)	0.306	(0.208)
day==Thursday	-0.176	(0.125)	0.0419	(0.109)
day==Tuesday	-0.201	(0.126)	-0.163	(0.121)
day==Wednesday	0.103	(0.121)	0.179	(0.141)
isHoliday	0.0256	(0.121)	-0.275	(0.157)
clima==cloudy	-0.252	(0.163)	-0.261*	(0.130)
clima==rain	0.0205	(0.263)	0.231	(0.193)
cool	0.888***	(0.233)	-0.347	(0.281)
warm	-0.0991	(0.170)	-0.0589	(0.115)
o.store==506	0	(.)	0	(.)
nweek==12	-0.559*	(0.226)	-0.214	(0.296)
nweek==19	0.0419	(0.162)	-0.455*	(0.226)
o.nweek==23	0	(.)	-0.566*	(0.250)
o.nweek==24	0	(.)	-0.451*	(0.202)
nweek==29	-0.156	(0.186)	-0.504*	(0.204)
nweek==30	-0.144	(0.153)	-0.519*	(0.231)
nwe Constant	-2.194*	(1.060)	-1.903***	(0.546)
Observations	107		141	
R^2	0.115		0.434	

Standard errors in parentheses

* $p < 0,05$, ** $p < 0,01$, *** $p < 0,001$

	department / Month of fdate											
	Auto Parts			Seasonal			Sporting Goods			Tools		
	January	February	March	January	February	March	January	February	March	January	February	March
Sales per walk-in	3.22	3.45	3.21	0.43	0.57	1.50	2.57	2.25	2.62	1.36	1.91	2.12
Conversion	0.130	0.141	0.13	0.027	0.034	0.046	0.055	0.051	0.058	0.047	0.057	0.058
Engagement				6%	14%	12%	33%	29%	27%	10%	22%	15%
Entrance	63,618	51,370	62,703	63,618	51,370	62,703	63,618	51,370	62,703	63,618	51,370	62,703
Sales (\$)	199,529	175,220	201,879	27,877	29,243	105,242	166,776	117,111	171,302	86,587	98,872	133,773
Transactions	8,241	7,246	7,864	1,700	1,736	2,902	3,516	2,609	3,664	2,979	2,921	3,618
Delta Conversion	-0.4%			2.0%			0.3%			1.1%		

Tab. 9: *Resultados del experimento realizado*. Fuente: Departamento de Marketing de la empresa en estudio y Scopix Solutions.

8. Medición de impacto mediante experimentos de campo

En esta sección se desarrolló una metodología basada en un experimento de campo para evaluar el impacto de la calidad de servicio al cliente en las ventas, llevando a cabo un experimento en otra tienda de retail por medio del método diferencia en diferencias. Éste consiste en observar resultados para dos grupos en dos períodos de tiempo, uno de los grupos se expone a un tratamiento en el segundo período de medición, pero no se aplica nada durante el primer período. El segundo grupo no es expuesto al tratamiento durante ninguno de los períodos.

Para el experimento se utilizó la tienda de retail en Canadá, la cual tiene mayor tamaño que la otra cadena analizada en la tesis, pero el servicio está basado en atención vía vendedores para asistir a los clientes. El experimento consistió en la medición por 3 meses (Enero, febrero y marzo del 2013) de la tasa de atención en distintos departamentos, de los cuales en dos de estos se tomaron acciones aumentando la tasa de atención, utilizando al resto como grupo de control. La métrica que se utilizó para hacer comparaciones fue conversión en la tienda. La entrada de personas se midió a nivel de tienda y no por departamento¹⁴. Las acciones se llevaron a cabo en los departamentos *Seasonal* y *Tools*, en los cuales se aumentó al tasa de atención un 6% y 5% respectivamente, mientras que en los departamentos *Auto Parts* y *Sporting Goods*, no se tomaron acciones en la tasa de atención. En los departamentos *seasonal* y *Tools*, la conversión aumentó en un 2% y un 1,1% respectivamente, mientras que en los departamentos que no se tomaron acciones la conversión casi no varió.

Finalmente se corrobora que existe una correlación positiva entre la tasa de atención y la conversión, incluso de magnitudes parecidas.

¹⁴ Se asumió que si aumenta el tráfico en tienda, también aumentará en cada departamento con la misma proporción

9. Análisis de sensibilidad

La motivación para llevar a cabo este análisis es mostrar que el modelo es robusto ante cambios tanto en la especificación del modelo como en la selección de la muestra. Es primordial este análisis, principalmente por que si se demuestra que el modelo utilizado es robusto, los resultados tendrán más credibilidad. Se estimó el modelo usando los controles por separado con el objetivo de analizar el impacto que tiene cada uno de estos en la tasa de atención.

