



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

**ESTIMACIÓN DE DEMANDA CON INFORMACIÓN INCOMPLETA PARA APOYAR
NEGOCIACIONES DE PRECIOS EN UNA EMPRESA INDUSTRIAL**

MEMORIA PARA OPTAR AL TÍTULO DE INGENIERO CIVIL INDUSTRIAL

HÉCTOR EDUARDO PAREJAS ESPINOZA

SANTIAGO DE CHILE



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

**ESTIMACIÓN DE DEMANDA CON INFORMACIÓN INCOMPLETA PARA APOYAR
NEGOCIACIONES DE PRECIOS EN UNA EMPRESA INDUSTRIAL**

MEMORIA PARA OPTAR AL TÍTULO DE INGENIERO CIVIL INDUSTRIAL

HÉCTOR EDUARDO PAREJAS ESPINOZA

PROFESOR GUÍA:

MARCEL GOIC FIGUEROA

MIEMBROS DE LA COMISIÓN:

RICARDO MONTOYA MOREIRA

LUIS ABURTO LAFOURCADE

SANTIAGO DE CHILE

AGOSTO 2012

RESUMEN DE LA MEMORIA PARA OPTAR AL
TÍTULO DE INGENIERO CIVIL INDUSTRIAL
POR: HÉCTOR EDUARDO PAREJAS ESPINOZA
FECHA: 29/08/2012
PROF. GUÍA: SR. MARCEL GOIC

ESTIMACIÓN DE DEMANDA CON INFORMACIÓN INCOMPLETA PARA APOYAR NEGOCIACIONES DE PRECIOS EN UNA EMPRESA INDUSTRIAL

Para identificar oportunidades en el mercado es fundamental aumentar el conocimiento de los clientes y de la competencia. En el caso del marketing industrial, esto presenta un grado de dificultad mayor debido a que variables importantes como el precio final de venta y las ventas de la competencia son desconocidas.

El presente trabajo se desarrolla para a un distribuidor de materias primas que vende a restaurantes. Desde la perspectiva del distribuidor es de interés entender cómo sus clientes deciden sus compras y optan por adquirir a un proveedor u otro. Para esto de gran utilidad poder inferir parte esta información no observable con la finalidad de mejorar sus políticas de precio. Como fuente de información se cuenta con datos transaccionales de venta de la empresa y una encuesta con características demográficas de los clientes. De esta forma surge como objetivo apoyar las negociaciones de precio del distribuidor mediante un mayor conocimiento del mercado en términos de su posición competitiva y un mayor entendimiento del comportamiento de compra de sus clientes. Para esto se desarrolla un modelo que considera el uso de la teoría económica para interpretar la elección de compra de los clientes. En la estimación de los parámetros del problema se utiliza un modelo lineal jerárquico que permita capturar heterogeneidad de los clientes identificando patrones de comportamiento.

Dentro de los resultados obtenidos se logra identificar patrones de comportamiento de compra comunes en los clientes según sus características demográficas, lo cual sumado al modelo de decisión de compra hace posible inferir las compras de los clientes tanto al distribuidor como a su competencia, además de recomendaciones para los procesos de negociación. La estimación de las cantidades vendidas por el distribuidor de referencia bajo el efecto de la competencia arroja un error MAPE de 92,77% y un ajuste R^2 de 0,6977. Finalmente dentro de las recomendaciones formuladas destaca que los precios de los clientes cuyo proveedor principal es el distribuidor de referencia son menos sensibles a compras realizadas en el pasado. Otro punto relevante es como la ubicación del cliente juega un papel relevante en prácticamente la totalidad de las variables explicativas del modelo, desde diferencias en los requerimientos básicos por categorías, variaciones en la sensibilidad al nivel de precios y las variables que describen las diferencias de precios.

Para: Teresa, Héctor,

Ingrid y Nicolás.

AGRADECIMIENTOS

Durante el desarrollo de esta memoria debo agradecer en primer lugar la fortaleza y perseverancia entregada por Dios que me ayudó en los momentos donde pensé que todo se hacía más difícil.

A mi familia, mi mamá en especial, Teresa, que con su amor, dedicación y esfuerzo fue un pilar fundamental en construir lo que estoy a punto de lograr, mi título como ingeniero. A mi papá Héctor y mi hermano Nicolás por apoyarme a pesar que estábamos lejos. A mi hermana Ingrid por su compañía durante estos años en que compartimos nuestro periodo universitario.

Otro agradecimiento especial es para los grandes amigos que gané en la Universidad, los que conocí al principio y durante el desarrollo de mi carrera, que sin duda hicieron de todo este proceso una de las mejores etapas de mi vida.

Finalmente debo agradecer a mis profesores, en especial a mi profesor guía Marcel Goic por su buena disposición y voluntad para guiarme en el desarrollo de este trabajo.

Índice

1	Introducción	9
2	Descripción y Justificación del Proyecto	10
3	Objetivos.....	11
3.1	Objetivo General	11
3.2	Objetivos Específicos.....	12
4	Marco Conceptual.....	12
4.1	Marketing Industrial.....	12
4.2	Administración de Inventario y elección de un distribuidor.....	13
4.3	Enfoque Bayesiano	14
5	Metodología	18
5.1	Análisis Contexto del Problema	18
5.2	Análisis y Selección de Datos.	19
5.3	Diseño y Desarrollo del Modelo de Decisión de Compra	19
5.4	Estimación Parámetros con Modelo Lineal Jerárquico Bayesiano.....	19
5.5	Predicción Demanda de Clientes.....	20
5.6	Análisis y validación del modelo.....	20
5.7	Interpretación de Resultados y Recomendaciones.....	21
6	Alcances	21
7	Resultados Esperados.....	21
8	Desarrollo Metodología.....	22
8.1	Análisis Contexto del Problema	22
8.2	Análisis y Selección de Datos	22
8.3	Diseño y Desarrollo del Modelo de Decisión de Compra	27
8.4	Estimación Parámetros con Modelo Lineal Jerárquico Bayesiano.....	30
8.5	Predicción de Demanda de los clientes.....	35
8.6	Análisis y validación del modelo.....	39
8.7	Interpretación de Resultados y Recomendaciones.....	41
9	Conclusiones	47
	Bibliografía	49

Anexos	51
Anexo A: Tipo de información	51
Anexo B: Análisis Estacionalidad	53
Anexo C: Descripción variables demográficas:	55
Anexo D: Determinación de Ψ	58
Anexo E: Parámetros estimados a nivel de población	59
Anexo F: Parámetros a nivel de población para $\phi = 1$ y $\phi = 0,75$	61
Anexo G: Parámetros estimados a nivel de población (Goic y Montgomery, 2011) ..	65

Índice de Gráficos

Grafico 1: Ventas semanales por categoría	24
Grafico 2: Ventas Carnes año 1 y 2	24
Grafico 3: Ventas Aves año 1 y 2	25
Grafico 4: Precios promedio negociados por categoría.....	25
Grafico 5: Posteriori draws de los parámetros a nivel de población	33
Grafico 6: Indicadores selección de phi.....	34
Grafico 7: Real vs Estimado- cliente 27 - carnes.....	35
Grafico 8: Estimación Distribuidor vs Competencia - cliente 27 - carnes	35
Grafico 9: Real vs Estimado – cliente 123 - lácteos	36
Grafico 10: Estimación Distribuidor vs Competencia - cliente 123 - lácteos.....	36
Grafico 11: Real vs Estimado – cliente 42 – lácteos	37
Grafico 12: Estimación Distribuidor vs Competencia – cliente 42 - lácteos.....	37
Grafico 13: Real vs Estimado – cliente 17 - lácteos	38
Grafico 14: Estimación Distribuidor vs Competencia - cliente 17- lácteos.....	38
Grafico 15: Ventas semanales por categoría	40
Grafico 16: Proporción de ventas por categoría	52
Grafico 17: Ventas Lácteos año 1 y 2	53
Grafico 18: Ventas Congelados año 1 y 2	53
Grafico 19: Ventas Enlatados y Secos año 1 y 2	54
Grafico 20: Clientes por tipo Restaurante	55
Grafico 21: Proporción de clientes donde la empresa es su proveedor principal.....	55

Índice de Tablas

Tabla 1: Estadísticos variable cantidad vendida.....	23
Tabla 2: Correlación ventas por categoría	26
Tabla 3: Coeficientes por intervalo de iteración.....	33
Tabla 4: Validación modelo	39
Tabla 5: MAPE por categoría	39
Tabla 6: Ajuste por categoría	40
Tabla 7: Ajuste por categoría a nivel individual	41
Tabla 8: Coeficientes estimados β_i, γ_i	42
Tabla 9: Necesidades por categoría restaurante comida mexicana	43
Tabla 10: Necesidad básica por categoría lácteo según ubicación.....	44
Tabla 11: Sensibilidad nivel de precio restaurante comida italiana	44
Tabla 12: Sensibilidad nivel de precio por ubicación del cliente.....	44
Tabla 13: Efecto diferencia de precios si el cliente tiene al distribuidor como proveedor principal.....	45
Tabla 14: Efecto diferencia de precios por ubicación del cliente	46
Tabla 15: Muestra base de dato transaccional.....	51
Tabla 16: Identificador por categoría.....	51
Tabla 17: Estadísticos número de empleados.....	56
Tabla 18: Clientes por ubicación	56
Tabla 19: Validación $\phi=0,25$	58
Tabla 20: Validación $\phi=0,5$	58
Tabla 21: Validación $\phi=0,75$	58
Tabla 22: Validación $\phi=1,5$	58
Tabla 23: Validación $\phi=1,75$	58
Tabla 24: Validación $\phi=2$	59
Tabla 25: Validación $\phi=1$	59
Tabla 26: Validación $\phi=1,25$	59
Tabla 27: Parámetros estimados a nivel de población $\phi=1,25$	59
Tabla 28: Parámetros estimados a nivel de población para $\phi=1$	61
Tabla 29: Parámetros estimados a nivel de población para $\phi=0,75$	63

Índice de Ilustraciones

Ilustración 1: Distribución Posteriori Modelo Jerárquico Bayesiano	17
Ilustración 2: Plan Metodológico.....	18
Ilustración 3: Ubicación Clientes	57
Ilustración 4: Parámetros estimados a nivel de población (Goic y Montgomery, 2011) .	65

1 Introducción

Dentro de los conocimientos importantes de abordar en marketing se encuentra el entendimiento de la competencia y de los clientes de una empresa, incluyendo aspectos como identificación de posición competitiva, impacto de las decisiones de cada actor, estimación de demanda, entre otros. En el contexto del marketing industrial, definido como la compra de bienes y servicios para su procesamiento y posterior venta, lo anterior plantea dificultades debido a la existencia de información no observable, como los precios transados y el volumen de venta de la competencia. (Goic y Montgomery, 2011)

En la estimación de demanda se deben considerar características propias del mercado, como estacionalidad y cambios en el consumo, y aquellas pertenecientes al contexto industrial, como por ejemplo la existencia de requerimientos relativamente fijos para la compra de materias primas que reduce la posibilidad de sustitución entre productos distintos y la tendencia al abastecimiento de más de un proveedor (dual sourcing) justificado por beneficios asociados a aumentar la competencia y evitar quiebres de stock.

Esta memoria pretende identificar cómo clientes industriales toman sus decisiones de compra afectados por consideraciones propias de su industria y de sus proveedores y determinar cómo un distribuidor, que vende esos productos, puede inferir su función de demanda con información incompleta de ventas y precios de la competencia. Para esto se cuenta con información transaccional de la empresa y variables demográficas de sus clientes.

Frente a este escenario se propone construir una metodología que permita inferir información no observable para apoyar las negociaciones de precios de un distribuidor que vende a restaurantes en EEUU. Más concretamente es de interés obtener las cantidades que compra cada cliente al distribuidor de referencia y a su competencia sobre un determinado nivel de agregación de categorías en un determinado periodo de tiempo, con lo cual desarrollar recomendaciones de negociación.

Este trabajo se estructura en primer lugar con una descripción del problema, incluyendo las peculiaridades del contexto industrial, el comportamiento de un comprador y las posibilidades de modelar la decisión con los datos disponibles, para luego pasar al diseño y formalización del modelo, el cual considera el uso de la teoría económica para interpretar la elección de compra de los clientes y la estimación de los parámetros del problema por medio de un modelo lineal jerárquico que permita capturar heterogeneidad de los clientes y así finalmente obtener las inferencias buscadas, las cuales permiten formular recomendaciones para las políticas de precio del distribuidor.

2 Descripción y Justificación del Proyecto

El presente trabajo es desarrollado para el caso de un distribuidor de alimentos que vende a restaurantes independientes en EEUU. La industria está conformada por 33.000 distribuidores y se encuentra altamente concentrada en cinco empresas que reúnen cerca del 50% del ingreso del total de la industria. Los mayores distribuidores son McLane Company, SUPERVALU, Sysco Corp, y US Foodservice¹. En este mercado los precios son negociados directamente entre distribuidores y clientes, razón por la cual las transacciones realizadas por la competencia son desconocidas.

La empresa, cuyo nombre se mantendrá en reserva, posee un alto porcentaje de la industria, cuenta con clientes pertenecientes a un área extensa de EEUU, los cuales mantienen operaciones más bien pequeñas con ventas entre US \$1-5 millones anuales.

El distribuidor de referencia posee su propia fuerza de venta, la cual realiza visitas periódicas a sus clientes teniendo autonomía para negociar precios. Esto último tiene como consecuencia una elevada variabilidad en los precios y en la calidad de servicio entre clientes.

Para determinar cómo sus clientes deciden sus compras y traducir esto en recomendaciones que permitan apoyar las negociaciones de venta, es necesario poder inferir la información no observable ventas de la competencia. Para esto se cuenta con dos fuentes de información; una base de datos transaccional de ventas, que corresponde a un periodo de dos años y abarca desde el segundo semestre del 2008 al primer semestre del 2010, además de una encuesta que recoge características como el tipo de restaurante, número de empleados, código postal, entre otros.

Dentro de las investigaciones en marketing industrial se han desarrollado trabajos desde una perspectiva cualitativa que describen el comportamiento de los clientes industriales al momento de elegir un proveedor identificando factores relevantes (Cameron y Shipley 1985, Dolan 1995) tales como: confiabilidad de la entrega, condiciones de crédito, servicio post-venta, personal de ventas, gama de productos, marca comercializadora. Esto también es descrito de una manera más resumida en que el comportamiento de un comprador industrial frente a la elección por un proveedor u otro debe principalmente a tres factores: el precio, la calidad y el servicio ofertado (Puto, Patton, King, 1985).

Lo anterior supone un punto de partida para modelar el comportamiento de un cliente industrial frente a la elección de compras hacia sus proveedores.

Respecto a modelos que trabajan con información incompleta de ventas de la competencia existen estudios que abordan la estimación de demanda con información

¹ <http://www.firstresearch.com/industry-research/Food-Wholesalers.html>

incompleta desde la perspectiva de la administración de inventario (Chen y Plambeck, 2008), en donde la información a inferir corresponde a las ventas que se pierden por no tener stock con el objeto de mejorar políticas de inventario. Esto se realiza estimando de manera probabilística la demanda del cliente añadiendo una componente de sustitución entre productos de un mismo vendedor. Dicho de otra forma se define que en caso de enfrentar quiebres de stock, los clientes con una probabilidad p adquieren un producto sustituto y con una probabilidad $1-p$ se pierde la venta.

En el contexto particular de la problemática a abordar existe un primer enfoque (Goic y Montgomery, 2011) que usa información transaccional de ventas y demográfica de los clientes, en el cual se plantea un modelo econométrico para describir la decisión de compra de los clientes industriales, minimizando los costos de adquisición sujeto a la satisfacción de los requerimientos de los clientes empleando condiciones de primer orden para caracterizar tres regiones de solución para los parámetros donde los clientes compran a cada distribuidor y a ambos. Para esto se empleó un enfoque bayesiano, estimando por medio de MCMC y utilizando metropolis-hastings para la obtención de los parámetros a nivel individual. Dentro de los resultados alcanzados destaca la obtención de un mecanismo para inferir participaciones de mercado.

Con motivo de enriquecer el modelo existente se plantea incorporar nuevas consideraciones de inventario y la inclusión de la variable ubicación de cada cliente. Además se desea estudiar el efecto de simplificar el modelo para admitir solo soluciones interiores. Esto quiere decir, que si bien el cliente tiene libertad para escoger a quién y cuánto comprar, tomando en cuenta que el mantener un abastecimiento dual conlleva beneficios asociados al aumento de la competencia y a la disminución de la posibilidad de enfrentar quiebres de stock, se impone la condición de que el cliente siempre va a adquirir a más de un proveedor. Para resolver este problema se utiliza un modelo lineal jerárquico bayesiano, para incorporar heterogeneidad de los clientes sobre la solución interior identificada en el problema econométrico.

Finalmente el modelo a presentar en este trabajo contempla un modelamiento de la decisión de compra del cliente como un problema de optimización, en donde los parámetros del problema se estiman con un modelo lineal jerárquico para incorporar heterogeneidad en los clientes identificando patrones comunes de consumo.

3 Objetivos

3.1 Objetivo General

Desarrollar un modelo que permita apoyar las negociaciones de precios de un distribuidor industrial mediante la estimación de demanda de sus clientes con información incompleta de la competencia.

3.2 Objetivos Específicos

- Desarrollar un modelo de decisión de cómo clientes industriales eligen a sus proveedores y permita obtener las cantidades compradas al distribuidor de referencia y a su competencia.
- Determinar qué otros efectos además de la competencia influyen en la demanda.
- Estimar demanda total de cada cliente, definiendo el conjunto de variables a incluir en el pronóstico.
- Identificar cómo influye el nivel de precio sobre la decisión de compra de los clientes.
- Identificar posición competitiva del distribuidor.
- Evaluar capacidad de pronóstico del modelo para demandas futuras.
- Elaborar recomendaciones para apoyar las políticas de negociación de venta del distribuidor.

4 Marco Conceptual

4.1 Marketing Industrial

En marketing industrial los bienes y servicios son comprados para procesamiento y su consiguiente venta a otros clientes (Kasturi e Isaacson, 1994). Dentro de las características principales y que diferencian al marketing b2b se encuentran:

- Los precios son negociados directamente entre proveedor y cliente, en consecuencia los precios y cantidades acordados por parte de la competencia no son observables.
- Existencia de requerimientos relativamente fijos para la compra de materias primas reduciendo la posibilidad de sustitución entre productos distintos.
- En el caso de los clientes industriales existe un beneficio asociado al mantener más de un proveedor por producto (*dual sourcing*), esto se debe a que es posible

aumentar la competencia entre los proveedores o disminuir la posibilidad de enfrentar quiebres de stock (Lyon 2006).

En la descripción y justificación se mostraron diversos estudios de carácter cualitativo referentes a la decisión de compra de los clientes industriales, resumiendo esto en tres factores principales: precio, calidad y servicio ofertado. Respecto al precio un aspecto relevante en el comportamiento del cliente es identificar aquellos elementos que afectan la sensibilidad (Dolan 1995), tales como:

- El efecto de la reputación de la empresa ofertante. Para efectos prácticos de la formulación del modelo esto es interesante de incorporar.
- Diferenciación de la oferta entre proveedores. En este punto la oferta de la industria es muy similar en cuanto que los productos que se comercializan no tienen gran diferencia entre proveedores.
- Costo de cambiar de proveedor. En virtud de que los clientes son visitados por los proveedores no incurren en mayores costos, puesto que aquellos relacionados con la transacción son absorbidos mayoritariamente por el proveedor.
- El comprador puede comparar precio y performance de las alternativas.
- ¿En el mercado precios altos indican mejor calidad?
- ¿El costo del ítem representa un porcentaje substancial del gasto total?

Para el diseño del modelo se ha optado por no incluir factores asociados a aspectos de calidad y servicio en la elección de un proveedor debido a que son difíciles de observar con la información disponible.

4.2 Administración de Inventario y elección de un distribuidor

La literatura de administración de inventario es útil para identificar criterios de elección de proveedores y costos en los que se incurre por parte del comprador. En lo que se refiere a la cantidad óptima a comprar presenta un complejo problema de control de inventario, el cual considera los siguientes factores (Meyer y Assuncao, 1990):

- El precio futuro del bien.
- La tasa de consumo futuro.
- El valor de las alternativas.
- El costo de almacenar el bien.
- El costo de transacción.
- El tamaño del inventario existente.

Para el caso del problema en estudio se deben considerar las siguientes características:

- Baja diferenciación entre los productos ofertados por los proveedores, dada las especificaciones demandadas por los clientes.
- Los costos de transacción son absorbidos por el proveedor, quien visita a sus clientes periódicamente y entrega las compras.
- Los clientes poseen una limitada capacidad de inventario.

En vista de lo anterior el modelamiento matemático se centrará en la minimización de los costos del cliente, descartándose incluir un costo fijo de orden para reflejar el costo de transacción. Siguiendo el argumento anterior respecto a la existencia de poca diferenciación de productos y a que el costo de transacción es absorbido principalmente por el proveedor tampoco se incluye un costo por cambiar de proveedor.

En la estimación de los requerimientos de los clientes se contempla la inclusión de variables asociadas al inventario, que permita incorporar la capacidad de almacenamiento del cliente, su stock y el crecimiento que experimentan sus compras.

Por otro lado desde la perspectiva de un comprador la elección de un distribuidor u otro depende de distintos criterios, siendo recurrentes el tipo de relación entre las partes, reputación y posicionamiento dentro de la industria, y capacidad de entrega (Cebi y Bayraktar, 2003 y Guneri, Yucel y Ayyilidz, 2009). Sin embargo, cabe recordar que el contexto a analizar en esta memoria difiere del anterior en el sentido de que mientras los modelos mencionados tratan el problema desde la visión del cliente, el cual posee información interna como capacidad de almacenamiento y demanda, mientras que en esta memoria el análisis se desarrolla desde la perspectiva de un proveedor que desea inferir comportamientos del cliente para obtener una idea de su entorno competitivo.

4.3 Enfoque Bayesiano

El objetivo de este enfoque es hacer inferencia para describir una situación probabilística respecto a cantidades desconocidas como parámetros, variables latentes o de estados, o resultados condicionales usando toda la información disponible. Esta información puede ser basada en datos o no, como teorías de comportamiento, conocimiento implícito previo, etc.

Dentro de los beneficios de la inferencia bayesiana se encuentra:

- No se necesita recurrir a argumentos asintóticos para hacer inferencia.
- Integración coherente de múltiples fuentes de información.

- Métodos de simulación entregan mucha flexibilidad para estimar modelos complejos.

Costos:

- Computacional
- Hay que definir una distribución a priori (costo o beneficio?)
- Podría ser relevante para modelos con muchos parámetros.
- Se puede controlar mediante el uso de jerarquías.

Los algoritmos MCMC son ampliamente usados en inferencia estadística, para tomar muestra desde distribuciones complicadas de alta dimensión.

Cadena de Markov Monte Carlo (MCMC) es una clase de algoritmo para la toma de muestras de distribuciones de probabilidad sobre la base de la construcción de una cadena de Markov que tiene la distribución deseada como su distribución de equilibrio.

Para ajustar automáticamente la cadena de markov durante una ejecución existen diversas adaptaciones de MCM tales como: the Adaptive Metropolis (AM) multivariate algorithm of Haario et al. (2001), Metropolis-within-Gibbs algorithms for non-conjugate hierarchical models, Regionally Adjusted Metropolis algorithms, y logarithmic scalings.

MCMC puede ser muy exitoso en la búsqueda de buenos valores de parámetros con poca intervención del usuario en términos de alguna medida apropiada de mezcla de la cadena de markov, como por ejemplo la autocorrelación integrada de una función de interés. (Roberts y Rosenthal, 2008)

Un tipo de MCMC es el modelo lineal jerárquico, el cual considera un sistema de regresiones relacionadas a través de correlaciones entre vectores de los coeficientes de regresión (Rossi, Allenby y McCulloch, 2005).

Modelo Lineal Jerárquico Bayesiano:

Un modelo lineal jerárquico postula que la variable independiente y_i de cada unidad i queda descrita por una función lineal y un error normalmente distribuido, como indica el conjunto de regresiones de la siguiente ecuación: (Rossi, Allenby y McCulloch, 2005)

$$y_i = X_i\theta_i + \varepsilon_i, \quad \varepsilon_i \sim iid N(0, \sigma_i^2 I_{n_i})$$

La jerarquía es introducida al asumir que los parámetros de las regresiones a nivel de la unidad i provienen de una distribución común, como indica la siguiente ecuación:

$$\theta_i = \Lambda \cdot z_i + v_i \quad v_i \sim N(0, V_\theta)$$

Las variables del vector de z_i representan características específicas de cada unidad i . Esto quiere decir que los parámetros θ no son completamente independientes entre sí, sino que existe una heterogeneidad importante de considerar.

El modelo se completa con la especificación de la distribución a priori de los parámetros σ_i^2 , Λ y V_θ .

$$\begin{aligned} V_\theta &\sim IW(v, V) \\ \text{vec}(\Lambda) | V_\theta &\sim N(\text{vec}(\bar{\Lambda}), V_\theta \otimes A^{-1}) \\ \sigma_i^2 &\sim v_i \text{SS} q_{0i}^2 / \chi_{v_i}^2 \end{aligned}$$

Así permite explicar de manera más completa el comportamiento de los datos, en comparación con los modelos simples como una regresión lineal.

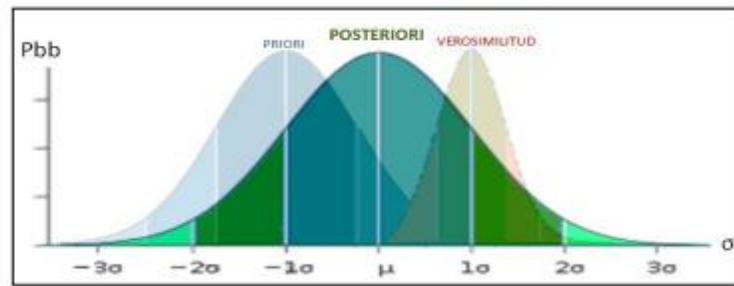
Para estimar los parámetros definidos mediante el Modelo Jerárquico Bayesiano se utiliza el algoritmo de iteración de Gibbs sampler, en el cual se realizan los siguientes 4 pasos iterativos que se repiten hasta que se obtiene la convergencia de las distribuciones a posterior:

$$\begin{aligned} \beta_i^1 &| y_i, X_i, Z\Delta_0, V_\beta^0, \sigma_i^0 \\ \sigma_i^1 &| y_i, X_i, Z\Delta_0, V_\beta^0, \beta_i^1 \quad i = 1, \dots, m \\ V_\beta^1 &| \{\beta_i^1\}, v, V \\ \Delta^1 &| \{\beta_i^1\}, A, \bar{\Delta}, V_\beta^1 \end{aligned}$$

Este método de iteración es una cadena de Markov cíclica que se obtiene a partir de una serie de distribuciones condicionales sobre una distribución inicial. Se separa en p grupos o bloques de parámetros y Gibbs sampler se define mediante la simulación repetitiva de cada una de las p distribuciones condicionales. El método Gibbs sampler va iterando entre las dos regresiones definidas, la regresión que relaciona Y con X , y la regresión lineal que analiza la relación entre θ y Z .

Finalmente se tiene que el comportamiento particular de cada cliente debiera parecerse al que tiene el grupo poblacional. El resultado posteriori resulta ser una multiplicación entre la verosimilitud y el priori, que corresponde a una actualización de las creencias a priori respecto al parámetro θ . En la siguiente figura, la curva color verde representa la distribución final resultante:

Ilustración 1: Distribución Posteriori Modelo Jerárquico Bayesiano



Fuente: Reuse, 2011

Medidas de evaluación para modelos de estimación.

Para evaluar la capacidad de estimación del modelo se tienen los siguientes estadísticos que son usados con frecuencia:

a) MAE: y_i = valor real; \bar{y}_i = valor estimado.

Error absoluto medio.

$$\text{MAE} = \frac{1}{N} \sum_{i=1}^N |y_i - \bar{y}_i|$$

b) MSE

Error cuadrático medio.

$$\text{MSE} = \frac{1}{N} \sum_{i=1}^N (y_i - \bar{y}_i)^2$$

c) MAPE

Medida de precisión de los valores pronosticados, error absoluto porcentual de la media.

$$\text{MAPE} = \frac{1}{N} \sum_{i=1}^N \left| \frac{y_i - \bar{y}_i}{y_t} \right|$$

d) WMAPE

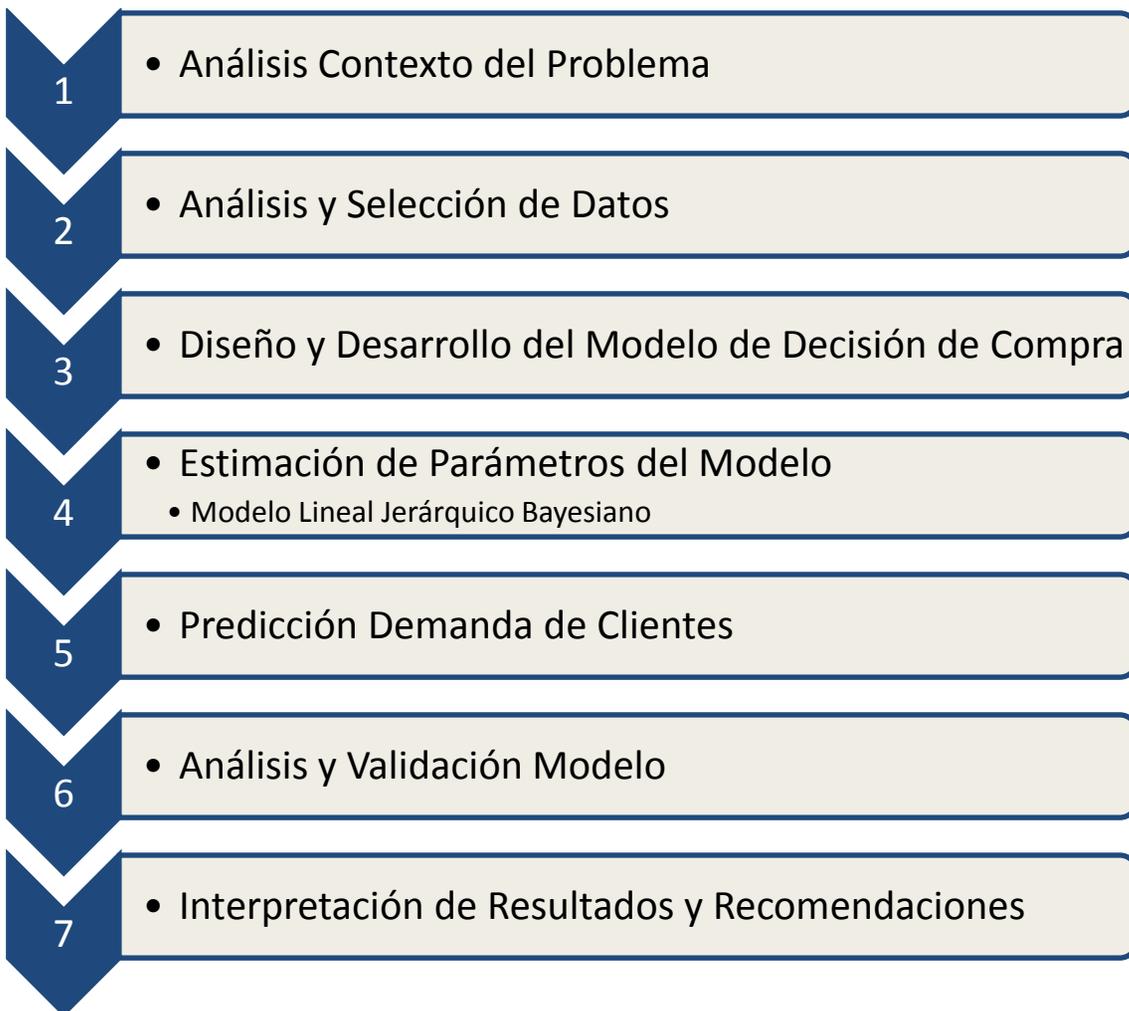
MAPE ponderado, evita que para valores reales = 0 quede indefinido.

$$\text{WMAPE} = \frac{\frac{1}{N} \sum_{i=1}^N |y_i - \bar{y}_i|}{\frac{1}{N} \sum_{i=1}^N y_i}$$

5 Metodología

La metodología a desarrollar en este trabajo recoge en parte la desarrollada anteriormente en trabajos similares (Reuse, 2011 y Troncoso, 2010) además de las particularidades propias del problema a resolver. Esto puede ser resumido con el siguiente diagrama:

Ilustración 2: Plan Metodológico



5.1 Análisis Contexto del Problema

Para un mayor conocimiento del contexto sobre el cual se enmarca el problema a resolver es necesario comprender el marketing industrial y sus características para poder modelar la forma en que los clientes toman decisiones sobre sus proveedores. El

contexto general del problema ya fue descrito en el capítulo 4, en donde se dan las primeras luces de las consideraciones que se tomarán en el modelo a partir de las características existentes en el marketing industrial ya establecidas.

De esta manera una vez identificado aquellos factores más relevantes estudiados en la literatura de marketing industrial se evalúan su inclusión en el modelo sujeto a la factibilidad otorgada por los recursos de información.

5.2 Análisis y Selección de Datos.

Previo al diseño del modelo es fundamental analizar la información que se posee mediante la inspección de los datos disponibles, la selección de aquellos que sean relevantes según el modelamiento tentativo, y finalmente un pre-procesamiento y transformación de datos según sea necesario, ya sea seleccionando una muestra balanceada, un determinado nivel de agregación de categorías, un horizonte de tiempo a emplear, entre otras. Esto con la finalidad de determinar los objetivos, resultados esperados y alcances de la memoria.

5.3 Diseño y Desarrollo del Modelo de Decisión de Compra

El diseño de un modelo de decisión tiene como finalidad establecer cómo un cliente industrial elige sus proveedores y las cantidades a comprar, considerando las particularidades presentes en marketing industrial y las propias de la industria de distribución. Además de las posibilidades que se derivan de la información disponible.

5.4 Estimación Parámetros con Modelo Lineal Jerárquico Bayesiano

Con la construcción del modelo de decisión se obtiene una caracterización de las soluciones del problema.

La solución encontrada estará en función de diversos parámetros que deben ser estimados, teniendo como finalidad la obtención patrones comunes entre los clientes que determinen sus comportamientos de compra. Esto se obtiene con un enfoque bayesiano, específicamente un modelo lineal jerárquico, el cual es capaz de incorporar heterogeneidad en los clientes.

Otro aspecto a definir tiene que ver con las variables explicativas a utilizar para la definición de los parámetros del modelo, los cuales debe permitir además de su determinación poder concluir en qué forma se ven afectados estos parámetros y como esto se puede interpretar desde una mirada industrial.

El modelo es programado en R, el cual es un conjunto integrado de servicios de software libre para manipulación de datos, cálculo y representación gráfica, para el cual existe una gran colección de *paquetes*² que permiten aplicar distintos métodos. Específicamente se utiliza el paquete *bayesm*, del cual se trabaja con la función *rhierLinearModel* para realizar las Regresiones Jerárquicas Bayesianas.

5.5 Predicción Demanda de Clientes.

Dadas las estimaciones de los parámetros a estimar sujeto a la solución encontrada con el modelo de decisión en esta etapa se muestran las cantidades estimadas de las compras de cada cliente por categoría de producto para el distribuidor de referencia y para su competencia, con lo cual se obtienen las participaciones de mercado

5.6 Análisis y validación del modelo

En esta etapa es necesario determinar la capacidad de pronóstico del modelo, de manera tal que sea posible su evaluación y valorización como herramienta efectiva que apoye las negociaciones del distribuidor.

Como método de validación se usará el indicador MAPE sobre los valores estimados, esto debido que para algunos valores reales se presentan valores cero. Un método alternativo que corrige este inconveniente es el MAPE ponderado o WMAPE, sin embargo, al realizar distintas pruebas variando algunas consideraciones en el modelo no se obtienen valores significativamente distintos. De esta forma el indicador a utilizar es el siguiente:

$$\text{MAPE}_{\text{est}} = \frac{1}{N} \sum_{i=1}^N \left| \frac{y_i - \bar{y}_i}{\bar{y}_i} \right|; \text{ donde: } y_i = \text{valor real}; \bar{y}_i = \text{valor estimado}$$

Este será calculado tanto sobre la muestra de entrenamiento como de testeo. Además se usará el indicador de ajuste R^2 , el cual es calculado tomando los valores reales de las ventas del distribuidor y su valor estimado con el modelo.

² http://cran.r-project.org/web/packages/available_packages_by_date.html

5.7 Interpretación de Resultados y Recomendaciones.

En esta sección se dará a conocer las conclusiones que se derivan de los resultados obtenidos por el modelo, identificando aspectos tales como funcionamiento del modelo, casos en los cuales no se obtienen resultados razonables y las recomendaciones para las políticas de negociación dentro del distribuidor.

6 Alcances

Según expuesto en el capítulo 5, específicamente en 5.1 y 5.2, respecto a las particularidades del mercado industrial sumado al tipo de información existente (base de dato transaccional y encuesta) se concluye que factores asociados a aspectos de calidad y servicio en la elección de un proveedor son difíciles de observar por tanto no se incluyen en el modelo.

Se desarrollará el modelo para un nivel de cinco grupos de categorías (lácteos, carnes, aves, congelados, enlatados y secos). En consecuencia no se obtendrán inferencias de precios ni estimaciones de cantidades individuales por productos, vale decir, los resultados estarán limitados al nivel de agregación de categorías escogido.

Se establecerán recomendaciones de negociación de precios identificando aquellos en los que el distribuidor es más y menos competitivo.

Se obtendrá la posición competitiva de la empresa mediante la estimación de las cantidades que vende y las de su competencia

7 Resultados Esperados

El presente trabajo plantea la construcción de un modelo que apoye las negociaciones de precio de un distribuidor industrial arrojando como resultado:

- La demanda agregada de los clientes a nivel de categorías.
- Inferencias de nivel de precios de la competencia (superior o inferior al del distribuidor de referencia)
- Inferencias de las ventas de la competencia.

- Recomendaciones para las negociaciones de precios para determinadas categorías y clientes.

8 Desarrollo Metodología

8.1 Análisis Contexto del Problema

Para representar los argumentos en la decisión de compra de un cliente industrial que permita identificar a qué proveedor y cuánto de cierta categoría de producto adquirirá en un determinado periodo de tiempo en este capítulo se presentaran las bases sobre las que será construido el modelo de optimización.