En la tabla¹⁵ 10 se muestran 3 especificaciones del modelo desarrollado, es decir tomando como variable dependiente el logit de la conversión y usando como variable dependiente la tasa de atención más controles, siendo instrumentada por la predicción del ratio. El primer modelo se controla solo por días de la semana, donde los días Lunes y Martes son significativos variando el parámetro de la tasa de atención. Al controlar por día, el efecto de la atención sobre la conversión se vuelve mayor. Controlar por feriados y clima por separados no afecta significativamente el valor obtenido del parámetro atención.

En la tabla 11 se controla por número de semana, tienda y cliente. De estas especificaciones, el control por tienda afecta el parámetro asociado a la tasa de atención, disminuyéndolo casi a la mitad. Por lo tanto controlar por tienda implica un menor efecto de la atención sobre conversión.

Como se muestra en la tabla 12, en la primera especificación no se utilizan controles, mientras que en la segunda se utilizan todos los controles disponibles. Es interesante notar que al no utilizar controles, el parámetro de la variable atención no varía con respecto al modelo principal, sin embargo, se vuelve significativa a un nivel de confianza del 99 %. Lo rescatable de la segunda especificación es que a pesar de utilizar una gran cantidad de controles, el parámetro de atención sigue siendo significativo.

En la tabla 13 se muestran 3 modelos en los cuales se modificó la cantidad de muestras para la tasa de atención. En el primer modelo se utiliza una muestra diaria mayor a 5 registros por día, en el segundo mayor a 10 registros por día y en el tercero mayor a 15 registros por día, lo cual es coherente con el número de observaciones, disminuyendo a medida que se aumenta la exigencia en este criterio (menor error). Al aumentar la exigencia, el parámetro estimado para la tasa de atención aumenta en conjunto con su significancia, pero por el contrario el R^2 disminuye.

El modelo seleccionado buscó equilibrar entre cantidad de la muestra y error, evitando tener muy pocos registros por día y no perder mucha varianza explicada para la variable dependiente por parte de las variables independientes.

¹⁵ Las tablas se encuentran en el apéndice.

El primer modelo de la tabla 14 es la especificación considerando todos los registros, mientras que en el segundo modelo no se consideran los registros no observados. Se desprende de los resultados que el sesgo de selección disminuye el efecto de la atención sobre la conversión, incluso la variable atención se vuelve no significativa, resultado que muestra la importancia de haber solucionado este problema.

Por último, en la tabla 15 se estimaron los modelos sin usar variables instrumentales, utilizando todos los registros (1) y sin considerar los registros no observados (2). Se repite lo anterior en donde el sesgo de selección disminuye el efecto. El problema de endogeneidad tiene un impacto más significativo en la conversión, es decir, al utilizar variables instrumentales el efecto de la atención sobre la conversión aumenta, e incluso vuelve la variable significativa con respecto a no utilizar variables instrumentales.

10. Implicancias para la gestión de la fuerza de venta

Los resultados obtenidos en la parte anterior permiten determinar que existen beneficios para la tienda si el nivel de atención es aumentado. Ahora bien, se realizó un análisis económico tomando como base los siguientes 3 casos:

1. Se agrega un vendedor a tiempo completo. Esto implica un escenario en donde se adiciona un trabajador en todo horario en que la tienda se encuentra en funcionamiento.
2. Se agrega un vendedor a tiempo parcial. En este escenario se adicionó un trabajador solo en días y horas de mayor tráfico y ratio. Específicamente se tomaron los días y horas donde el ratio está sobre 2, que corresponde a los Días Viernes, Sábado y Domingo entre las 12 PM y las 4 PM.
3. No se agrega vendedor adicional, pero sí se aumenta tasa de atención. En este caso, el aumento se efectúa mediante una capacitación a los vendedores, a quienes en un caso hipotético se les entrega instrucciones relacionadas en preocuparse más por asistir clientes en el caso que requieran ayuda para encontrar algún producto, dejando un poco de lado otras tareas como por ejemplo ordenar productos en tienda. Se simuló un aumento de la tasa de atención en una desviación estándar, valor dentro del rango de atención para un día promedio.

El sueldo promedio para un vendedor de retail a nivel país en Estados Unidos es de 12,17 dólares por hora¹⁶. Específicamente para los vendedores de retail en la categoría de vestimen-

¹⁶ Ocupación del empleo y salario, Estados Unidos, Mayo del 2012, <http://www.bls.gov/oes/current/oes412031.htm>.