En el capítulo 4 ya se especificaron las bases sobre las cuales se sustenta el modelo, las cuales se resumen de la siguiente forma:

- Se establecerá un modelo con dos proveedores, el distribuidor de referencia y un “competidor” para cada producto que represente aquel más influyente.
- El modelo se basará principalmente en la minimización del costo de adquisición que enfrenta el comprador, descartándose incluir un costo fijo de orden para reflejar el costo de transacción, y en consecuencia tampoco un costo por cambiar de proveedor.
- Se incorpora en el costo de adquisición un beneficio asociado a la compra a más de un proveedor, que con la finalidad de estudiar el efecto de simplificar el modelo se supondrá que un cliente siempre estará dispuesto a comprar a más de uno para comparar la significancia de esta condición dados los resultados obtenidos por Goic y Montgomery (2011).
- Debido a que los clientes son muy pequeños en relación al distribuidor se asume que siempre habrá disponibilidad. Mismo caso para la competencia, en cuyo caso se debe a que esta se modelará como la agregación de los competidores.
- En la estimación de los requerimientos de los clientes se contempla la inclusión de variables asociadas al inventario, que permita incorporar la capacidad de almacenamiento del cliente, su stock y el crecimiento que experimentan sus compras.

8.2 Análisis y Selección de Datos

Los datos transaccionales corresponden a un periodo de dos años y abarca desde el segundo semestre del 2008 al primer semestre del 2010, donde se incluyen las ventas transadas por semana, cliente y categoría. Estos datos están agrupados en 5 categorías, las cuales representan un 80% de las transacciones y sobre un 90% de los ingresos. El total de semanas contenidas la base de datos transaccional es 104 y el total de clientes 130.

Para efectos del problema a resolver consideraran las siguientes variables transaccionales:

- cust: identificador del cliente.
- cat: identificador de categoría.
- week: semana de la venta.
- totQty: cantidad vendida.

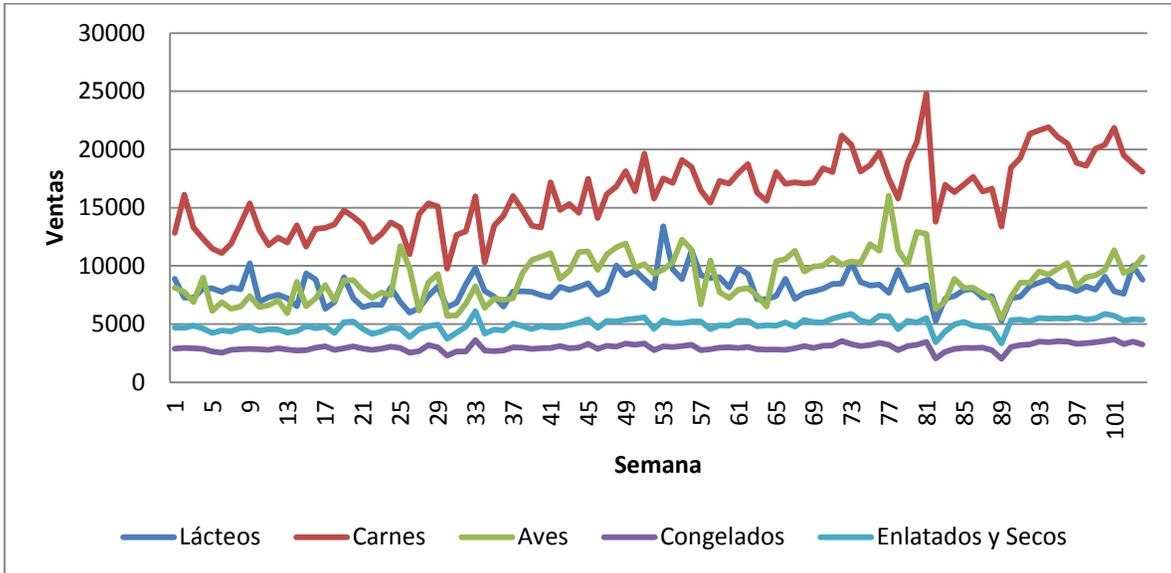
La siguiente tabla entrega una breve descripción de la variable cantidad vendida:

Tabla 1: Estadísticos variable cantidad vendida

Categoría	Promedio	Máx.	Mín.	Desv. Est.
Lácteos	64	5568	1	169
Carnes	130	2819	1	170
Aves	74	3567	1	203
Congelados	24	332	1	23
Enlatados y Secos	39	575	1	34

Las ventas experimentadas por categoría agregadas al total de clientes son mantienen un comportamiento similar a lo largo de las 104 semanas para las categorías lácteos, congelados, y enlatados y secos. Teniendo una mayor variación las carnes y aves principalmente la segunda mitad del periodo contenido en los datos.

Grafico 1: Ventas semanales por categoría



Al inspeccionar las ventas por categoría comparando las cantidades correspondientes al primer y segundo año de los datos transaccionales no se identifica la presencia de una estacionalidad importante, salvo en las categorías carnes y aves a fin de año o aves a mediados de año. De todas formas se ha optado por no incorporar estacionalidad al modelo.

Grafico 2: Ventas Carnes año 1 y 2

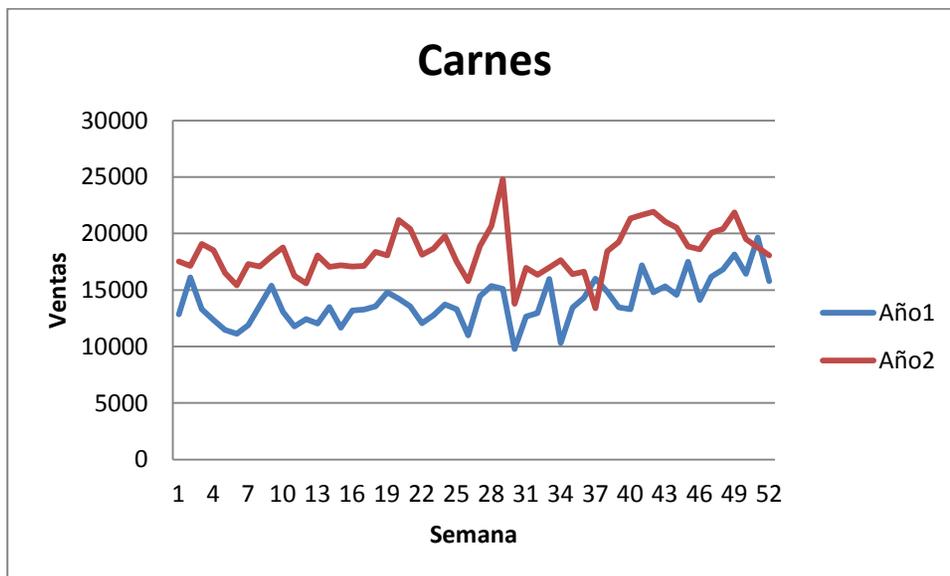
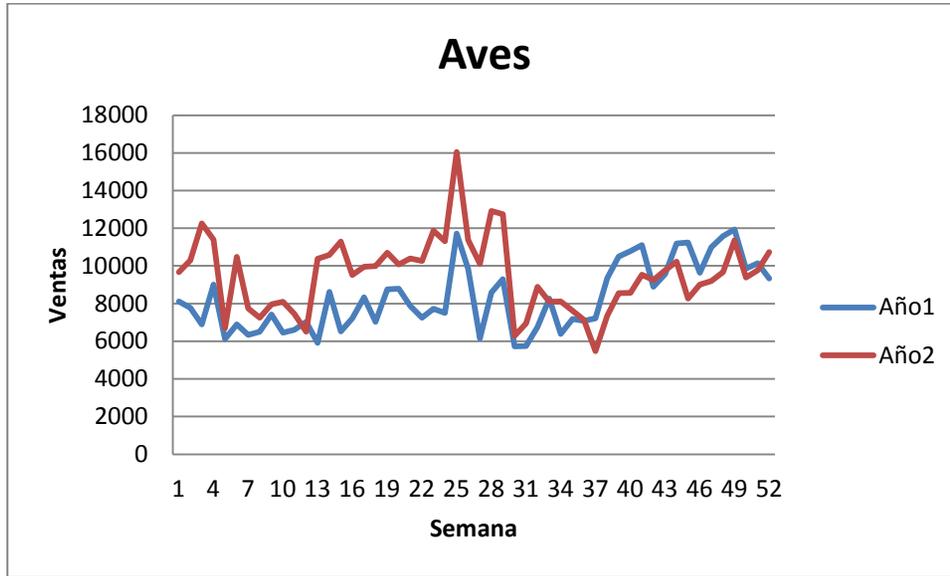


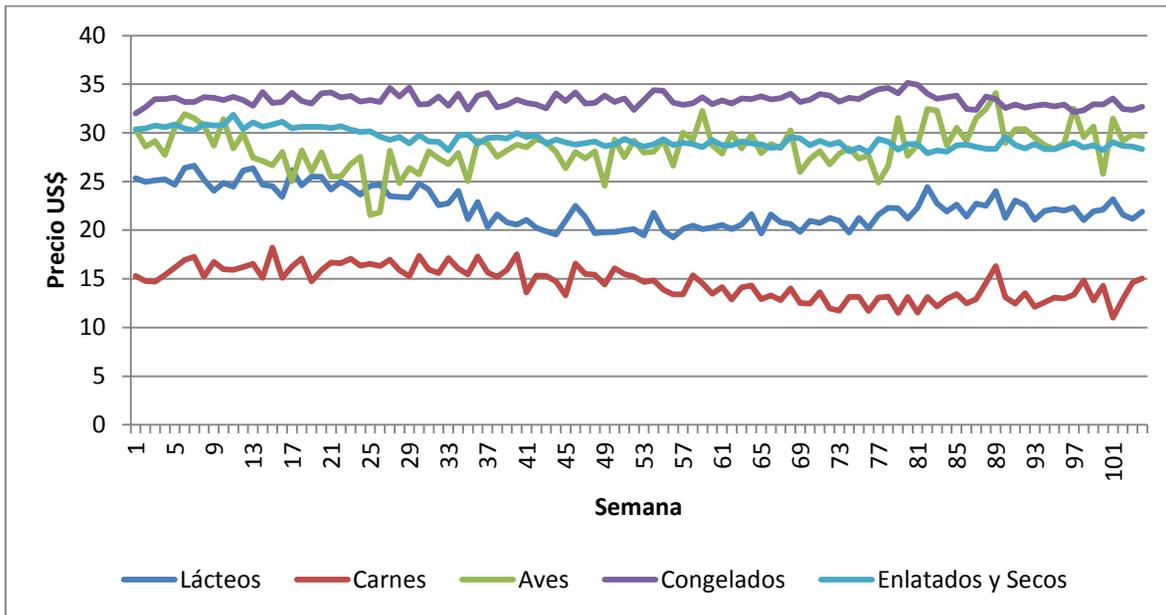
Grafico 3: Ventas Aves año 1 y 2



Para el caso del resto de las categorías se puede revisar en el anexo B.

Con respecto a los precios promedio negociados con los clientes son estables por categoría, pese a esto es de conocimiento en la empresa que dado que los precios se negocian directamente entre fuerza de venta y clientes, estos presentan una gran variabilidad para cada par SKU cliente a lo largo del tiempo.

Grafico 4: Precios promedio negociados por categoría



Mediante un simple análisis de correlación es posible identificar si las ventas por categorías se encuentran relacionadas.

Tabla 2: Correlación ventas por categoría

	Lácteos	Carnes	Aves	Congelados	Enlatados y Secos
Lácteos	1,000	0,388	0,342	0,470	0,580
Carnes	0,388	1,000	0,569	0,751	0,781
Aves	0,342	0,569	1,000	0,571	0,570
Congelados	0,470	0,751	0,571	1,000	0,893
Enlatados y Secos	0,580	0,781	0,570	0,893	1,000

Lo interesante de esta tabla es que de alguna forma las ventas de distintas categorías (en gran parte de ellas) están relacionadas, lo cual se puede deber a que los clientes prefieren comprar varios productos a un mismo proveedor lo cual se determinará a partir de los modelos de regresión a definir más adelante. Esto además permitirá analizar si estas correlaciones están igualmente presentes a nivel de cliente. Visto de otra forma parte de la correlación también se puede explicar por el tamaño de la operación entre el cliente y el proveedor, lo cual puede interpretarse como descuentos por cantidad, sin embargo, en la empresa consideran más importante el efecto de comprar a más de un proveedor que los descuentos no lineales debido a la limitada capacidad de almacenamiento de los clientes y su alta frecuencia de compra. Además si se optará por incluir los descuentos de precio, estos debiesen ser descritos como una función de covariables distintas a las que afectan al efecto de comprar a más de un proveedor.

Las variables demográficas consideradas son:

- Tipo de restaurante (comida americana, italiana, mexicana y pizzería).
- Si el distribuidor de referencia es el proveedor principal.
- El número de empleados del restaurante.
- La ubicación del cliente.

De las variables demográficas antes mencionadas las tres primeras se extraen directamente de la base de dato que contiene la información de la encuesta. Para la ubicación del cliente se tiene el código postal, contabilizándose un total de 84 códigos distintos, para reducir el número de ubicaciones posibles se recopilaron las latitudes y longitudes³ asociadas a cada código postal. Estas geo-referencias fueron agrupadas de acuerdo a la cercanía entre ellas empleando un método k-means apoyado en la ubicación visualizada en Google Earth⁴ para la elección del k. Si bien es cierto las

³ Latitudes y longitudes obtenidas en: <http://www.geopostcodes.com/>

⁴ Google Earth:

http://www.google.com/intl/es/earth/index.html#utm_campaign=es&utm_medium=ha&utm_source=es-ha-sk-eargen&utm_term=earth

coordenadas latitud y longitud no son lineales dada la cercanía de los restaurantes se trabajo como si así fuesen. Finalmente se agruparon según 6 posibles ubicaciones.

Para apreciar la conformación de la data, revisar Anexo A y C.

8.3 Diseño y Desarrollo del Modelo de Decisión de Compra

Anteriormente se ha descrito el contexto del problema con las consideraciones previas a tener en cuenta en el modelamiento y la selección de datos a utilizar. Con esto se procede a formular el modelo propiamente tal, el cual como ya ha sido mencionado considera que el cliente escoge a quién y cuanto comprar a través de la minimización del costo de adquisición considerando un descuento asociado a la compra a más de un proveedor.

Como uno de los objetivos es identificar la posición de la empresa en el mercado, en cuanto a las ventas que concreta y a las que pierde considerando los precios de la competencia de manera endógena, se establecerá un modelo con dos proveedores el distribuidor de referencia y un “competidor” para cada producto que represente aquel más influyente. Esto porque la intención del modelo es estimar la demanda total de los clientes y el porcentaje que está captando la empresa y el que no, además de poder entender los motivos por los cuales ese porcentaje no está siendo capturado, ya sea por precios poco competitivos, por la importancia de las transacciones históricas de las partes, intensidad de la competencia según una determinada ubicación o bien debido a que la oferta en términos de la totalidad de los requerimientos es menos atractiva.

Debido a que los clientes son muy pequeños en relación al distribuidor se asume que siempre habrá disponibilidad. Mismo caso para la competencia, en cuyo caso se debe a que esta se modelará como la agregación de los competidores.

VARIABLES DE DECISIÓN:

M = total de categorías.

i = Índice cliente.

$m = 1, \dots, M$ índice categoría.

$t=1, \dots$, período de tiempo (semanal).

q_{imt} = cantidad comprada por el cliente i en la categoría m en el periodo t al distribuidor de referencia.

q_{imt}^c = cantidad comprada al competidor, asumiendo que corresponde a la cantidad agregada que adquiere a la competencia y se entiende como aquel competidor más atractivo para el cliente.

Restricciones:

Satisfacción de los requerimientos del cliente i en el periodo t por categorías:

$$q_{imt} + q_{imt}^c \geq \tau_{imt}$$

Función objetivo:

De acuerdo a los principios del *dual sourcing* descritos en el capítulo 4, en el mercado industrial existe un beneficio que proviene del abastecimiento de más de un proveedor debido principalmente a influenciar una mayor competencia entre los proveedores o disminuir la posibilidad de enfrentar quiebres de stock. Esto es incorporado en la función de utilidad del modelo. Adicionalmente, con la finalidad de estudiar el efecto de simplificar el modelo se supondrá que un cliente siempre estará dispuesto a comprar a más de uno para comparar la significancia de esta condición dados los resultados obtenidos por Goic y Montgomery (2011). Es así como con la finalidad de minimizar los costos de adquisición de cada cliente en cada periodo se tiene:

$$\text{Min CT } (\{q_{imt}, q_{imt}^c\}_{m=1}^M)$$

Donde:

$$CT = \sum_{m=1}^M \{p_{imt} q_{imt} + p_{imt}^c q_{imt}^c - \psi q_{imt} q_{imt}^c\}$$

La caracterización de la región de solución se determina derivando las condiciones de primer orden del problema antes definido. Cabe señalar que por la forma funcional de la función de costo el problema de optimización es separable por categoría:

Derivando las condiciones de primer orden del modelo y simplificándolo para que admita solo soluciones interiores se obtiene que la solución esté dada por el siguiente sistema de ecuaciones:

$$\begin{aligned} p_{imt} - \psi q_{imt}^c - \lambda_{imt} &= 0 // q_{imt} > 0 \\ p_{imt}^c - \psi q_{imt} - \lambda_{imt} &= 0 // q_{imt}^c > 0 \\ q_{imt} + q_{imt}^c &= \tau_{imt} \end{aligned}$$

Desarrollando el sistema de ecuación anterior y definiendo $\delta_{imt} = p_{imt} - p_{imt}^c$, se obtiene que las cantidades que compra a cada proveedor están dadas por:

$$\begin{aligned} q_{imt} &= 1/2 \left(\tau_{imt} - \delta_{imt}/\psi \right) \\ q_{imt}^c &= 1/2 \left(\tau_{imt} + \delta_{imt}/\psi \right) \end{aligned}$$

La simplificación impuesta a la adquisición de más de un proveedor se debe a que de esta forma, por medio de la condición de primer orden, se considera una solución para la cantidad demandada al distribuidor de referencia, sobre la cual es posible usar el modelo lineal jerárquico dado la definición de una ecuación lineal que se define a partir de esta solución para la estimación conjunta de los parámetros τ_{imt} y δ_{imt} .

Parámetros:

τ_{imt} = requerimientos de materias primas para el cliente i en la categoría m en el periodo t .

ψ = propensión del cliente para adquirir a más de un proveedor, vale decir, la condición de que con $\psi > 0$ el cliente prefiere comprar a más de un distribuidor. (Dual sourcing). Por simplicidad se asumirá constante en el tiempo y entre categorías.

δ_{imt} = diferencia de precios entre el distribuidor de referencia y su competencia ($p_{imt} - p_{imt}^c$), de esta forma endogeneizar la variable no observable precio facilita la estimación de las cantidades que adquiere el cliente q_{imt} y su competencia q_{imt}^c . (Shugan, 2004).

De esta forma se definen las siguientes ecuaciones de regresión para la cantidad demandada por los clientes (τ_{imt}) y la diferencia de precios (δ_{imt}):

$$\tau_{imt} = XT'_{imt}\beta_i + \varepsilon_{1imt}; \varepsilon_1 \sim iid N(0, \sigma_1^2 I_{n_1})$$

$$\delta_{imt} = XD'_{imt}\gamma_i + \varepsilon_{2imt}; \varepsilon_2 \sim iid N(0, \sigma_2^2 I_{n_2})$$

Donde XT y XD son matrices de variables explicativas observables, β_i y γ_i parámetros a estimar y ε_{1imt} y ε_{2imt} errores asociados a la estimación.

8.4 Estimación Parámetros con Modelo Lineal Jerárquico Bayesiano

Dadas las ecuación de regresión definidas para de τ_{imt} y δ_{imt} se procede a la elección de las variables explicativas, las cuales deben cumplir como condición ser distintas para cada uno de los parámetros a fin de evitar problemas de multicolinealidad, que se puede manifestar en que parámetros de nivel individual no alcancen estacionariedad, valores estimados para los coeficientes β_i y γ_i reflejan efectos compartidos u otro.