ta, el salario es más bajo llegando a un promedio de 10,35 dólares. Las tiendas analizadas se encuentran en San Francisco, para esta ciudad solo fue posible obtener el sueldo promedio para vendedores de retail, llegando a \$13,63 por hora, es decir \$1,46 más alto que el promedio nacional. Con los datos anteriores se estimó el sueldo promedio para un vendedor de retail en la categoría de vestimenta para San Francisco:

$$\text{Sueldo/hora} = \$10,35 + (\$13,63 - \$12,17) = \$11,81$$

El tráfico promedio en tienda es 213,71 clientes al día, con un ticket promedio de 132,25 dólares. En el primer caso el aumento en la conversión es del 0,9% lo que trae consigo un aumento en las ventas de 257,05 dólares. Al agregar un vendedor se generarían utilidades para la tienda si el margen obtenido por los productos¹⁷ es más de un 46%¹⁸. En el caso de agregar un vendedor *part-time*, el aumento promedio es de 301 dólares al día, ya que la conversión aumenta en un 1,1%. En este caso, la tienda debería tener un margen superior al 39% para que esta alternativa sea rentable. En el último caso, el aumento promedio diario es de 1285 dólares, tomando como base que la conversión aumenta en un 4,5%.

No cabe dudas que, combinando estas opciones el beneficio resultante es positivo (margen necesario es mucho menor) e incluso se mejoraría la calidad de servicio en la tienda por medio del aumento de las horas disponibles por parte del personal para ordenar productos, y la asistencia a los clientes, es decir, se mejoraría el nivel de servicio. Según un estudio¹⁹ el margen en la industria del retail para el segmento de ropa fue de un 36,57% para el año 2012. Si se combina el caso de agregar un vendedor en conjunto con una capacitación se logra avalar la última hipótesis, es decir, existe oportunidad para asignar personal adicional, aumentando las utilidades de la tienda.

¹⁷ Se define como el valor de venta menos los costos asociados al producto

¹⁸ $\$11,81 * 10/257,05$

¹⁹ Publicado por Demand Media, ref: <http://yourbusiness.azcentral.com/profit-margin-retail-clothes-26395.html>

11. Conclusiones

El análisis de videos permitió entender el mecanismo por el cual los vendedores pueden ayudar al aumento de las ventas en la tienda, por medio de un aumento en la conversión. Con este estudio, una cadena de retail puede tomar decisiones basándose en información sobre acontecimientos que ocurren antes de que se realice la venta, es decir, tomando conocimiento de las variables más relevantes que afectan al proceso de compra.

Por medio de un análisis exploratorio, se determinó que las acciones por parte del vendedor que más afectan la conversión son: estar disponible en el caso que algún cliente necesite ayuda (ya sea para encontrar algún producto, preguntar por un precio, solicitar un producto u otra talla) y facilitar o persuadir a un cliente para que ingrese al probador. Se identificó que estas acciones permitirían a la tienda cambiar el enfoque de atención en el cual se basan actualmente, es decir, ordenar productos y atender clientes con la misma prioridad.

Se postulaba por parte de la tienda que sus vendedores eran capaces de reconocer a los clientes más rentables, no pudiendo verificar esta afirmación. Si bien la conversión en la muestra en el caso que un vendedor tomó la iniciativa y existió atención es más alta con respecto al caso cuando el cliente no recibe atención, la diferencia no es estadísticamente significativa. Un aspecto que llama mucho la atención es que no existe un muy buen manejo del personal de ventas, ya que la cantidad de vendedores durante la semana no varía en la misma proporción que el tráfico, disminuyendo la capacidad para atender clientes en los días con mayor demanda.

Para validar las hipótesis se desarrolló un modelo econométrico, específicamente transformando la variable dependiente (tomando el logit), para luego estimar el modelo usando variables instrumentales por el método de mínimos cuadrados en dos etapas. La razón de aplicar variables instrumentales responde a un posible sesgo de endogeneidad y error de muestreo que puede presentar la tasa de atención, ya que no es claro el efecto de atención sobre la probabilidad de compra, es decir, los clientes pueden buscar contactar a un vendedor ya habiendo tomado la decisión de comprar o los mismos vendedores podrían elegir a qué cliente contactar (si supieran cuales son los clientes mas rentables). Por lo tanto, estas decisiones pueden estar relacionadas con características del cliente que no son observables y que determinan la intención de compra, siendo no necesariamente un efecto causal.

Otro problema abordado por la metodología fue el sesgo en la selección de la muestra para la variable atención, en donde se incluyeron los registros en que no se sabía qué había ocurrido con el cliente, estimando la probabilidad de ser atendido con ciertas condiciones. La metodología abordó los problemas anteriormente mencionados con el fin de evitar sesgo en las estimaciones de los parámetros, lo que podría traer consigo conclusiones erróneas.

De esta forma, se validó la primera hipótesis, ya que al aumentar la tasa de atención en un 11 % la conversión lo hace en un 4,5 %. El aumento en la tasa de atención equivale a una desviación estándar, es decir, la tasa de atención está dentro de un rango factible de alcanzar para un día promedio. Si se considera como variable dependiente las ventas por entrada, se tiene un aumento de 7,4 dólares por cada persona que entra en la tienda²⁰. Esta última métrica es bastante potente, ya que resume dos efectos importantes que dependen del compromiso de conversión: cantidad de personas que compran y promedio de dinero gastado por cada cliente. Si bien al variar la tasa de atención las ventas por entrada aumentan, esto se debe a que existe un aumento en la conversión (aumenta el número de personas que compran) y no porque los clientes gastan más en cada compra, donde el ticket promedio no se ve afectado.

Un resultado importante fue que a pesar de que las tiendas presentan características similares en la forma de atender a los clientes y en los estadísticos descriptivos, hay diferencias en la efectividad de la atención para convertir no compradores en compradores.

También se probó que al adicionar un vendedor la tasa de atención aumenta y por lo tanto la conversión, sin embargo el aumento de la tasa de atención es solo de un 2,3 %. Posiblemente lo anterior se puede deber a que el ratio entre clientes y vendedores se mantiene bajo durante la semana, por ende el efecto no es de gran magnitud al considerar todos los días. Este efecto se puede acentuar aún más si se agrega un vendedor en el período donde existe mayor demanda, obteniendo más beneficios económicos, en donde al agregar un vendedor los días viernes, sábados y domingos en las horas con mayor demanda (específicamente ratio entre clientes y vendedores mayor a dos) genera más ganancias a la tienda que al agregar un trabajador para toda la semana.

Se concluye que la tienda obtiene más beneficios si combina el agregar un trabajador con un entrenamiento del personal para elevar la atención (por medio de las acciones encontradas en el análisis exploratorio).

Fue posible validar los resultados por medio de un experimento realizado en una cadena de retail con características distintas. El experimento consistió en aumentar la tasa de atención en dos departamentos, mientras que otros dos se tomaron como grupo de control, es decir, no se varió la tasa de atención. Como resultado se obtuvo que los departamentos en donde se aumentó la tasa de atención también lo hizo la conversión (incluso la variación fue de magnitud parecida a los resultados del modelo desarrollado), mientras que en los departamentos en los cuales no se tomaron acciones la conversión casi no varió.

Finalmente se desarrolló un análisis de sensibilidad en donde se validó la robustez del modelo en términos de la especificación utilizada. Se encontró que la corrección de la endogeneidad

²⁰ Aumentando la tasa de atención en una desviación estándar sobre la media, es decir, un 11 %

(por medio de variables instrumentales) aumenta significativamente el valor del parámetro asociado a la atención en el modelo de conversión, mientras que la corrección del sesgo de selección también aumenta el efecto de atención sobre conversión pero en menor magnitud.

El estudio realizado tiene algunas limitaciones relacionadas con las mediciones. La primera limitación es la pérdida de muestras por la existencia de puntos ciegos, es decir, si un cliente se va a un punto ciego y se queda ahí un rato (más de 10 segundos), el cliente de ese registro no es tomado como válido, lo que significa buscar otro cliente en la misma muestra si esto ocurre. Lo anterior trae como consecuencia el uso de más recursos para medir la métrica asociada.

Otra limitación es la medición de los videos de seguimiento completo, en los cuales se le hace seguimiento a los clientes. Como se registra el trayecto completo del cliente, toma tiempo medir a cada cliente, resultando costoso tomar un tamaño de muestra considerable. Adicionalmente otra limitación es que estos videos de seguimiento completo no se pueden obtener en todas las horas durante el día, sino que solo a una hora en el día, por temas de capacidad.

La metodología desarrollada en esta tesis muestra un importante avance en esta área, ya que no existe literatura asociada a combinación de métricas de desempeño utilizando esta tecnología con datos del punto de venta. Para trabajos futuros se recomienda tomar estos resultados e integrarlos para el diseño de un modelo de dotación de personal, el cual tenga como objetivo reubicar a los vendedores eficientemente de acuerdo a la demanda presentada en la tienda.

En el análisis de la sección 10 se mostraron algunos resultados de cómo el estudio empírico realizado puede apoyar decisiones de dotación, lo que permitiría a los gerentes de local ubicar su fuerza de ventas de manera más óptima (basados en la maximización de utilidades). También se recomienda llevar el análisis en una gran cantidad de tiendas para identificar cómo se diferencian éstas en términos de impacto de atención sobre conversión.

Por último, se propone la implementación de los resultados obtenidos en la medición de impacto utilizando videos de seguimiento (completo o parcial), para luego medir los efectos asociados a esta implementación.