XT:

- Considera un intercepto para cada categoría, entendido como las necesidades de cada cliente por cada una de las categorías.
- Coeficiente precio asumido constante entre categorías que considera la significancia del nivel de precios en los requerimientos del cliente. (Precio promedio por categoría y cliente a lo largo del tiempo)
- Crecimiento de las compras, definido como:

$$\frac{\text{Unidades vendidas semana anterior}}{\text{Unidades vendidas en promedio semanal último mes}} = \frac{q_{imt-1}}{\frac{1}{4} \times \sum_{t-4}^{t-1}(q_{imt})}$$

- Exceso de compras por parte del cliente en periodo anterior dado un nivel de compras promedio (Donde t_{oi} corresponde al periodo de inicio de las compras del cliente i):

Unidades vendidas semana anterior – Promedio histórico unidades vendidas

$$= q_{imt-1} - \left(\frac{1}{(t - t_{0i})}\right) \times \sum_{t_{0i}}^{t-1} (q_{imt})$$

Las últimas dos variables explicativas se derivan de la descripción realizada por Meyer y Assuncao (1990) respecto a los factores a considerar para la determinación de cantidades óptimas a comprar sujeto a consideraciones de inventario rescatando particularmente la tasa de consumo futuro y el tamaño de inventario existente. Además para la construcción propiamente tal se tomó en cuenta lo modelado por Figueroa (2009) quien intenta capturar efectos similares de crecimiento de ventas y exceso de compras para la estimación de quiebres de stock.

XD:

- Ventas en el periodo anterior en la categoría de producto.
- Volumen total de ventas en todas las categorías en periodos anteriores. para entender el efecto del conocimiento previo entre las partes y como influiría la intensidad de la relación.

Las Variables asociadas a XD fueron normalizadas para obtener diferencias de precios centradas en cero.

El resto de las variables explicativas siguen las consideraciones establecidas en Goic y Montgomery (2011).

ψ (Beneficio dual sourcing): El parámetro ψ es determinado minimizando el error asociado a la estimación y de acuerdo al ajuste del modelo probando distintos valores entre 0,25 y 2.

Estimación:

Para la estimación conjunta de los parámetros definidos anteriormente se define la siguiente ecuación lineal por cliente y categoría basada en la solución encontrada a partir del modelo de decisión.

$$y_i = X_i \theta_i + \varepsilon_i, \quad \varepsilon_i \sim iid N(0, \sigma_i^2 I_{n_i})$$

Donde X_i corresponde a las variables explicativas detalladas anteriormente, que en el caso de las que determinan la diferencia de precios se encuentra incorporado el factor ψ . θ_i es definido como el vector de parámetros individuales

$$\theta_i = (\text{vec}(\beta_i), \text{vec}(\gamma_i))$$

Recordar que los coeficientes β_i corresponden a la ecuación de regresión para la estimación de los requerimientos de los clientes y los coeficientes γ_i corresponden a la ecuación de regresión para la diferencia de precios entre el distribuidor y su competencia.

Introduciendo heterogeneidad de los clientes para especificar que los parámetros de los modelos siguen una distribución de población común, la cual por simplicidad asumimos normal:

$$\theta_i = \Lambda \cdot z_i + v_i \quad v_i \sim N(0, V_\theta)$$

Donde z_i corresponde a las características demográficas de los clientes obtenidas con la encuesta, siendo las escogidas para el modelo: tipo de restaurante, número empleados, si es el distribuidor el proveedor principal y la ubicación de los clientes. (En el Anexo C se encuentra una descripción de las variables demográficas)

La inclusión de una estructura jerárquica juega un rol central en la identificación de los parámetros con información incompleta. Al observar la demanda de otros clientes con perfiles demográficos similares es posible evaluar cuan probables es que los clientes compren a un competidor versus al distribuidor de referencia. Para completar el modelo se utilizan priores conjugados, vale decir, la posterior está en la misma familia de distribución que el prior.

$$\begin{aligned} V_\theta &\sim IW(v, V) \\ \text{vec}(\Lambda) | V_\theta &\sim N(\text{vec}(\bar{\Lambda}), V_\theta \otimes A^{-1}) \\ \sigma_{1i}^2 &\sim v_1 \text{ssq}_1 / \chi_{v_1}^2 \\ \sigma_{2i}^2 &\sim v_2 \text{ssq}_2 / \chi_{v_2}^2 \end{aligned}$$

Donde $v, v_1, v_2, \text{ssq}_1, \text{ssq}_2, V, \bar{\Lambda}$ y A son escogidos para tener relativamente diffuse priors.

De las 104 semanas que completan la base de datos transaccional se han destinado las primeras 88 para la calibración del modelo y las última 16 para el testeo.

Para obtener los parámetros necesarios para la estimación de demanda, se ejecutaron un total de 40.000 iteraciones del modelo, usándose las últimas 10.000 para las estimaciones. Esto debido a que las iteraciones iniciales no han alcanzado estacionariedad. A modo de ejemplo, tres de los posterior draws de los parámetros a nivel de población:

Grafico 5: Posteriori draws de los parámetros a nivel de población

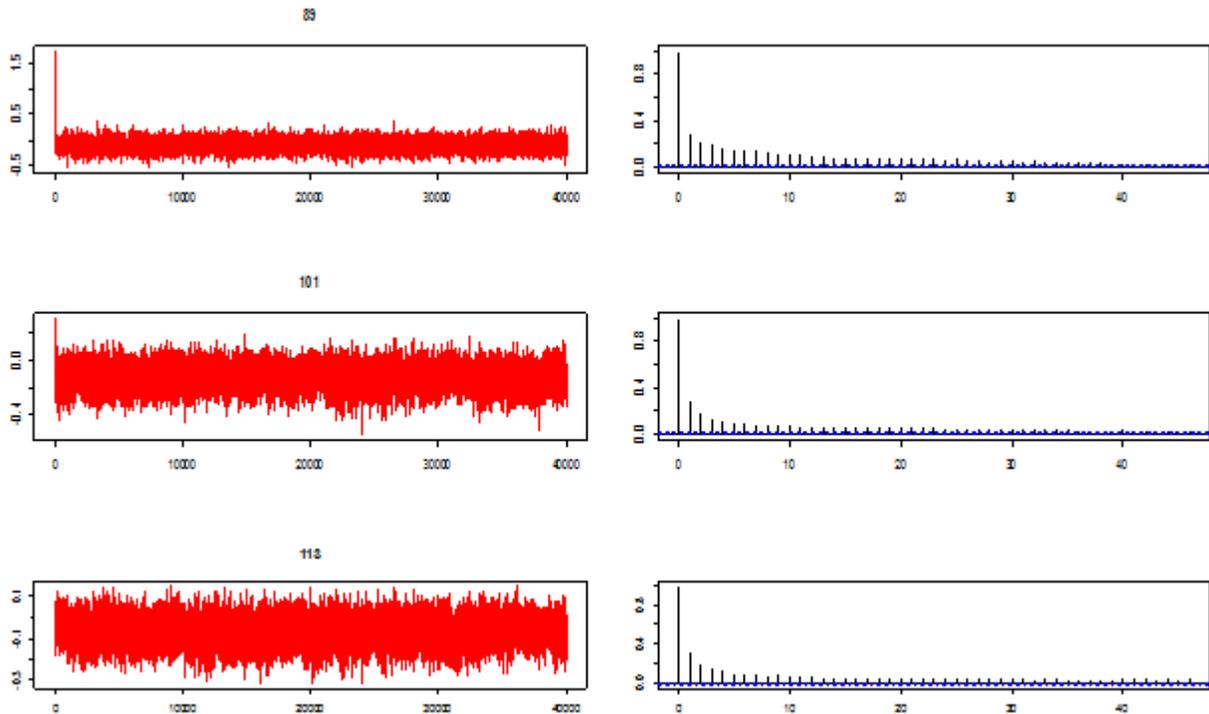


Tabla 3: Coeficientes por intervalo de iteración

Intervalo iteración	Media									
	β_1	β_2	β_3	β_4	β_5	β_6	β_7	β_8	γ_1	γ_2
0-5.000	1322,663	252,447	247,336	262,727	315,654	-6,467	0,117	-30,984	-79,623	-48,897
5.000-10.000	1328,146	259,142	253,399	273,403	323,117	-6,569	0,116	-38,408	-84,093	-47,392
10.000-15.000	1328,313	259,170	256,074	275,123	323,411	-6,652	0,117	-36,468	-79,860	-51,070
15.000-20.000	1326,063	257,479	252,132	272,258	320,803	-6,662	0,113	-32,692	-78,922	-50,316
20.000-25.000	1329,824	256,937	251,912	270,824	321,098	-6,479	0,115	-39,092	-84,293	-47,840
25.000-30.000	1328,901	254,391	248,452	268,527	318,456	-6,474	0,116	-34,075	-84,923	-45,470
30.000-35.000	1327,802	256,239	253,762	270,006	320,214	-6,572	0,115	-35,136	-80,490	-49,315
35.000-40.000	1329,440	255,710	251,116	268,379	318,047	-6,389	0,114	-38,028	-88,426	-42,240
30.000-40.000	1328,620	255,967	252,437	269,193	319,132	-6,480	0,115	-36,581	-84,459	-45,778

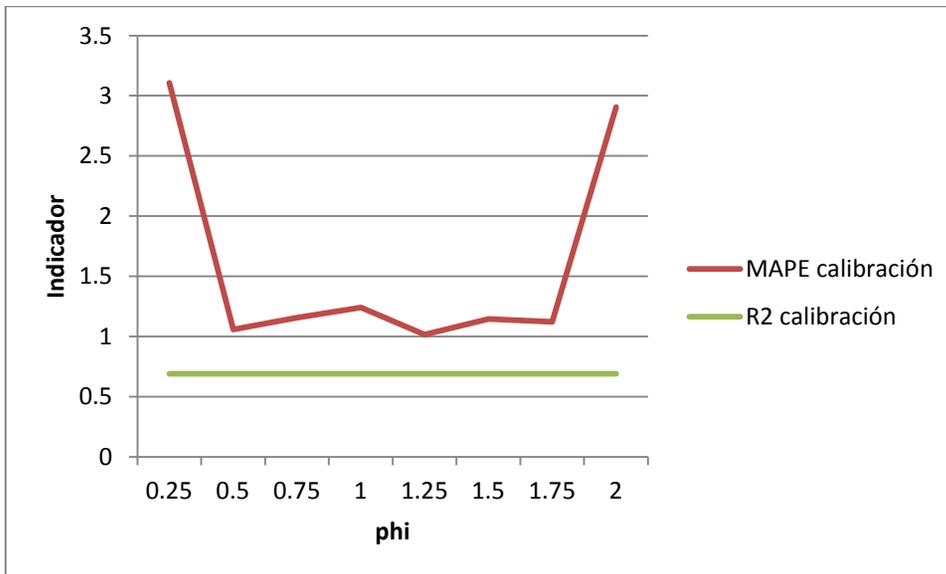
Al analizar las medias de los coeficientes estimados por intervalo de iteración es posible observar que los coeficientes asociados a consideraciones de inventario y al

parámetro diferencia de precio varían de forma no despreciable. Es por esto que, pese a que los parámetros a nivel de población no parecieran tardar en converger se escogió un gran número de iteraciones.

Determinación de ψ : Como ya fue mencionado ψ se obtiene a partir del mayor ajuste de la estimación (R^2) y su menor error asociado (MAPE estimado). Para esto se probó con valores 0,25 – 0,5 – 0,75 – 1 – 1,25 – 1,5 – 1,75 – 2.

De manera preliminar se optó por realizar cinco ejecuciones por valor de ψ , obteniéndose en promedio los siguientes resultados (para ver el detalle de los valores revisar anexo D):

Grafico 6: Indicadores selección de phi



En los valores analizados el criterio de elección por calidad de ajuste del modelo es irrelevante debido a que rondan valores muy similares, en torno al 69%. En consecuencia se empleara el criterio de mínimo error para la determinación del phi.

Para los valores extremos $\psi = 0,25$ y $\psi = 2$ se obtienen valores considerablemente altos respecto al resto para ciertas ejecuciones, pese a que también hay casos que toman valores bajos se descartan por la poca consistencia de los resultados.

Para los valores intermedios $\psi = 0,5$, $\psi = 0,75$, $\psi = 1$, $\psi = 1,25$ y $\psi = 1,75$ no hay mucha diferencia en los errores asociados a las estimaciones, tampoco en cuanto a los resultados obtenidos en los parámetros a nivel de población (Ver anexo E y F)

Finalmente el valor de ψ con el cual se obtiene un menor error de manera consistente en las 5 ejecuciones realizadas es 1,25, por tanto se realizarán las estimaciones y los análisis de resultados para ese valor.

Con el valor escogido para ψ se realiza la estimación definitiva para los parámetros τ_{imt} y δ_{imt} . (Para ver el detalle de los valores obtenidos de error y ajuste por valor de ψ ver Anexo D)

8.5 Predicción de Demanda de los clientes.

Teniendo los valores calculados para los parámetros del problema ψ , τ_{imt} y δ_{imt} se procede a estimar q_{imt} y q_{imt}^c .

A modo de graficar algunos resultados se tienen:

Grafico 7: Real vs Estimado- cliente 27 - carnes

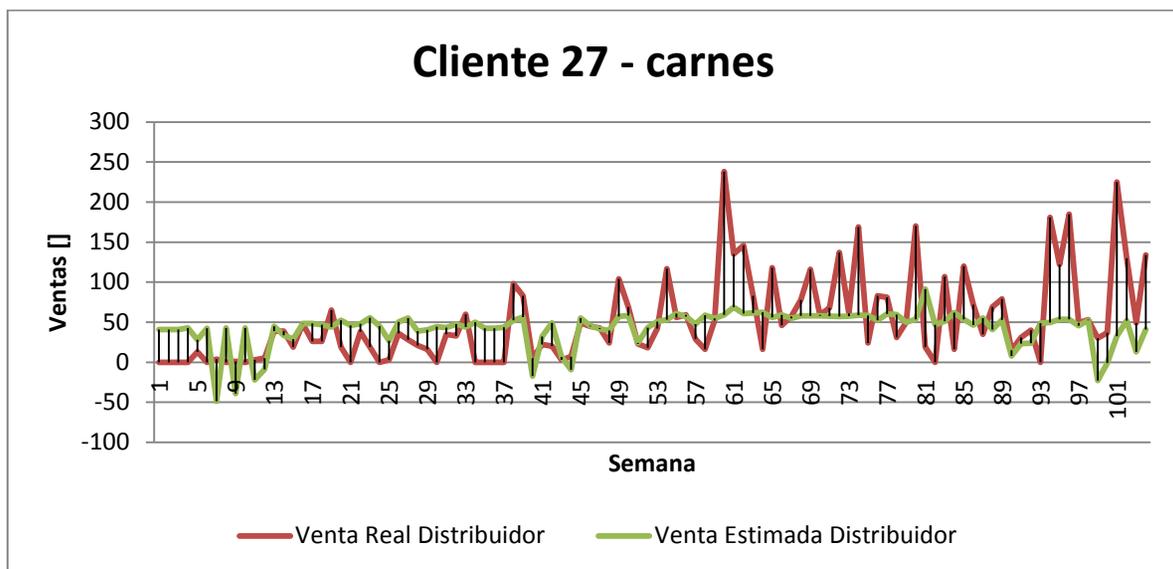


Grafico 8: Estimación Distribuidor vs Competencia - cliente 27 - carnes

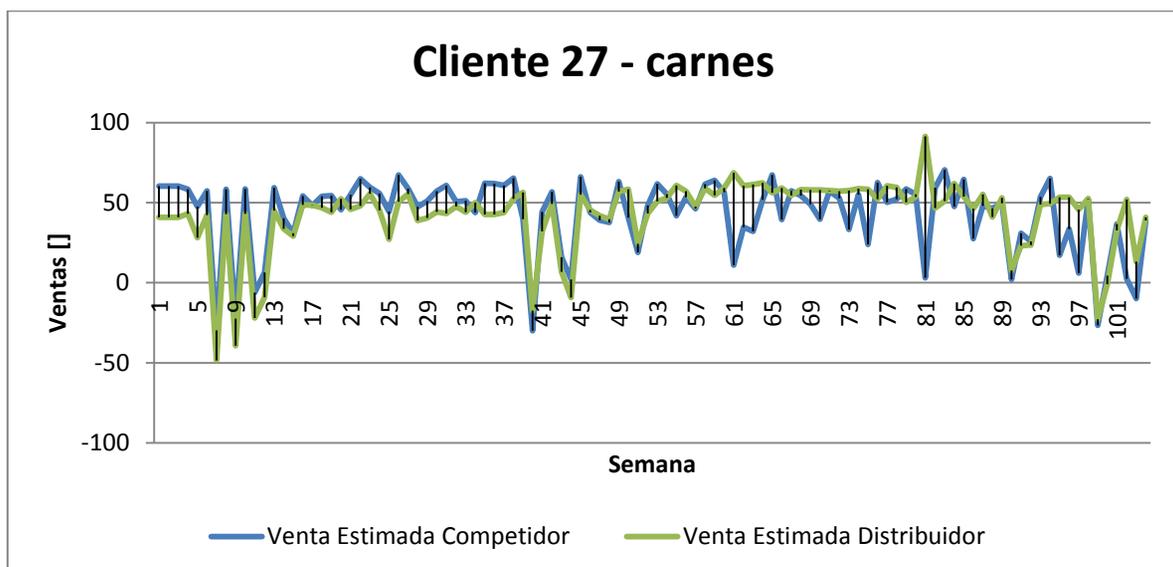


Grafico 9: Real vs Estimado – cliente 123 - lácteos

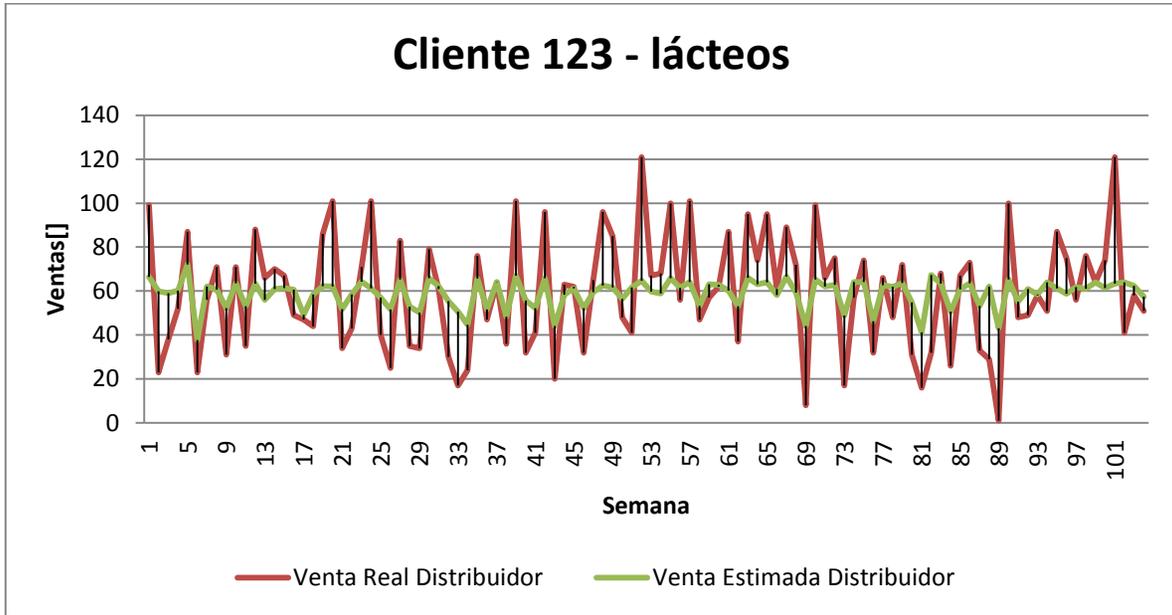


Grafico 10: Estimación Distribuidor vs Competencia - cliente 123 - lácteos

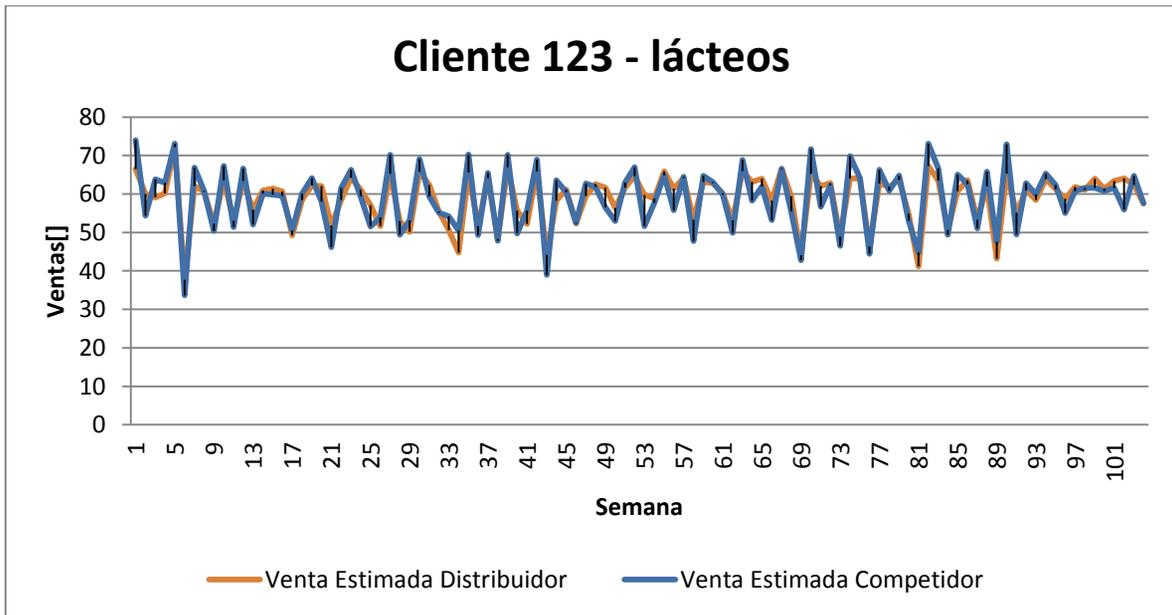


Grafico 11: Real vs Estimado – cliente 42 – lácteos

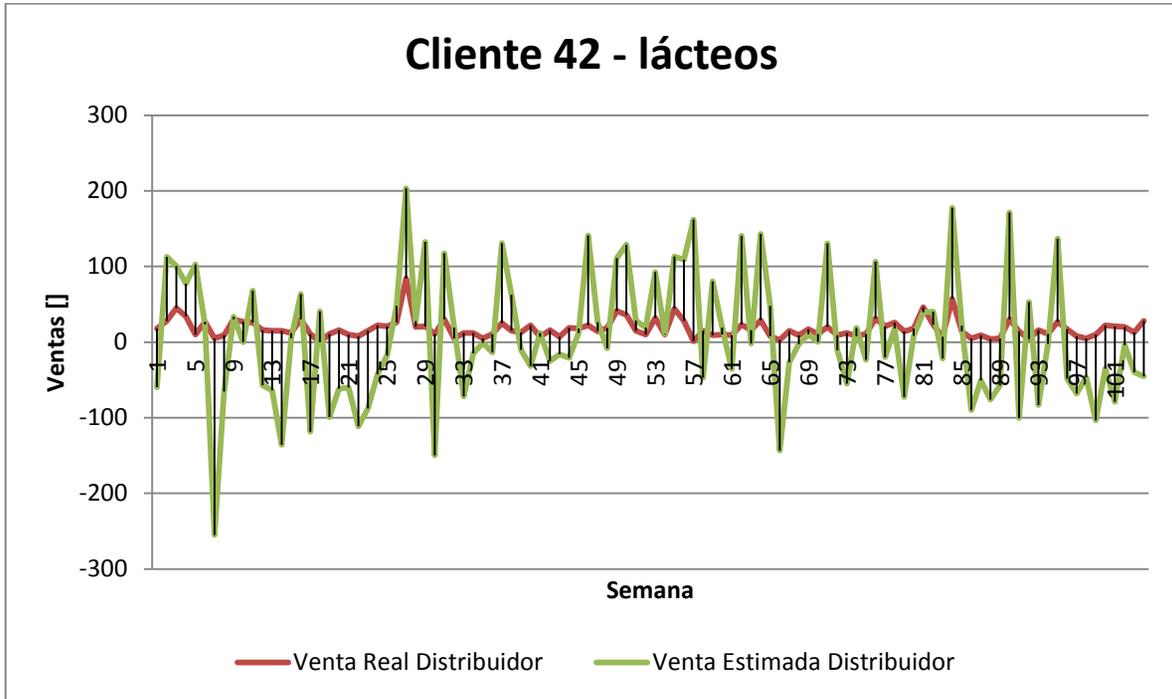


Grafico 12: Estimación Distribuidor vs Competencia – cliente 42 - lácteos

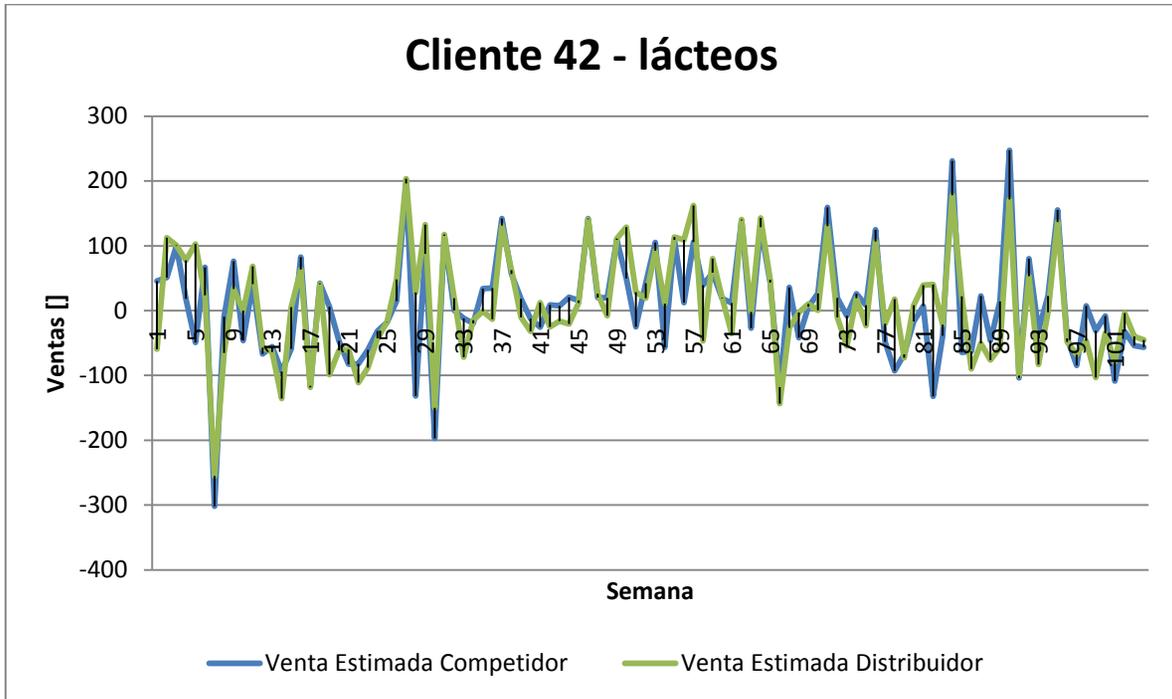


Grafico 13: Real vs Estimado – cliente 17 - lácteos

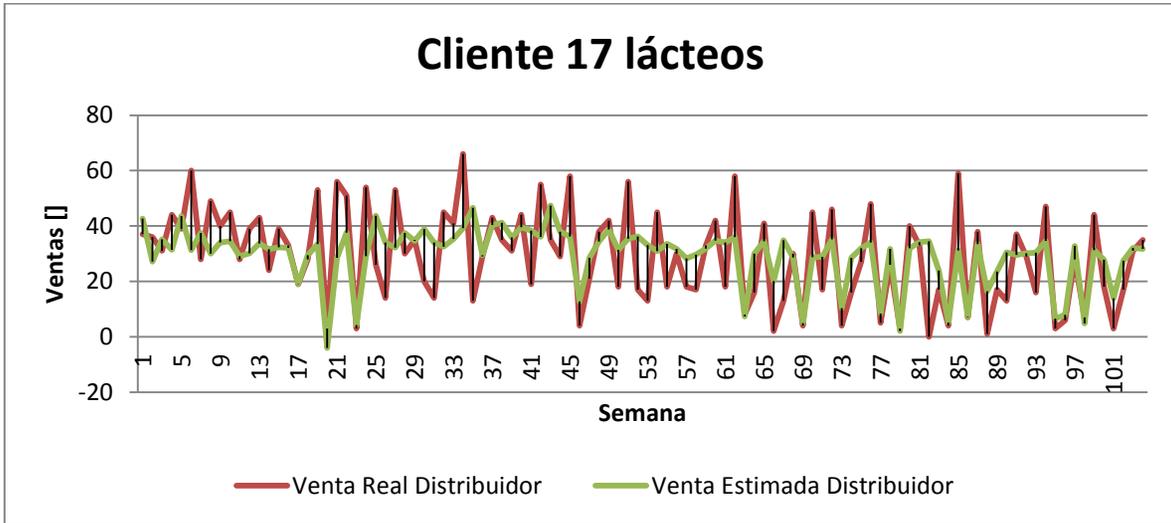
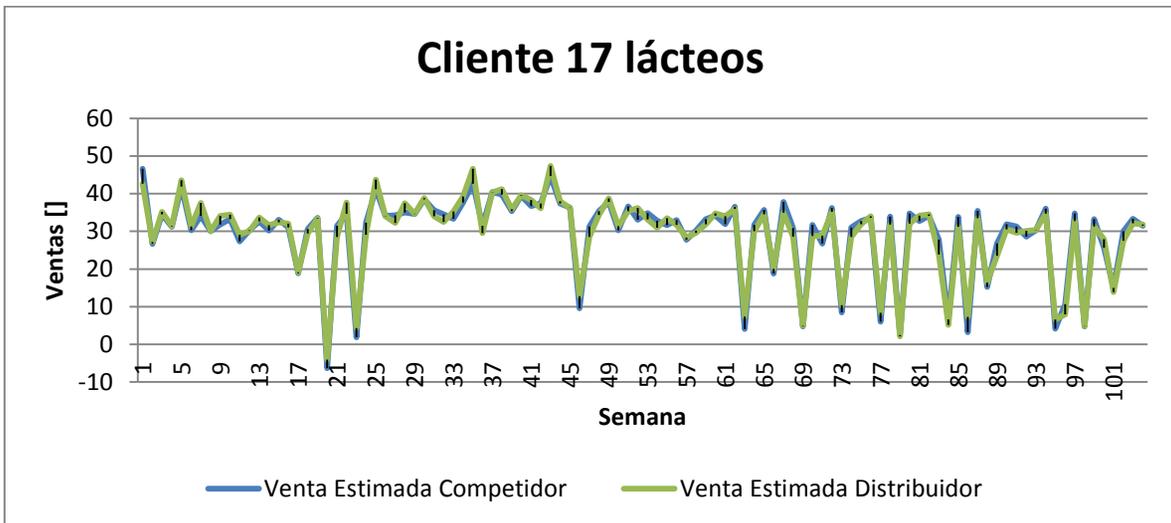


Grafico 14: Estimación Distribuidor vs Competencia - cliente 17- lácteos



De los gráficos mostrados se rescata que hay casos en los cuales la competencia y el distribuidor disputan participación de mercado, mientras en otros las ventas experimentadas son muy similares y siguen la tendencia dada por los restaurantes, los cuales a su vez compran de manera equitativa a los proveedores.

Por otro lado queda de manifiesto errores del modelo para estimar obteniendo resultados negativos, esto se da principalmente en periodos donde las ventas reales son nulas.

8.6 Análisis y validación del modelo

Los criterios que se utilizarán para la validación del modelo serán el error MAPE promedio sobre los valores estimados de las cantidades compradas al distribuidor y el ajuste R^2 entre las ventas reales del distribuidor y las cantidades estimadas.

Para efecto de la estimación y validación del modelo se emplea una de las 5 ejecuciones realizadas para $\psi=1,25$. De todas formas cabe recordar que como quedó de manifiesto en el capítulo 8.4 los valores obtenidos son consistentemente cercanos en las distintas ejecuciones para para $\psi=1,25$.

Tabla 4: Validación modelo

ψ	1,25
MAPE calibración	92,77%
MAPE testeo	119,67%
R^2	0,6977

El MAPE corresponde al promedio de todos los clientes y todas las categorías sobre los valores estimados dado que la estimación con el modelo lineal jerárquico es a nivel individual, es decir, a nivel de categoría y cliente.

El R^2 contempla las 104 semanas, incluye las estimaciones en la etapa de calibración y en la etapa de testeo.

Es interesante de analizar como se ve modificado el error MAPE y el R^2 por categoría para ver si hay diferencias significativas.

Tabla 5: MAPE por categoría

id categoría	Categoría	MAPE estimado	
		Calibración	Testeo
2	Lácteos	110,91%	204,69%
3	Carnes	55,45%	65,46%
5	Aves	122,50%	201,84%
6	Congelados	107,49%	82,58%
7	Enlatados y Secos	67,44%	43,79%

De esta tabla se extrae que las categorías Carnes, Congelados, y Enlatados y Secos son las que tienen un menor error asociado a la estimación. Si consideramos aquellas categorías donde la venta agregada tiene una menor variación en el tiempo y retomamos el gráfico ventas por categorías por semana, podemos ver que el menor

error se condice con una menor variación en las ventas, al menos para los Congelados, y Enlatados y Secos por lo tanto esto se puede ver a que hay algún factor que no se está incluyendo en la estimación de las categorías más variables.

Grafico 15: Ventas semanales por categoría

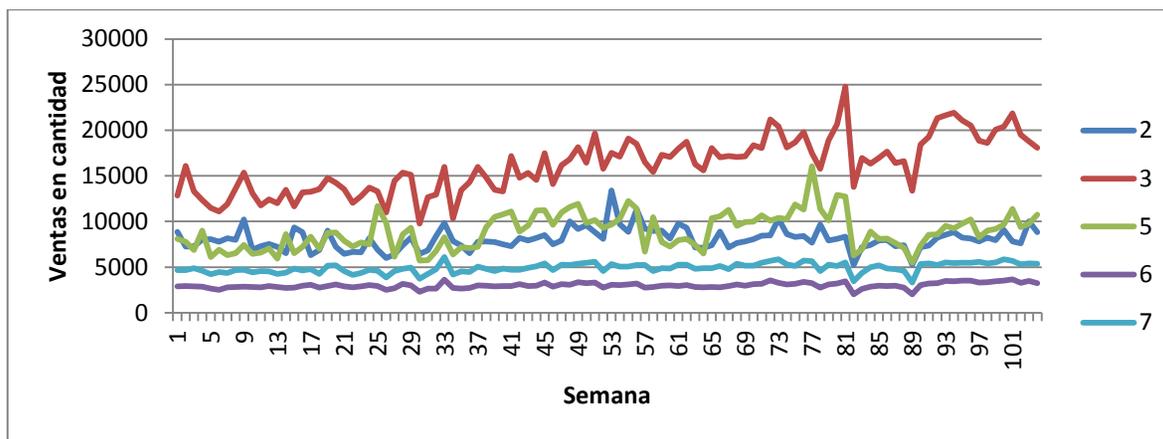


Tabla 6: Ajuste por categoría

Categoría	R ²
Lácteos	0,373938
Carnes	0,348136
Aves	0,402592
Congelados	0,134083
Enlatados y Secos	0,157275

Ahora si tomamos el ajuste R² por categoría podemos ver que las estimaciones que explican de mejor forma son las asociadas a las categorías Lácteos, Carnes y Aves. En consecuencia reuniendo los dos criterios (ajuste y error) es posible afirmar que el modelo predice de buena forma las ventas de la categoría Carnes y cómo influiría la competencia, aunque cabe señalar que esta categoría representa cerca del 40% de las ventas a lo largo del periodo en estudio del distribuidor (ver Anexo A) por tanto la calidad de la estimación puede estar influenciada por la representatividad de la categoría.

Alternativamente realizando una regresión lineal (no jerárquica bayesiana) entre las ventas del distribuidor de referencia y los regresores: tipo restaurante, nivel de precio, exceso de compras, crecimiento en las ventas, ventas en la categoría en el periodo anterior, ventas totales en el periodo anterior, si es el proveedor principal y

número de empleados se tienen los siguientes resultados a nivel individual por categoría:

Tabla 7: Ajuste por categoría a nivel individual

Categoría	R ²
Lácteos	0,30175
Carnes	0,46176
Aves	0,24637
Congelados	0,53416
Enlatados y Secos	0,57621

En este caso los ajustes correspondientes a las categorías congelados, y enlatados y secos son considerablemente mejores que con el modelo lineal jerárquico. Pese a esto dos de las categorías que tienen mayor presencia en las ventas (Lácteos y Aves) poseen un mejor ajuste. Lo cual sumado a que el modelo jerárquico permite identificar patrones comunes de comportamiento y que con las condiciones de primer orden es posible inferir las ventas de la competencia el modelo propuesto es mejor alternativa que una estimación simple de las ventas del distribuidor.

8.7 Interpretación de Resultados y Recomendaciones.

Dentro de los resultados es relevante entender el efecto que tienen las variables explicativas escogidas en la estimación de los parámetros τ_{imt} y δ_{imt} . Para esto se usarán las estimaciones promedios de los coeficientes asociados. (Cabe recordar que los coeficientes β_i corresponden a la ecuación de regresión para la estimación de los requerimientos de los clientes y los coeficientes γ_i corresponden a la ecuación de regresión para la diferencia de precios entre el distribuidor y su competencia.

Tabla 8: Coeficientes estimados β_i, γ_i

		Media
Coeficiente		
Intercepto Lácteos	β_1	202,81
Intercepto Carne	β_2	291,28
Intercepto Aves	β_3	213,79
Intercepto Congelados	β_4	177,6
Intercepto Enlatados y Secos	β_5	190,06
Nivel de Precios	β_6	-3,6168
Exceso compras	β_7	-0,01577
Crecimiento compras	β_8	-11,809
Venta anterior	γ_1	-13,54
Venta total anterior	γ_2	-8,0613

Los interceptos por categoría indican las necesidades básicas en promedio por categoría para los clientes.

Como es de esperar el nivel de precios tiene un efecto negativo en la estimación de las cantidades a comprar por el cliente, vale decir, aumento en los precios disminuye las cantidades a comprar.

Los excesos de compra tienen un efecto poco significativo a nivel general, esto se puede deber a que como se mostro en la descripción de los datos las ventas son relativamente estables al estar los datos agregados.

Los crecimientos en las compras en el periodo anterior respecto a un nivel de compras histórico promedio se traducen en disminución de las compras presentes, esto también es razonable en la medida que un aumento sustancial en las compras en un periodo tiene efecto negativo en las del periodo siguiente.

Finalmente los parámetros asociados a las diferencias de precios tienen efecto negativo en las compras del cliente. Esto quiere decir que las compras en periodos anteriores tiene un efecto negativo, sin embargo, para un mayor entendimiento es útil su razonamiento desde los parámetros obtenidos a nivel de población que se analizarán más adelante.