12. Bibliografía

Referencias

- [1] P Conroy and S Bearse. Customer conversion. *The Changing*, 2006.
- [2] Marshall Fisher, Marshall L Fisher, and Ananth Raman. *The new science of retailing: How analytics are transforming the supply chain and improving performance*. Harvard business press, 2010.
- [3] Marshall L Fisher, Jayanth Krishnan, and Serguei Netessine. Retail store execution: An empirical study. *University of Pennsylvania, the Wharton School and Research Center: Operations and Information Management Department*. Retrieved December, 1:2006, 2006.
- [4] Marshall L Fisher, Jayanth Krishnan, and Serguei Netessine. Are your staffing levels correct? *International Commerce Review*, 8(2):110–115, 2009.
- [5] John Graham and Campbell Harvey. The equity risk premium in january 2006: Evidence from the global cfo outlook survey. *Available at SSRN 871105*, 2005.
- [6] Yina Lu, Andrés Musalem, Marcelo Olivares, and Ariel Schilkrut. Measuring the effect of queues on customer purchases. *Management Science*, 2013.
- [7] Puneet Manchanda, Peter E Rossi, and Pradeep K Chintagunta. Response modeling with nonrandom marketing-mix variables. *Journal of Marketing Research*, pages 467–478, 2004.
- [8] Vidya Mani, Saravanan Kesavan, and Jayashankar M Swaminathan. Understaffing in retail stores: Drivers and consequences. Technical report, Working Paper, Smeal College of Business, Pennsylvania State University, State College, PA, 2011.
- [9] John Neter, William Wasserman, Michael H Kutner, et al. *Applied linear statistical models*, volume 4. Irwin Chicago, 1996.
- [10] Serguei Netessine, Marshall L Fisher, and Jayanth Krishnan. Labor planning, execution, and retail store performance: An exploratory investigation. *University of Pennsylvania*, 2010.
- [11] Olga Perdikaki, Saravanan Kesavan, and Jayashankar M Swaminathan. Effect of traffic on sales and conversion rates of retail stores. *Manufacturing & Service Operations Management*, 14(1):145–162, 2012.

- [12] Jagdish N Sheth, Bruce I Newman, and Barbara L Gross. Why we buy what we buy: a theory of consumption values. *Journal of business research*, 22(2):159–170, 1991.
- [13] Harish Sujan, Barton A Weitz, and Mita Sujan. Increasing sales productivity by getting salespeople to work smarter. *The Journal of Personal Selling and Sales Management*, pages 9–19, 1988.
- [14] Zeynep Ton. The effect of labor on profitability: The role of quality. Technical report, Harvard Business School, 2009.
- [15] Zeynep Ton and Ananth Raman. The effect of product variety and inventory levels on retail store sales: A longitudinal study. *Production and Operations Management*, 19(5):546–560, 2010.
- [16] Kenneth Train. *Discrete choice methods with simulation*. Cambridge university press, 2009.

13. Anexos

Anexo A: Análisis de sensibilidad

Tab. 10: Modelo Logit utilizando controles de día, feriados y clima

	(1)		(2)		(3)	
	logitcr		logitcr		logitcr	
(mean) RAtt	2.624***	(0.496)	2.049***	(0.445)	2.022***	(0.433)
day==Monday	-0.182*	(0.0898)				
day==Saturday	-0.0124	(0.0848)				
day==Sunday	0.164	(0.130)				
day==Thursday	-0.0961	(0.109)				
day==Tuesday	-0.199*	(0.101)				
day==Wednesday	0.0497	(0.119)				
isHoliday			0.0755	(0.0856)		
clima==cloudy					-0.190	(0.120)
clima==rain					-0.0971	(0.205)
cool					0.173	(0.207)
warm					-0.158	(0.114)
Constant	-2.286***	(0.233)	-2.059***	(0.222)	-1.850***	(0.239)
Observations	248		248		248	

Standard errors in parentheses

* $p < 0,05$, ** $p < 0,01$, *** $p < 0,001$

Tab. 11: Modelo Logit utilizando controles de tienda, cliente y número de semana

	(1)	(2)	(3)			
	logitcr		logitcr			
(mean) RAtt	1.739***	(0.488)	1.035*	(0.502)	2.222***	(0.451)
nweek==9	-0.347**	(0.109)				
nweek==13	-0.262	(0.145)				
nweek==14	-0.301*	(0.133)				
nweek==15	-0.252*	(0.125)				
nweek==19	-0.262*	(0.111)				
nweek==20	-0.408***	(0.114)				
nweek==23	-0.497*	(0.193)				
nweek==24	-0.565***	(0.0736)				
nweek==26	-0.230*	(0.0989)				
nweek==27	-0.211*	(0.0978)				
nweek==28	-0.320***	(0.0960)				
nweek==29	-0.421***	(0.116)				
nweek==30	-0.455***	(0.116)				
nweek==34	-0.229*	(0.100)				
nweek==35	-0.235*	(0.108)				
nweek==36	-0.423**	(0.149)				
nweek==38	-0.307**	(0.110)				
store==506			-0.194**	(0.0601)		
(mean) Cust					0.0668*	(0.0263)
Constant	-1.676***	(0.247)	-1.454***	(0.272)	-2.361***	(0.250)
Observations	248		248		248	
R^2	0.175		0.092		.	