Cabe señalar que el modelo propuesto por Goic y Montgomery (2011) no considera los coeficientes asociados a inventario (exceso de compras y crecimiento en las ventas), que en los desarrollos preliminares del modelo arrojaron mejor ajuste y un menor error en las estimaciones, con lo cual se deduce la utilidad que tienen para una estimación más completa de la demanda de los clientes.

En esta sección se discutirá las estimaciones posteriores de los parámetros de nivel de población Λ , los cuales permitirán entender el comportamiento de los clientes asociados a determinados perfiles demográficos sujetos a las variables escogidas que se traducirán en recomendaciones para las futuras negociaciones de precios. A modo de contextualización es necesario recordar que dentro de los resultados que entrega la estimación lineal jerárquica se encuentra el efecto de las variables demográficas en los coeficientes asociados a la ecuación lineal previamente definida, que en este caso se deriva de la solución interior del modelo de decisión de compra. Para un mayor entendimiento de este análisis se puede apoyar en los valores posteriores de los parámetros de nivel de población contenidos en el Anexo E

En primer lugar y como era de esperar se presentan diferencias en las necesidades por categoría en función del tipo de restaurante. Así por ejemplo los restaurantes de comida mexicana tienen una necesidad considerablemente menor en promedio de abastecerse de productos Congelados, Aves y Enlatados y Secos.

Tabla 9: Necesidades por categoría restaurante comida mexicana

				MEX	
Intercepto Aves	β_3	Media	Desv. stand.	-29,749	854,803
		90%Intervalo Conf.		-1444,279	1366,172
Intercepto Congelados	β_4	Media	Desv. stand.	-29,903	441,172
		90%Intervalo Conf.		-764,850	699,758
Intercepto Enlatados y Secos	β_5	Media	Desv. stand.	8,155	443,020
		90%Intervalo Conf.		-724,963	733,535

Siguiendo la línea anterior destaca que existan requerimientos base por categoría distintos dada la ubicación de los clientes, esto se puede deber a una mayor concentración de ciertos tipos de restaurantes en determinadas zonas geográficas, lo cual tendría que ser corroborado de manera empírica.

Tabla 10: Necesidad básica por categoría lácteo según ubicación

				U1		U2		U3	
Intercepto Lácteos	β_1	Media	Desv. stand.	65,788	870,361	107,670	871,706	-10,563	873,736
		90%Intervalo Conf.		-1377,062	1507,675	-1330,181	1554,790	-1435,239	1431,440

Respecto a la sensibilidad al nivel de precio experimentada por determinados perfiles de clientes, y que posiblemente sea una de las más interesantes de entender, se tiene que dentro de los tipos de restaurante aquel que presenta una mayor y significativa sensibilidad al nivel de precio es el de comida italiana. (Este coeficiente mide como se modifican los requerimientos con el precio). El valor que toma en los otros restaurantes se puede ver en el anexo E.

Tabla 11: Sensibilidad nivel de precio restaurante comida italiana

				ITL	
Nivel de Precios	β_6	Media	Desv. stand.	-2,881	10,434
		90%Intervalo Conf.		-20,400	14,577

En cuanto a la sensibilidad al nivel de precio dada la ubicación del cliente, aquella definida como U2 es la que presenta una mayor sensibilidad en relación al resto de las ubicaciones. Esta ubicación puede ser asignada al Condado de Fairfax, el cual se caracteriza por tener la media más elevada de ingresos domésticos de las ubicaciones estudiadas y la 2da a nivel de EEUU, además es el centro de la comunidad de negocios y tecnología del área de Washington, D.C⁵. Además manifiestan un grado de sensibilidad apreciable las ubicaciones U0 y U1, de las cuales es importante destacar que al igual que U2, U0 contiene parte del Condado de Fairfax. Esto lleva a tener en mente que la incorporación de una variable asociada a los ingresos de una ubicación en particular puede enriquecer aún más el modelo y las conclusiones que se pueden obtener de él.

Tabla 12: Sensibilidad nivel de precio por ubicación del cliente

				U0		U1		U2	
Nivel de Precios	β_6	Media	Desv. stand.	-1,017	10,352	-1,219	10,356	-2,170	10,377
		90%Intervalo Conf.		-18,366	16,042	-18,528	15,949	-19,387	14,961

⁵ www.fairfaxcountyeda.org/international-business-spanish
http://es.wikipedia.org/wiki/Condado_de_Fairfax

Los coeficientes β_7 y β_8 pretenden capturar el efecto asociado al inventario de los clientes, según el tipo de variable demográfica al que se asocia puede tener efectos negativos o positivos. De la literatura de inventario se extrae que el hecho de que tomen efectos negativos se puede deber a que mayores niveles de inventario puede subir la tasa de consumo, aumentando así la probabilidad de incidencia de compra, por otro lado mayores niveles de inventario reducen la probabilidad de compra como el consumidor siente menos presión para comprar (Boztuğ y Bell, 2004). Posiblemente dentro de los clientes hay casos para los que se impone uno de los efectos. En los resultados no se encontraron diferencias significativas en que se pueda afirmar que el tipo de restaurante sea un determinante, salvo para los restaurantes pizzerías en donde el crecimiento de las compras tiene un efecto negativo superior al del resto de los restaurantes. De todas formas es de destacar para la mayoría de los casos se impone un coeficiente negativo, vale decir, que mayores niveles de inventario reducen las probabilidad de compra, lo cual es razonable al considerar el contexto industrial de los clientes.

Los coeficientes γ_1 y γ_2 están asociados a la diferencia de precios que enfrenta el cliente. En vista que las variables explicativas usadas en la ecuación que determina las diferencias de precios están en función de ventas pasadas, las estimaciones negativas y cercanas a cero que se tienen como resultado, manifiestan que los precios de los clientes, cuyo proveedor principal es el distribuidor de referencia son menos sensibles a compras de los periodos anteriores.

Tabla 13: Efecto diferencia de precios si el cliente tiene al distribuidor como proveedor principal

				Prov. Principal	
Venta anterior	γ_1	Media	Desv. stand.	-0,125	0,079
		90%Intervalo Conf.		-0,256	0,002
Venta total anterior	γ_2	Media	Desv. stand.	-0,061	0,057
		90%Intervalo Conf.		-0,156	0,030

Continuando con los coeficientes γ_1 y γ_2 se repiten las ubicaciones U0, U1 y U2 respecto a la importancia de los precios para estas zonas, son sensibles respecto a las compras en periodos anteriores, lo cual refuerza la tesis que una zona caracterizada por altos ingresos sería más exigente en los procesos de negociación.

Tabla 14: Efecto diferencia de precios por ubicación del cliente

				U0		U1		U2	
Venta anterior	Y ₁	Media	Desv. stand.	-10,973	51,016	-5,553	51,046	-13,405	51,141
		90%Intervalo Conf.		-95,032	72,150	-89,997	78,233	-98,637	69,983
Venta total anterior	Y ₂	Media	Desv. stand.	-5,434	35,844	-1,681	35,873	-8,041	35,972
		90%Intervalo Conf.		-64,374	53,139	-60,611	57,075	-67,237	51,367

Finalmente si se comparan los resultados obtenidos a nivel de parámetros de población con los obtenidos por Goic y Montgomery (2011) (Ver Anexo G) no se tienen mayores diferencias en términos cualitativos, si es efectivo que los parámetros toman valores distintos, pero esto se debe a que se han incluido nuevas variables al problema, tanto demográficas como es el caso de la ubicación de los clientes y explicativas en relación a las consideraciones de inventario incorporadas. Además no se debe olvidar que el modelo tratado en este trabajo en una versión simplificada de la solución al modelo de decisión.

Críticas al modelo: De los resultados obtenidos destaca que en promedio los efectos sean similares a los resultados obtenidos anteriormente, sin embargo, el los mayores intervalos de confianza para los valores de los parámetros de población y considerablemente grandes sin necesidad de hacer comparación refleja un problema que no pudo ser detectado, en donde incluso reduciendo la exigencia del intervalo de confianza a un 80 %, 70% e incluso 60% no se logran disminuciones considerables.

Dentro de las estimaciones para las cantidades del distribuido y de la competencia existen algunos valores que toman signo negativo, esto se presenta principalmente en aquellos instantes en que las ventas reales son cercanas a cero. En principio se intento imponer la condición de que las estimaciones fuesen positivas por medio de que la ecuación derivada de la condición de primer orden impusiera que las cantidades se definieran como la exponencial del resultado finalmente usado. Sin embargo, se obtuvieron valores de error MAPE cercano a 400 % incluso para casos en que el modelo final daba valores de un 92%, e incluso el ajuste R^2 entre el valor real de las ventas del distribuidor y las estimaciones obtenidas era considerablemente bajo (0,0012).

9 Conclusiones

Este trabajo planteó como objetivo principal desarrollar una metodología para apoyar las negociaciones de precios de un distribuidor que abastece con materias primas a restaurantes. La principal dificultad radica en el desconocimiento de la información de las ventas y precios de su competencia debido a que en este contexto industrial las negociaciones son directas entre vendedores y clientes, e incluso desde el mismo proveedor a sus distintos clientes se presentan diferencias en los precios negociados.

Con la finalidad de entregar recomendaciones para los procesos de negociación se debió enfrentar la carencia de información por medio de un modelo econométrico que permitiera endogeneizar la variable no observable precio transado por la competencia y poder estimar la variable ventas de la competencia igualmente desconocida. Para esto fue necesario modelar la forma en que los clientes industriales llevan a cabo sus decisiones de adquisición de materias primas para su procesamiento y por consiguiente venta.

A lo largo del desarrollo del trabajo se llegó en variadas oportunidades a la necesidad de contar con mayor información para poder formular un modelo lo más completo posible, sin embargo, aspectos como calidad del servicio y capacidad de almacenamiento de cada cliente fueron imposibles de utilizar directamente, teniendo que ser descartado el caso del primero y estimado en caso del segundo.

Luego de ser imposible incorporar nuevas apreciaciones al modelo de decisión de compra planteado en trabajos anteriores. Dentro de lo más destacable del modelo desarrollado se encuentra la utilización de un modelo lineal jerárquico bayesiano que estimaba conjuntamente los parámetros del problema sujeto a la solución interior del modelo de decisión utilizado que caracteriza la demanda del distribuidor considerando el efecto de la competencia. Lo positivo de la utilización del modelo antes mencionado es que ha permitido incorporar heterogeneidad de los clientes, vital para identificar clientes con comportamientos de compra similar y formular recomendaciones para negociar sujeto a estos perfiles encontrados como tipo de restaurante, ubicación, entre otros.

En relación a la calidad de la estimación de demanda entregada por el modelo esta es posible validarla en función de la cantidad vendida por distribuidor de referencia al sólo poseer su información transaccional y no la de su competencia. De todas formas se obtuvo un modelo que explica cerca del 70% de la varianza con un error MAPE de un 93% aprox. que es bastante bueno si se considera que se ha trabajado asimilando información desconocida. En donde también se ha dejado en manifiesto que para ciertas categorías la exactitud del modelo cambia.

Dentro de las recomendaciones formuladas destaca que los precios de los clientes cuyo proveedor principal es el distribuidor de referencia son menos sensibles a compras realizadas en el pasado. Otro punto relevante es como la ubicación del cliente juega un papel relevante en prácticamente la totalidad de las variables explicativas del modelo, desde diferencias en los requerimientos básicos por categorías, variaciones en la sensibilidad al nivel de precios y las variables que describen las diferencias de precios.

Para trabajos futuros existen dos aristas por mejorar, la primera de ellas tiene que ver con un enriquecimiento del modelo que acá se plantea en términos de incorporar la variable demográfica definida de otra forma asignándoles valores como el ingreso promedio de la zona. En esta misma línea otro aspecto a mejorar es las consideraciones que se imponen en el modelo de estimación asociada a las categorías, pues como se pudo ver existen diferencias en la calidad de la estimación según el tipo de categoría, quizás un aspecto a incluir sería un efecto de estacionalidad específico por categoría que en este trabajo no se tomó en consideración.

El otro aspecto que no pudo ser abordado radica en el diseño del modelo de decisión sobre el cual podría estudiarse formas funcionales más complejas o incorporar otros factores que permitan diferenciar los beneficios asociados a la adquisición de distintos proveedores y no determinar la elección por uno u otro a la minimización de los costos de adquisición, siempre y cuando existan estas diferenciaciones puesto que como se mencionó en este trabajo los productos que comercializa el distribuidor son más bien commodities, sin embargo, el servicio asociado a la distribución puede significar una diferenciación en los beneficios que observa el cliente.

Bibliografía

Boztuğ, Yasemin; Bell, David R. (2004) "The Effect of Inventory on Purchase Incidence: Empirical Analysis of Opposing Forces of Storage and Consumption", Center for Applied Statistics and Economics (CASE), No. 2004,43.

Cameron S. y Shipley D. D. (1985), "A Discretionary Model of Industrial Buying", *Managerial y Decision Economics*, Vol. 6, Issue 2, Pages 102-111, June 1985.

Cardozo R. N., Csgley J. W. (1985) "Experimental study of industrial buyer behavior", *Journal of Marketing*. Vol. 8, No. 3 (Aug., 1971), Pages 329-334.

Çebi F. y Bayraktar D., (2003) "An integrated approach for supplier selection", *Logistics Information Management*, Vol. 16 Issue 6, Pages 395-400.

Chen L. y Plambeck E. L. (2008), "Dynamic Inventory Management with Learning About the Demand Distribution y Substitution Probability", *Journal Manufacturing & Service Operations Management*. Vol. 10, Pages 2236-256.

Christophe A., y Thoms J. (2008) "A tutorial on adaptive MCMC," *Journal Statistics y Computing*, Vol 18, No 4, Pages 343-373.

Dolan R. J. (1995), How do you know when the Price is right?, *Harvard Business Review*.

Figuroa Gostín, Trinidad (2009). Modelo predictivo de quiebres de stock en un supermercado. Memoria para optar al título de Ingeniero Civil Industrial. Santiago, Universidad de Chile, Facultad de Ciencias Físicas y Matemáticas, Departamento de Ingeniería Industrial

Goic M. y Montgomery A. L., *Inferring Competitor Pricing with Incomplete Information*, Working Paper, January 2011.

Goic M., Jerath K. y Srinivasan K. (2011), "Cross-Market Discounts", *Marketing Science*. Vol. 30, Issue 1, Pages 134-148.

Guneri A. F, Yucel A. y Ayyildiz G. (2009), "An integrated fuzzy-lp approach for a supplier selection problem in supply chain management", *Expert Systems with Applications*, Vol 36, Issue 5, Pages 9223–9228.

Heikki H., Saksman E. y Tamminen J. (2001) "An Adaptive Metropolis Algorithm" *Bernoulli*, Vol. 7, No. 2, Pages 223-242.

Lyon T. (2006) "Does Dual Sourcing Lower Procurement Costs?" *The Journal of Industrial Economics*. *Journal of Industrial Economics*, Vol. 54, No. 2, Pages 223-252, June 2006.

Meyer R. J. and Assuncao J., (1990) "The Optimality of Consumer Stockpiling Strategies", Marketing Science, Vol. 9, No. 1 (Winter, 1990), Pages 18-41.

Puto C. P., Wesley E. Patton III, King R. H. (1985), "Risk Handling Strategies in Industrial Vendor Selection Decision", Journal of Marketing. Vol. 49, No. 1 (Winter, 1985), Pages 89-98.

Rangan V. K., Isaacson B. (1994), ¿Qué es marketing industrial?, HBR 23 de Septiembre de 1994.

Reuse Tello, Felipe Ignacio (2011). Estimación del valor de clientes fidelizados en una cadena de supermercados utilizando modelos jerárquicos bayesianos. Memoria para optar al título de Ingeniero Civil Industrial. Santiago, Universidad de Chile, Facultad de Ciencias Físicas y Matemáticas, Departamento de Ingeniería Industrial.

Roberts G. O. y Rosenthal J. S. (2008), "Examples of Adaptive MCMC", Journal of Computational y Graphical Statistics, Vol. 18, No. 2, Pages 349-367.

Rossi E. Peter, Allenby M. Greg, McCulloch Robert (2005). Bayesian Statistics and Marketing, John Wiley & Sons.

Shugan, Steven (2004) "Endogeneity in Marketing Decision Models", Marketing Science, Vol 23, No 1, Pages 1-3.

Troncoso Rojas, Catalina Pía (2011). Determinación de precios óptimos para una cadena de supermercado utilizando modelos jerárquicos bayesianos Memoria para optar al título de Ingeniero Civil Industrial. Santiago, Universidad de Chile, Facultad de Ciencias Físicas y Matemáticas, Departamento de Ingeniería Industrial.

Anexos

Anexo A: Tipo de información

Información contenida en la encuesta:

- Tipo de restaurante (comida americana, comida italiana, comida mexicana y pizzería)
- Código postal
- Número de empleados
- Volumen actual de ventas
- Compra “estimada”
- Gasto total por semana
- Metros cuadrados
- Si es el distribuidor de referencia el proveedor principal
- Frecuencia de visita
- Frecuencia de entrega

La base de datos transaccional está conformada de la siguiente forma:

Tabla 15: Muestra base de dato transaccional

Cust	Cat	Week	totQty	totExp	totCost	NDiffProd
5173	2	1,00	6,00	165,47	114,24	5,00

Donde:

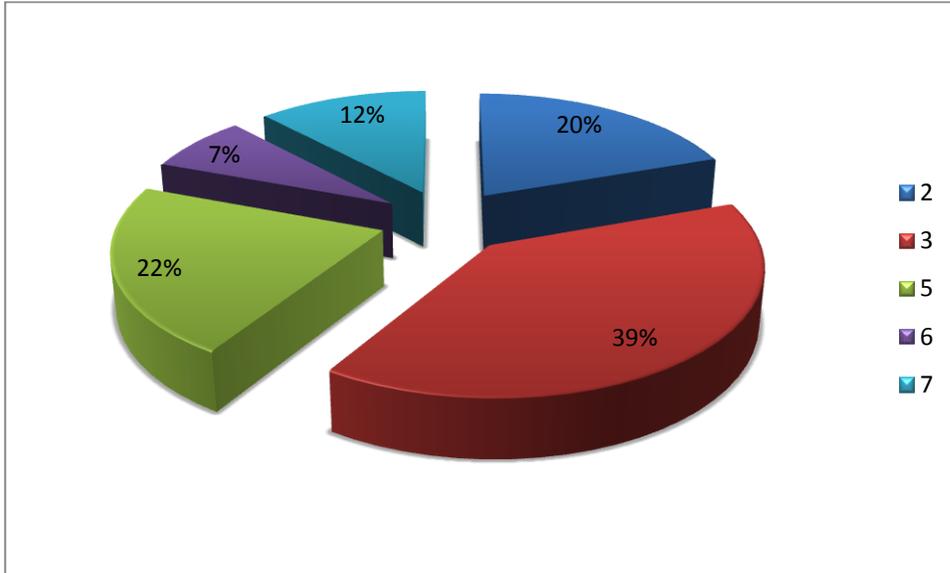
- cust: identificador del cliente.
- cat: identificador de categoría.
- week: semana de la venta.
- totQty: cantidad vendida.
- totExp: ingreso por venta.
- totCost: costo por venta.
- NDiffProd: A cuantos productos (SKUs distintos) corresponde la orden.