Standard errors in parentheses

* $p < 0,05$, ** $p < 0,01$, *** $p < 0,001$

Tab. 12: Modelo Logit con y sin controles

	(1)		(2)	
	logitcr		logitcr	
(mean) RAtt	2.006***	(0.440)	1.998*	(0.873)
(mean) Cust			0.0560	(0.0446)
day==Monday			-0.123	(0.0827)
day==Saturday			-0.0736	(0.0843)
day==Sunday			0.175	(0.133)
day==Thursday			-0.0215	(0.0872)
day==Tuesday			-0.166	(0.0931)
day==Wednesday			0.143	(0.110)
isHoliday			-0.127	(0.108)
clima==cloudy			-0.215	(0.133)
clima==rain			0.221	(0.233)
cool			0.0251	(0.297)
warm			-0.000574	(0.111)
store==506			-0.139	(0.0747)
nweek==20			-0.347*	(0.174)
nweek==23			-0.528*	(0.220)
nweek==24			-0.524**	(0.181)
nweek==30			-0.359*	(0.179)
nweek==36			-0.441*	(0.198)
Constant	-2.028***	(0.218)	-1.753***	(0.521)
Observations	248		248	
R^2	.		0.270	

Standard errors in parentheses

* $p < 0,05$, ** $p < 0,01$, *** $p < 0,001$

Tab. 13: Modelo Logit cambiando la muestra para la tasa de atención

	(1)		(2)		(3)	
	logitcr		logitcr		logitcr	
(mean) RAtt	1.280	(0.785)	1.998*	(0.873)	2.510*	(1.233)
(mean) Cust	0.0438	(0.0436)	0.0560	(0.0446)	0.0731	(0.0531)
day==Monday	-0.104	(0.0785)	-0.123	(0.0827)	-0.136	(0.0921)
day==Saturday	-0.0644	(0.0775)	-0.0736	(0.0843)	-0.0888	(0.0923)
day==Sunday	0.0961	(0.123)	0.175	(0.133)	0.166	(0.137)
day==Thursday	0.0118	(0.0811)	-0.0215	(0.0872)	-0.0747	(0.104)
day==Tuesday	-0.120	(0.0849)	-0.166	(0.0931)	-0.211	(0.113)
day==Wednesday	0.107	(0.106)	0.143	(0.110)	0.113	(0.119)
isHoliday	-0.103	(0.111)	-0.127	(0.108)	-0.161	(0.123)
clima==cloudy	-0.0874	(0.127)	-0.215	(0.133)	-0.186	(0.158)
clima==rain	-0.00202	(0.251)	0.221	(0.233)	0.218	(0.247)
cool	0.254	(0.330)	0.0251	(0.297)	-0.217	(0.263)
warm	0.0593	(0.104)	-0.000574	(0.111)	-0.0427	(0.127)
store==506	-0.180*	(0.0715)	-0.139	(0.0747)	-0.112	(0.0924)
nweek==20	-0.328*	(0.144)	-0.347*	(0.174)	-0.372	(0.196)
nweek==23	-0.421*	(0.182)	-0.528*	(0.220)	-0.599*	(0.279)
nweek==24	-0.528***	(0.139)	-0.524**	(0.181)	-0.487*	(0.194)
nweek==36	-0.414*	(0.165)	-0.441*	(0.198)	-0.458*	(0.220)
Constant	-1.492**	(0.497)	-1.753***	(0.521)	-2.053**	(0.694)
Observations	253		248		230	
R^2	0.285		0.270		0.198	

Standard errors in parentheses

* $p < 0,05$, ** $p < 0,01$, *** $p < 0,001$

Tab. 14: Modelos utilizando variables instrumentales

	(1)		(2)	
	logitcr		logitcr	
(mean) RAtt	1.998*	(0.873)		
(mean) Eng			1.409	(0.764)
(mean) Cust	0.0560	(0.0446)	0.0454	(0.0428)
day==Monday	-0.123	(0.0827)	-0.100	(0.0819)
day==Saturday	-0.0736	(0.0843)	-0.0607	(0.0802)
day==Sunday	0.175	(0.133)	0.164	(0.136)
day==Thursday	-0.0215	(0.0872)	0.00943	(0.0829)
day==Tuesday	-0.166	(0.0931)	-0.141	(0.0869)
day==Wednesday	0.143	(0.110)	0.139	(0.107)
isHoliday	-0.127	(0.108)	-0.147	(0.125)
clima==cloudy	-0.215	(0.133)	-0.120	(0.121)
clima==rain	0.221	(0.233)	0.195	(0.239)
cool	0.0251	(0.297)	0.0458	(0.276)
warm	-0.000574	(0.111)	0.0250	(0.109)
store==506	-0.139	(0.0747)	-0.210***	(0.0588)
nweek==20	-0.347*	(0.174)	-0.359*	(0.153)
nweek==23	-0.528*	(0.220)	-0.463*	(0.197)
nweek==24	-0.524**	(0.181)	-0.530***	(0.144)
nweek==29	-0.331	(0.181)	-0.361*	(0.162)
nweek==30	-0.359*	(0.179)	-0.383*	(0.159)
nweek==36	-0.441*	(0.198)	-0.420*	(0.175)
Constant	-1.753***	(0.521)	-1.608**	(0.530)
Observations	248		248	
R^2	0.270		0.293	

Standard errors in parentheses

* $p < 0,05$, ** $p < 0,01$, *** $p < 0,001$

Tab. 15: Modelos sin usar variables instrumentales

	(1)		(2)	
	logitcr		logitcr	
(mean) RAtt	0.576	(0.298)		
(mean) Eng			0.395	(0.258)
(mean) Cust	0.0173	(0.0361)	0.0139	(0.0362)
day==Monday	-0.117	(0.0850)	-0.111	(0.0857)
day==Saturday	-0.0741	(0.0834)	-0.0704	(0.0833)
day==Sunday	0.118	(0.131)	0.115	(0.131)
day==Thursday	0.0214	(0.0845)	0.0306	(0.0854)
day==Tuesday	-0.133	(0.0871)	-0.126	(0.0861)
day==Wednesday	0.125	(0.109)	0.124	(0.108)
isHoliday	-0.0606	(0.117)	-0.0653	(0.121)
clima==cloudy	-0.136	(0.119)	-0.108	(0.114)
clima==rain	0.138	(0.233)	0.130	(0.236)
cool	0.111	(0.304)	0.118	(0.297)
warm	0.0683	(0.107)	0.0762	(0.108)
store==506	-0.238***	(0.0565)	-0.259***	(0.0571)
nweek==9	-0.792*	(0.360)	-0.844*	(0.365)
nweek==12	-0.779*	(0.360)	-0.868*	(0.364)
nweek==13	-0.689	(0.350)	-0.765*	(0.357)
nweek==14	-0.721*	(0.345)	-0.766*	(0.350)
nweek==15	-0.671	(0.343)	-0.728*	(0.348)
nweek==16	-0.710	(0.363)	-0.755*	(0.367)
nweek==18	-0.669	(0.348)	-0.739*	(0.351)
nweek==19	-0.794*	(0.359)	-0.845*	(0.363)
nweek==20	-0.871*	(0.347)	-0.939**	(0.349)
nweek==23	-0.902**	(0.341)	-0.947**	(0.348)
nweek==24	-1.125**	(0.357)	-1.192**	(0.358)
nweek==26	-0.648	(0.346)	-0.711*	(0.349)
nweek==28	-0.822*	(0.348)	-0.886*	(0.351)
nweek==29	-0.920*	(0.358)	-0.994**	(0.360)
nweek==30	-0.905*	(0.356)	-0.976**	(0.358)
nweek==35	-0.700	(0.359)	-0.766*	(0.363)
nweek==36	-0.972**	(0.360)	-1.031**	(0.363)
nweek==38	-0.768*	(0.336)	-0.799*	(0.343)
Constant	-0.443	(0.440)	-0.330	(0.443)
Observations	248		248	
R^2	0.350		0.344	