Categorías:

Tabla 16: Identificador por categoría

id categoría	descripción categoría
2	productos lácteos
3	carnes
5	aves de corral
6	congelado
7	enlatados y secos

Grafico 16: Proporción de ventas por categoría



Anexo B: Análisis Estacionalidad

Grafico 17: Ventas Lácteos año 1 y 2

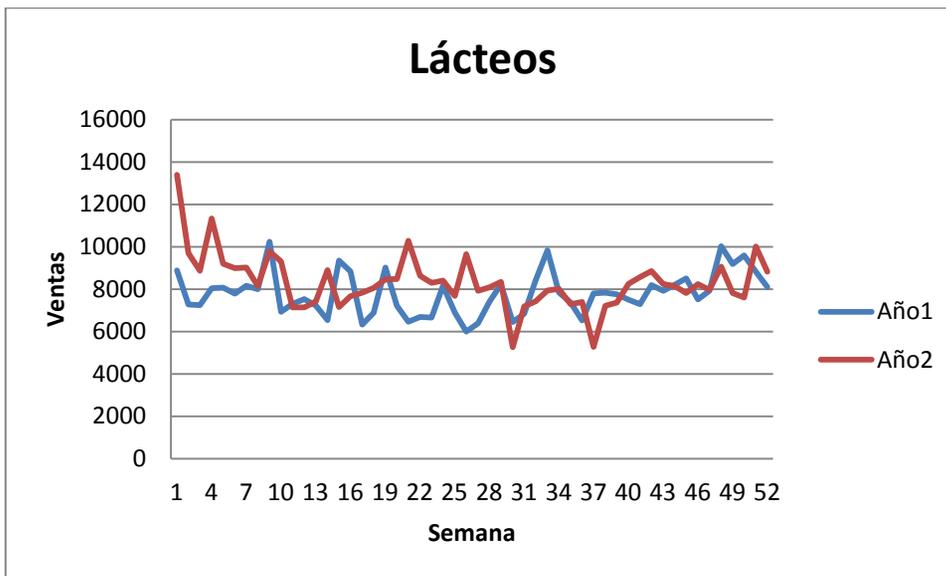


Grafico 18: Ventas Congelados año 1 y 2

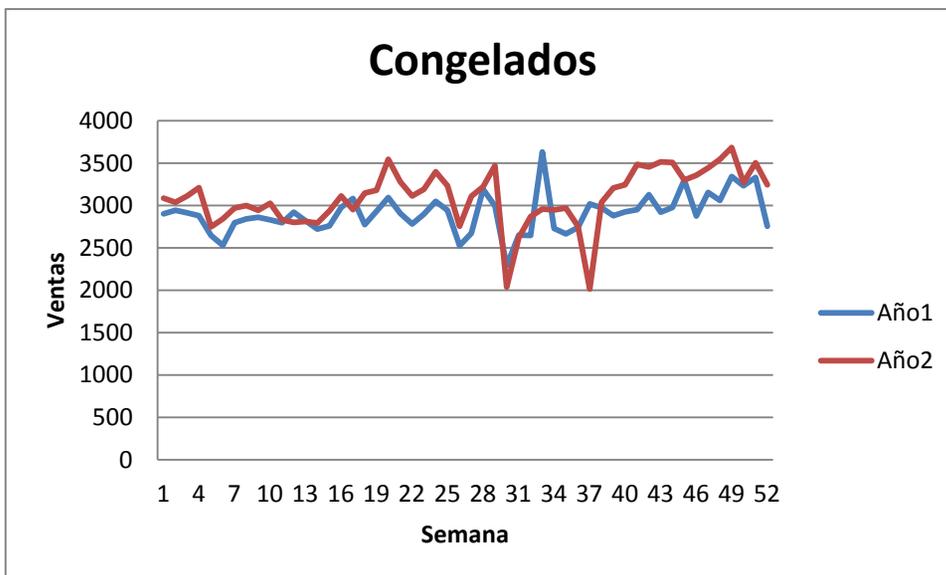
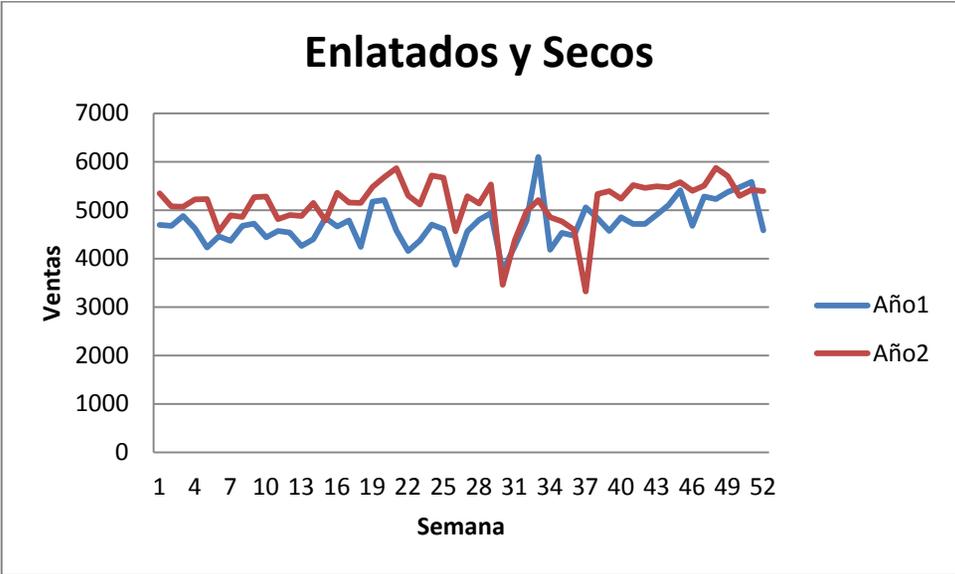


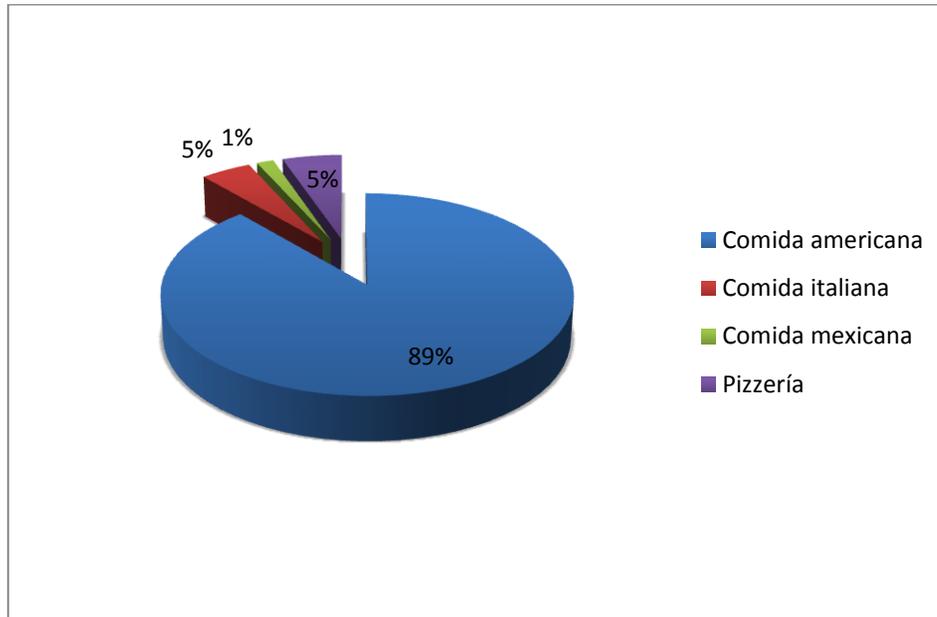
Grafico 19: Ventas Enlatados y Secos año 1 y 2



Anexo C: Descripción variables demográficas:

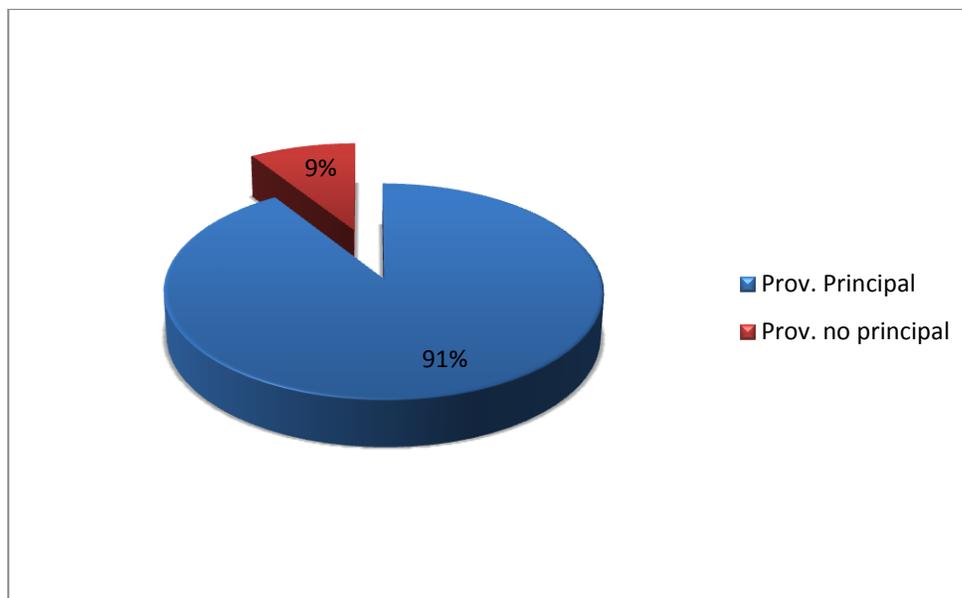
Tipo restaurante al que pertenecen los clientes:

Grafico 20: Clientes por tipo Restaurante



Proveedor principal:

Grafico 21: Proporción de clientes donde la empresa es su proveedor principal



Número de empleados:

Tabla 17: Estadísticos número de empleados

Promedio	28,07
Desv. Estándar	21,34
Min	1
Max	125

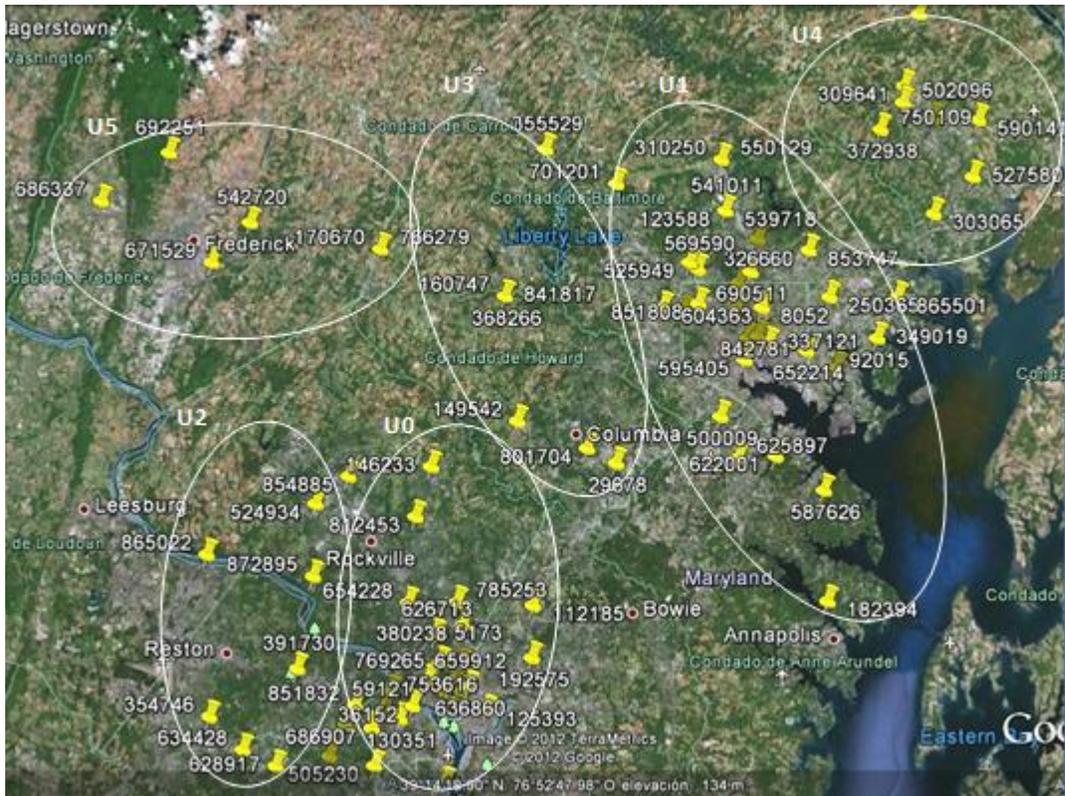
Ubicación:

Tabla 18: Clientes por ubicación

Ubicación	Total
0	47
1	47
2	12
3	8
4	10
5	6
Total general	130

Estás ubicaciones referenciadas en Google Earth se distribuyen de la siguiente forma:

Ilustración 3: Ubicación Clientes



Anexo D: Determinación de ψ

Se definen:

MAPE calibración y R2 calibración: indicadores sobre los valores estimados en los resultados dados en la etapa de calibración (88 semanas).

Tabla 19: Validación $\phi=0,25$

Phi = 0,25	Ejecución 1	Ejecución 2	Ejecución 3	Ejecución 4	Ejecución 5	Promedio
MAPE calibración	1,1262	10,898	1,1705	1,173	1,173	3,1081
R2 calibración	0,69057	0,69052	0,6905	0,69044	0,69044	0,6905

Tabla 20: Validación $\phi=0,5$

Phi = 0,5	Ejecución 1	Ejecución 2	Ejecución 3	Ejecución 4	Ejecución 5	Promedio
MAPE calibración	1,0139	1,2568	0,87969	1,0691	1,0558	1,0574
R2 calibración	0,69028	0,69029	0,69026	0,69028	0,69035	0,6903

Tabla 21: Validación $\phi=0,75$

Phi = 0,75	Ejecución 1	Ejecución 2	Ejecución 3	Ejecución 4	Ejecución 5	Promedio
MAPE calibración	1,1299	1,4549	1,1908	0,99921	0,99921	1,1548
R2 calibración	0,69022	0,69016	0,69019	0,69019	0,69019	0,6902

Tabla 22: Validación $\phi=1,5$

Phi = 1,5	Ejecución 1	Ejecución 2	Ejecución 3	Ejecución 4	Ejecución 5	Promedio
MAPE calibración	1,1427	1,1358	1,2559	1,1258	1,0679	1,1456
R2 calibración	0,69015	0,69019	0,69016	0,69029	0,69014	0,6902

Tabla 23: Validación $\phi=1,75$

Phi = 1,75	Ejecución 1	Ejecución 2	Ejecución 3	Ejecución 4	Ejecución 5	Promedio
MAPE calibración	1,4541	1,1134	0,95805	0,98485	1,095	1,1211
R2 calibración	0,69018	0,69025	0,19012	0,69013	0,6903	0,5902

Tabla 24: Validación phi=2

Phi = 2	Ejecución 1	Ejecución 2	Ejecución 3	Ejecución 4	Ejecución 5	Promedio
MAPE calibración	5,6779	0,98819	0,9439	5,6779	1,2436	2,9063
R2 calibración	0,69028	0,68997	0,69027	0,69028	0,69022	0,6902

Tabla 25: Validación phi=1

Phi = 1	Ejecución 1	Ejecución 2	Ejecución 3	Ejecución 4	Ejecución 5	Promedio
MAPEest calibración	1,1592	0,96028	1,276	1,1421	1,667	1,2409
R2 calibración	0,69026	0,69028	0,69033	0,6902	0,69025	0,6903

Se tenía presupuestado realizar 5 ejecuciones por valor de phi, pero considerando que uno de los indicadores de la ejecución 5 para phi = 1 se escapa de los resultados que se estaban obteniendo, se realizaron 6 ejecuciones para phi = 1,25.

Tabla 26: Validación phi=1,25

Phi = 1,25	Ejecución 1	Ejecución 2	Ejecución 3	Ejecución 4	Ejecución 5	Ejecución 6	Promedio
MAPE calibración	0,92768	1,1325	0,98044	1,1451	0,94166	1,0163	1,0239
R2 calibración	0,69029	0,69023	0,69016	0,69009	0,6903	0,69016	0,6902

Anexo E: Parámetros estimados a nivel de población

Tabla 27: Parámetros estimados a nivel de población $\phi=1,25$

			AMR		ITL		MEX		PZZ		Prov. Principal		N° empleados	
β_1	Media	Desv. stand.	-27,281	871,785	192,103	875,816	-71,348	886,845	387,131	878,599	1,651	1,169	81,616	93,560
	90%Intervalo Conf.		-1455,456	1414,308	-1255,131	1632,462	-1533,929	1375,370	-1070,892	1832,232	-0,241	3,590	-74,185	234,934
β_2	Media	Desv. stand.	169,747	741,377	133,074	743,723	172,162	757,118	146,689	747,942	1,257	0,993	-72,535	80,994
	90%Intervalo Conf.		-1058,809	1370,811	-1101,342	1321,286	-1082,398	1391,161	-1080,825	1369,902	-0,379	2,888	-204,320	61,773
β_3	Media	Desv. stand.	41,190	838,205	103,177	844,343	-29,749	854,803	76,117	845,147	2,500	1,123	34,426	91,460
	90%Intervalo Conf.		-1359,060	1413,081	-1304,658	1486,985	-1444,279	1366,172	-1322,403	1456,809	0,646	4,355	-117,038	185,096
β_4	Media	Desv. stand.	43,069	432,393	117,524	435,262	-29,903	441,172	93,425	436,046	1,293	0,605	33,894	48,407
	90%Intervalo Conf.		-677,745	762,939	-609,199	842,926	-764,850	699,758	-626,382	815,833	0,297	2,268	-44,137	113,356
β_5	Media	Desv. stand.	47,622	434,664	153,894	437,434	8,155	443,020	90,604	438,455	1,235	0,601	28,648	48,120
	90%Intervalo Conf.		-667,239	758,892	-569,313	870,443	-724,963	733,535	-635,222	807,664	0,251	2,214	-49,975	107,745
β_6	Media	Desv. stand.	-0,288	10,365	-2,881	10,434	0,766	10,573	-0,729	10,454	-0,032	0,015	-1,459	1,158
	90%Intervalo Conf.		-17,394	17,011	-20,400	14,577	-16,808	18,299	-18,018	16,572	-0,056	-0,008	-3,348	0,420
β_7	Media	Desv. stand.	0,144	1,832	-0,198	1,842	0,227	1,875	0,448	1,847	0,003	0,002	-0,377	0,202
	90%Intervalo Conf.		-2,916	3,089	-3,275	2,749	-2,936	3,214	-2,627	3,425	-0,001	0,007	-0,709	-0,042
β_8	Media	Desv. stand.	-7,575	65,681	-5,830	65,904	-7,767	67,139	-24,643	66,371	-0,092	0,099	9,090	7,924
	90%Intervalo Conf.		-114,686	101,106	-113,392	104,279	-118,728	103,917	-134,097	83,104	-0,256	0,069	-3,929	22,273
γ_1	Media	Desv. stand.	0,498	51,045	-18,887	51,304	3,868	52,207	-18,026	51,552	-0,125	0,079	-1,746	6,248
	90%Intervalo Conf.		-82,556	84,917	-103,859	64,835	-80,857	89,619	-102,319	66,526	-0,256	0,002	-12,040	8,358
γ_2	Media	Desv. stand.	-0,616	35,843	-5,942	36,008	2,393	36,796	-7,711	36,193	-0,061	0,057	-2,254	4,551
	90%Intervalo Conf.		-59,796	57,788	-65,532	52,846	-57,417	63,027	-67,518	51,592	-0,156	0,030	-9,757	5,160

			U0		U1		U2		U3		U4		U5	
β1	Media	Desv. stand.	118,169	871,100	65,788	870,361	107,670	871,706	-10,563	873,736	-8,547	872,573	114,141	874,122
	90%Intervalo Conf.		-1320,616	1551,180	-1377,062	1507,675	-1330,181	1554,790	-1435,239	1431,440	-1446,593	1429,393	-1328,223	1553,057
β2	Media	Desv. stand.	199,640	739,877	160,453	739,825	173,137	741,480	16,788	743,327	66,374	741,793	46,306	742,416
	90%Intervalo Conf.		-986,946	1409,536	-1032,045	1374,884	-1025,500	1386,844	-1187,302	1231,259	-1139,531	1283,210	-1160,869	1258,333
β3	Media	Desv. stand.	96,437	838,158	72,253	838,942	72,914	841,118	2,992	840,787	4,509	840,663	-21,736	845,566
	90%Intervalo Conf.		-1279,016	1494,239	-1313,499	1469,266	-1319,603	1466,767	-1374,681	1405,811	-1382,870	1409,821	-1418,839	1377,362
β4	Media	Desv. stand.	78,518	432,080	71,956	432,248	68,790	433,235	9,951	433,729	17,586	433,218	-5,665	435,275
	90%Intervalo Conf.		-637,261	798,880	-649,519	791,278	-651,046	793,761	-711,608	733,694	-706,314	741,259	-733,708	716,212
β5	Media	Desv. stand.	94,789	434,214	82,864	434,299	90,925	435,241	16,804	435,915	15,909	435,427	5,745	437,082
	90%Intervalo Conf.		-617,518	814,204	-624,691	803,389	-623,520	815,159	-702,409	740,476	-705,527	738,724	-713,108	729,113
β6	Media	Desv. stand.	-1,017	10,352	-1,219	10,356	-2,170	10,377	0,148	10,392	0,020	10,378	0,647	10,430
	90%Intervalo Conf.		-18,366	16,042	-18,528	15,949	-19,387	14,961	-17,232	17,424	-17,293	17,262	-16,547	18,105
β7	Media	Desv. stand.	0,168	1,832	0,114	1,830	-0,199	1,836	0,100	1,837	-0,037	1,838	0,212	1,838
	90%Intervalo Conf.		-2,752	3,242	-2,820	3,195	-3,151	2,877	-2,860	3,171	-2,985	3,034	-2,741	3,306
β8	Media	Desv. stand.	-15,321	65,661	-7,627	65,542	1,033	65,752	-2,191	65,849	-3,354	65,763	-11,991	65,878
	90%Intervalo Conf.		-124,565	92,028	-115,729	99,944	-107,998	108,067	-111,281	104,649	-112,344	104,337	-121,109	95,924
γ1	Media	Desv. stand.	-10,973	51,016	-5,553	51,046	-13,405	51,141	1,401	51,199	-1,249	51,160	0,424	51,275
	90%Intervalo Conf.		-95,032	72,150	-89,997	78,233	-98,637	69,983	-82,452	85,538	-85,804	82,385	-84,407	84,180
γ2	Media	Desv. stand.	-5,434	35,844	-1,681	35,873	-8,041	35,972	2,420	35,949	-0,060	35,946	0,993	36,074
	90%Intervalo Conf.		-64,374	53,139	-60,611	57,075	-67,237	51,367	-56,722	61,551	-59,352	58,973	-58,273	60,130

Anexo F: Parámetros a nivel de población para $\phi = 1$ y $\phi = 0,75$

Tabla 28: Parámetros estimados a nivel de población para $\phi=1$

			AMR		ITL		MEX		PZZ		Prov.Principal		Nempleados	
β_1	Media	Desv.stand.	-35,386	869,032	185,638	872,376	-77,280	885,580	379,305	875,514	1,662	1,157	82,437	94,967
	90%Intervalo Conf.		-1438,828	1416,207	-1220,090	1648,436	-1515,009	1383,173	-1028,196	1840,695	-0,232	3,558	-72,675	239,862
β_2	Media	Desv.stand.	169,962	742,186	133,741	745,379	173,672	755,712	147,558	746,821	1,261	1,000	-71,851	81,334
	90%Intervalo Conf.		-1041,474	1391,390	-1085,169	1357,688	-1065,883	1415,736	-1079,179	1369,487	-0,392	2,915	-206,225	59,376
β_3	Media	Desv.stand.	24,820	820,317	87,176	824,637	-43,011	839,032	59,963	825,932	2,504	1,118	36,565	92,194
	90%Intervalo Conf.		-1305,347	1380,610	-1243,255	1465,195	-1389,612	1336,219	-1274,562	1437,268	0,656	4,332	-114,049	188,864
β_4	Media	Desv.stand.	35,902	431,515	110,952	433,023	-35,819	440,601	86,057	434,768	1,314	0,603	35,721	48,984
	90%Intervalo Conf.		-661,975	751,226	-581,585	832,792	-746,334	696,234	-614,287	808,534	0,327	2,312	-44,922	116,377
β_5	Media	Desv.stand.	41,659	430,221	147,743	431,630	3,514	438,629	85,012	433,586	1,244	0,595	29,839	48,853
	90%Intervalo Conf.		-652,336	749,763	-536,191	865,781	-707,670	724,006	-622,211	799,443	0,247	2,221	-50,401	110,799
β_6	Media	Desv.stand.	-0,152	10,283	-2,761	10,318	0,860	10,493	-0,619	10,357	-0,033	0,015	-1,499	1,170
	90%Intervalo Conf.		-17,287	16,549	-20,117	13,966	-16,516	17,858	-17,827	16,284	-0,057	-0,009	-3,450	0,395
β_7	Media	Desv.stand.	0,119	1,851	-0,224	1,863	0,208	1,887	0,430	1,863	0,003	0,003	-0,375	0,204
	90%Intervalo Conf.		-2,881	3,108	-3,252	2,800	-2,878	3,292	-2,597	3,439	-0,001	0,007	-0,716	-0,041
β_8	Media	Desv.stand.	-7,268	64,665	-5,556	64,915	-7,304	66,023	-23,332	65,237	-0,098	0,096	8,921	7,850
	90%Intervalo Conf.		-114,914	99,078	-113,622	99,434	-116,332	100,576	-132,010	82,107	-0,256	0,058	-3,993	21,733
γ_1	Media	Desv.stand.	0,547	39,343	-14,535	39,470	2,952	40,116	-12,724	39,724	-0,097	0,060	-1,387	4,982
	90%Intervalo Conf.		-64,022	63,918	-79,545	49,409	-62,481	68,727	-77,799	51,669	-0,196	0,001	-9,684	6,700
γ_2	Media	Desv.stand.	-0,076	30,281	-4,827	30,336	2,370	30,881	-7,228	30,549	-0,053	0,047	-1,889	3,874
	90%Intervalo Conf.		-49,139	49,677	-54,099	44,914	-47,500	53,098	-56,942	43,122	-0,133	0,024	-8,256	4,411

			U0		U1		U2		U3		U4		U5	
β1	Media	Desv. stand.	125,657	866,949	72,900	866,873	115,874	869,016	-2,105	870,047	-0,254	869,432	121,586	872,118
	90%Intervalo Conf.		-1438,828	-1313,968	1522,955	-1369,792	1474,917	-1328,462	1519,250	-1451,760	1410,184	-1451,592	1410,015	-1335,518
β2	Media	Desv. stand.	198,505	742,032	159,137	742,478	172,249	743,321	15,199	745,345	67,016	744,421	45,222	746,691
	90%Intervalo Conf.		-1041,474	-1023,053	1407,098	-1059,387	1360,383	-1051,425	1389,962	-1212,396	1230,603	-1163,816	1270,511	-1178,172
β3	Media	Desv. stand.	110,369	819,930	86,927	819,084	86,075	821,936	17,704	824,393	19,110	821,832	-6,051	825,418
	90%Intervalo Conf.		-1305,347	-1258,971	1439,685	-1287,228	1402,719	-1283,549	1397,282	-1351,561	1345,993	-1352,586	1341,688	-1388,393
β4	Media	Desv. stand.	83,697	431,544	77,312	430,958	73,700	432,246	15,246	433,425	22,720	432,378	-1,798	434,212
	90%Intervalo Conf.		-661,975	-631,630	780,748	-639,889	771,763	-645,798	770,748	-702,350	715,328	-689,482	719,532	-723,809
β5	Media	Desv. stand.	99,583	429,997	87,641	429,592	95,626	430,853	21,591	432,008	21,144	431,153	10,238	432,861
	90%Intervalo Conf.		-652,336	-613,784	790,459	-622,510	776,023	-618,094	790,938	-692,643	711,315	-699,751	712,886	-705,322
β6	Media	Desv. stand.	-1,109	10,285	-1,314	10,270	-2,257	10,300	0,056	10,338	-0,079	10,307	0,561	10,342
	90%Intervalo Conf.		-17,287	-17,851	15,974	-17,946	15,822	-18,972	15,037	-16,750	17,206	-16,745	17,146	-16,075
β7	Media	Desv. stand.	0,196	1,852	0,140	1,850	-0,180	1,856	0,125	1,858	-0,016	1,854	0,238	1,862
	90%Intervalo Conf.		-2,881	-2,788	3,238	-2,842	3,166	-3,197	2,850	-2,897	3,167	-3,025	3,021	-2,761
β8	Media	Desv. stand.	-15,495	64,530	-7,861	64,592	0,754	64,695	-2,483	64,739	-3,450	64,676	-11,590	65,018
	90%Intervalo Conf.		-114,914	-120,811	92,478	-113,204	99,499	-104,493	107,902	-108,139	106,361	-108,019	104,084	-117,752
γ1	Media	Desv. stand.	-8,755	39,240	-4,719	39,186	-10,982	39,340	0,651	39,377	-1,378	39,423	0,357	39,463
	90%Intervalo Conf.		-64,022	-72,991	55,652	-68,780	59,934	-75,499	53,720	-63,503	65,603	-65,603	63,660	-64,145
γ2	Media	Desv. stand.	-4,846	30,176	-1,700	30,153	-6,705	30,248	1,913	30,277	-0,216	30,314	0,144	30,371
	90%Intervalo Conf.		-49,139	-54,260	43,825	-50,789	47,009	-56,132	42,124	-48,010	51,154	-49,859	48,751	-49,724

Tabla 29: Parámetros estimados a nivel de población para $\phi=0,75$

			AMR		ITL		MEX		PZZ		Prov.Principal		Nempleados	
β_1	Media	Desv.stand.	-35,577	868,282	185,662	871,780	-77,058	884,098	379,362	875,241	1,663	1,160	82,529	94,987
	90%Intervalo Conf.		-1447,352	1410,705	-1232,038	1653,808	-1520,656	1384,564	-1047,417	1837,514	-0,242	3,555	-71,935	238,948
β_2	Media	Desv.stand.	168,570	742,976	132,364	746,208	172,315	756,303	146,142	747,539	1,262	1,001	-71,750	81,185
	90%Intervalo Conf.		-1041,754	1392,952	-1081,886	1365,334	-1066,587	1415,736	-1091,285	1366,430	-0,394	2,922	-206,519	60,227
β_3	Media	Desv.stand.	24,374	820,477	87,165	824,943	-43,899	839,958	59,649	826,231	2,502	1,119	36,662	92,277
	90%Intervalo Conf.		-1300,748	1379,737	-1247,632	1461,107	-1401,994	1334,638	-1272,922	1428,273	0,649	4,339	-115,565	189,394
β_4	Media	Desv.stand.	35,698	431,417	111,039	433,145	-36,089	440,726	86,187	434,710	1,314	0,605	35,825	49,005
	90%Intervalo Conf.		-660,144	746,709	-585,818	835,006	-749,184	692,440	-615,445	803,046	0,323	2,318	-45,109	116,469
β_5	Media	Desv.stand.	41,461	430,314	147,821	431,918	3,246	438,803	85,103	433,674	1,245	0,597	29,933	48,864
	90%Intervalo Conf.		-652,591	749,256	-547,433	862,946	-707,025	723,277	-616,218	800,490	0,240	2,231	-50,262	110,704
β_6	Media	Desv.stand.	-0,148	10,282	-2,764	10,321	0,866	10,496	-0,623	10,357	-0,033	0,015	-1,501	1,171
	90%Intervalo Conf.		-17,279	16,468	-20,186	13,824	-16,502	17,722	-17,828	16,019	-0,057	-0,009	-3,458	0,402
β_7	Media	Desv.stand.	0,118	1,853	-0,223	1,865	0,208	1,888	0,429	1,865	0,003	0,003	-0,374	0,204
	90%Intervalo Conf.		-2,879	3,097	-3,269	2,802	-2,876	3,277	-2,607	3,444	-0,001	0,007	-0,715	-0,039
β_8	Media	Desv.stand.	-7,218	64,760	-5,567	65,014	-7,314	66,141	-23,324	65,331	-0,098	0,096	8,896	7,846
	90%Intervalo Conf.		-114,825	99,068	-113,029	100,149	-116,880	100,751	-132,956	81,825	-0,259	0,060	-4,041	21,850
γ_1	Media	Desv.stand.	0,426	29,372	-10,904	29,471	2,189	29,948	-9,510	29,657	-0,071	0,045	-1,064	3,718
	90%Intervalo Conf.		-47,767	47,850	-60,065	36,694	-46,626	50,978	-58,076	38,583	-0,144	0,002	-7,219	4,968
γ_2	Media	Desv.stand.	-0,061	23,020	-3,625	23,056	1,804	23,473	-5,464	23,232	-0,041	0,036	-1,400	2,951
	90%Intervalo Conf.		-37,256	37,593	-41,088	33,913	-36,055	40,348	-43,374	32,660	-0,102	0,017	-6,247	3,395

			U0		U1		U2		U3		U4		U5	
β1	Media	Desv.stand.	125,749	866,404	72,886	866,189	115,734	868,426	-2,194	869,479	-0,044	868,937	121,613	871,377
	90%Intervalo Conf.		-1314,667	1536,803	-1359,996	1491,824	-1327,514	1536,842	-1444,572	1423,736	-1448,055	1420,691	-1329,545	1540,795
β2	Media	Desv.stand.	199,715	742,807	160,398	743,132	173,517	743,894	16,519	745,838	68,425	745,074	46,459	747,406
	90%Intervalo Conf.		-1028,772	1402,797	-1060,944	1362,238	-1059,722	1381,166	-1211,430	1226,919	-1155,315	1273,606	-1192,241	1253,278
β3	Media	Desv.stand.	110,610	820,057	87,298	819,225	86,647	822,104	18,270	824,595	19,366	821,867	-5,422	825,294
	90%Intervalo Conf.		-1243,278	1439,689	-1282,781	1410,476	-1269,203	1404,203	-1335,821	1346,970	-1344,978	1354,742	-1376,374	1312,785
β4	Media	Desv.stand.	83,725	431,502	77,394	430,877	73,927	432,113	15,322	433,479	22,815	432,356	-1,630	434,150
	90%Intervalo Conf.		-631,348	779,416	-638,169	771,567	-641,389	770,693	-704,025	716,320	-690,616	719,809	-723,269	704,736
β5	Media	Desv.stand.	99,615	430,156	87,713	429,687	95,838	430,906	21,648	432,268	21,220	431,365	10,354	432,941
	90%Intervalo Conf.		-607,848	794,069	-620,377	783,148	-615,651	793,707	-692,323	715,945	-696,435	718,434	-708,303	709,522
β6	Media	Desv.stand.	-1,109	10,284	-1,316	10,268	-2,262	10,298	0,056	10,339	-0,079	10,305	0,560	10,342
	90%Intervalo Conf.		-17,677	16,077	-17,759	15,777	-18,929	14,871	-16,623	17,176	-16,644	17,118	-16,135	17,883
β7	Media	Desv.stand.	0,195	1,854	0,139	1,852	-0,179	1,858	0,125	1,860	-0,017	1,856	0,238	1,864
	90%Intervalo Conf.		-2,792	3,244	-2,845	3,164	-3,210	2,852	-2,897	3,156	-3,025	3,012	-2,747	3,312
β8	Media	Desv.stand.	-15,495	64,641	-7,853	64,691	0,733	64,801	-2,513	64,839	-3,498	64,785	-11,641	65,102
	90%Intervalo Conf.		-120,987	91,265	-113,165	99,793	-104,512	108,243	-108,191	105,531	-108,381	104,086	-117,134	96,797
γ ₁	Media	Desv.stand.	-6,529	29,303	-3,535	29,260	-8,148	29,372	0,443	29,405	-1,105	29,436	0,207	29,476
	90%Intervalo Conf.		-54,273	41,603	-51,192	44,722	-56,054	40,324	-47,442	49,000	-48,823	47,755	-47,549	48,853
γ ₂	Media	Desv.stand.	-3,667	22,932	-1,286	22,915	-5,125	22,983	1,466	23,009	-0,105	23,040	0,172	23,079
	90%Intervalo Conf.		-41,074	33,164	-38,647	35,733	-42,604	32,131	-36,472	39,084	-37,897	37,276	-37,649	37,830

Anexo G: Parámetros estimados a nivel de población (Goic y Montgomery, 2011)

Ilustración 4: Parámetros estimados a nivel de población (Goic y Montgomery, 2011)

		Type A	Type B	Type C	Type D	IsLargest	NEmployee
β_1	Mean (s.d)	65.2 (48.2)	164 (61.2)	12.6 (114)	222 (70.9)	0.767 (0.561)	53.2 (45.7)
	90% C.I	(-28.7, 156)	(42.4, 289)	(-204, 239)	(79.4, 361)	(-0.306, 1.82)	(-31.8, 142)
β_2	Mean (s.d)	158 (44.4)	178 (56.5)	108 (106)	133 (63.9)	0.581 (0.506)	-5.46 (42.3)
	90% C.I	(69.3, 248)	(69.3, 296)	(-105, 319)	(7.31, 256)	(-0.448, 1.59)	(-88.4, 81.6)
β_3	Mean (s.d)	42.9 (53.9)	96.7 (67.1)	-9.7 (117)	19.4 (78.4)	2.17 (0.631)	51.7 (52.4)
	90% C.I	(-61.9, 150)	(-38.6, 223)	(-232, 217)	(-129, 168)	(0.912, 3.46)	(-50.1, 155)
β_4	Mean (s.d)	85.8 (30.5)	138 (38.8)	26.5 (72.2)	71.5 (44.4)	0.724 (0.354)	31.4 (29.0)
	90% C.I	(26.9, 145)	(62.7, 216)	(-119, 168)	(-14.4, 157)	(0.001, 1.41)	(-24.8, 86.4)
β_5	Mean (s.d)	101 (29.5)	150 (37.4)	51.6 (70.1)	82 (42.7)	0.555 (0.341)	23.1 (28)
	90% C.I	(44.7, 158)	(75.2, 222)	(-85.4, 188)	(-0.144, 164)	(-0.157, 1.21)	(-33.1, 75.8)
β_6	Mean (s.d)	0.0797 (0.69)	-2.49 (0.85)	0.561 (1.59)	1.06 (1.03)	-0.011 (0.00776)	-0.978 (0.667)
	90% C.I	(-1.26, 1.5)	(-4.11, -0.867)	(-2.7, 3.8)	(-0.913, 3.09)	(-0.0272, 0.0034)	(-2.33, 0.254)
γ_1	Mean (s.d)	-31.9 (37.1)	-2.97 (48.3)	-15.0 (87.3)	-83.2 (55.2)	-1.12 (0.436)	-23.6 (35.4)
	90% C.I	(-106, 41.1)	(-102, 91.5)	(-186, 154)	(-194, 22.5)	(-1.93, -0.231)	(-92.3, 44.5)
γ_2	Mean (s.d)	11.5 (15.1)	11.0 (18.8)	18.3 (33.7)	-27.8 (21.4)	-0.494 (0.169)	-2.44 (14.5)
	90% C.I	(-18.7, 41.7)	(-28.6, 47.1)	(-52.2, 78.6)	(-69.3, 12.0)	(-0.82, -0.152)	(-31.9, 26.2)