Standard errors in parentheses

* $p < 0,05$, ** $p < 0,01$, *** $p < 0,001$

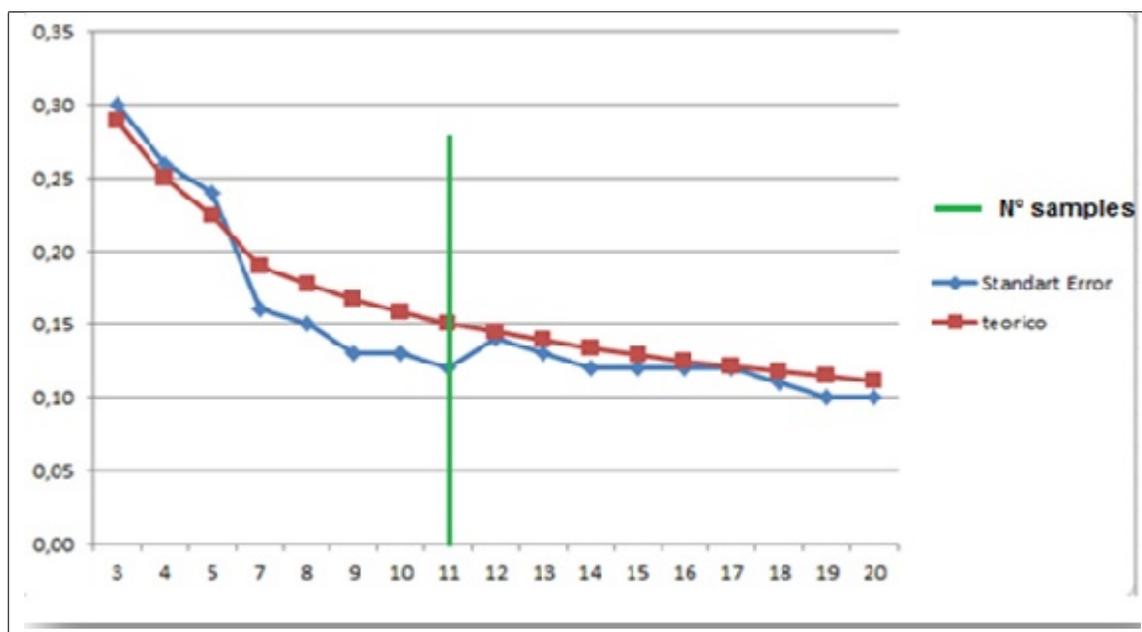


Fig. 6: Error estándar versus número de muestras. Fuente: Elaboración propia.

Anexo B: Muestreo Engagement: ¿Cuántas muestras considerar?

Elección muestra: Basada en estándar Error

Error estándar teórico: $SE = \sqrt{0,5 * 0,5/n}$, en rojo

Error estándar muestral, consideramos: $TENG_{jt} \sim Normal$ en azul

El *trade off* es que al disminuir el error se pierden observaciones.

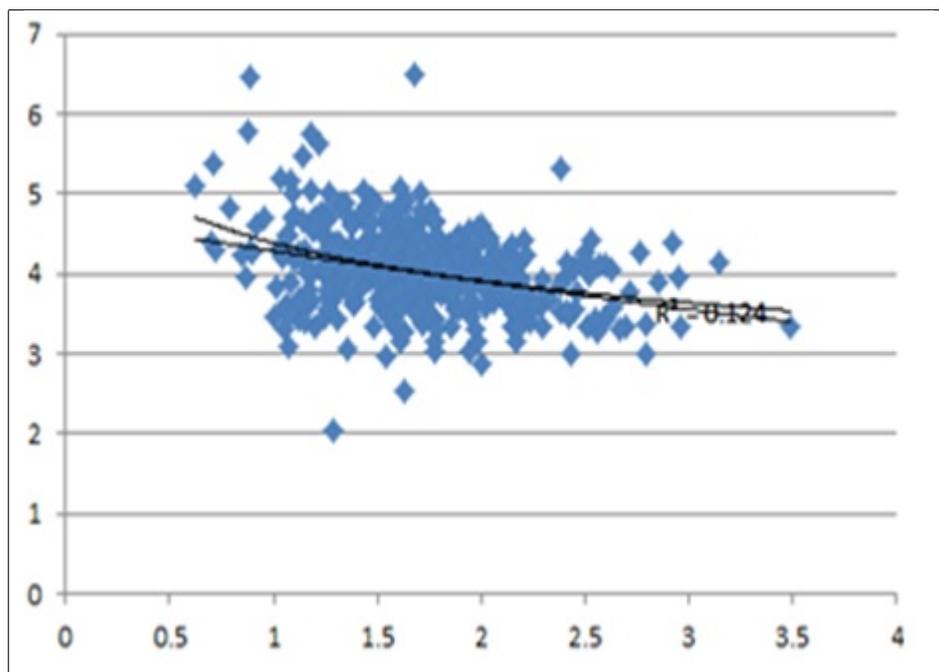


Fig. 7: Modelo de Logit de la atención versus tasa de atención. No hay evidencia para rechazar linealidad. Fuente: Elaboración propia.

Anexo C: Linealidad $\text{logit}(ENG_i)$ en el ratio

Se utilizaron múltiples especificaciones, pero todas estas ajustaron muy parecido al modelo lineal. No hay motivos para rechazar la linealidad entre el ratio y la atención.

Anexo D: Resumen principales resultados

	Sesgo corregido	Sesgo no corregido
IV	1.99*	1.41(mg)
No IV's	0.57*	0.39

Fig. 8: Resultados para la estimación de distintos criterios: Efecto de compromiso en la conversión. Fuente: Elaboración propia.

Solo control	Parámetro tasa de atención	P-valor
Día	2.62	0,000
feriado	2,04	0,000
clima	2,02	0,000
semana	1,73	0,000
tienda	1,03	0,034
Clientes	2.22	0,000
Todos	1.99*	0,022

Fig. 9: Tabla resumen con los valores que toma el parámetro de la tasa de atención al cambiar los controles en el modelo. Fuente: Elaboración propia.