



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

VALORACIÓN DE PRODUCTOS DE ARROZ DE VARIEDAD DESARROLLADA EN  
CHILE Y SU EFECTO EN LAS GANANCIAS DEL RETAIL Y EMPRESAS  
PRODUCTORAS

TESIS PARA OPTAR AL GRADO DE  
MAGÍSTER EN ECONOMÍA APLICADA

ÁLVARO RODRIGO LORENZO ESPINOZA HERNÁNDEZ

PROFESOR GUÍA:  
CARLOS NOTON NORAMBUENA

MIEMBROS DE LA COMISIÓN:  
ANDRÉS MUSALEM SAID  
ANDRÉS ELBERG SHEWARD

SANTIAGO DE CHILE  
2018

RESUMEN DE LA MEMORIA PARA OPTAR  
AL TÍTULO DE MAGÍSTER EN ECONOMÍA APLICADA  
POR: ÁLVARO RODRIGO LORENZO ESPINOZA HERNÁNDEZ  
FECHA: 2018  
PROF. GUÍA: CARLOS NOTON NORAMBUENA

VALORACIÓN DE PRODUCTOS DE ARROZ DE VARIEDAD DESARROLLADA EN  
CHILE Y SU EFECTO EN LAS GANANCIAS DEL RETAIL Y EMPRESAS  
PRODUCTORAS

Este trabajo analiza los efectos que las preferencias por características de los productos de arroz en el equilibrio de mercado generan sobre las utilidades de distribuidores y productores, en el contexto de una industria con productos diferenciados. Se utilizan técnicas no paramétricas de estimación estructural de un modelo mixed logit, de desarrollo reciente en la literatura, usando datos transaccionales del retail.

Se obtiene evidencia de heterogeneidad en las preferencias por atributos de los productos, que influye en los patrones de sustitución. Los escenarios analizados por el modelo dan justificación económica a aspectos de organización de esta industria, como el interés de las empresas de mantener el cultivo de variedad local y el ingreso del retail al mercado de este producto en los últimos años, con marcas propias en el segmento que tiene las características más preferidas. Por otra parte, las ganancias generadas plantean un cuestionamiento al uso del precio de variedades de origen importado que las empresas de *national brands* utilizan para valorizar la compra de materia prima a los agricultores locales.



# Tabla de Contenido

<b>Introducción</b>	<b>1</b>
<b>1. Organización de la industria</b>	<b>3</b>
1.1. Producción primaria . . . . .	3
1.2. Industria procesadora . . . . .	4
1.3. Productos . . . . .	5
1.4. Distribución: retail y consumidores . . . . .	6
<b>2. Datos disponibles</b>	<b>9</b>
<b>3. Marco analítico</b>	<b>14</b>
3.1. Demanda . . . . .	14
3.1.1. Preferencias homogéneas: Logit . . . . .	14
3.1.2. Preferencias heterogéneas: Mixed Logit . . . . .	17
3.1.3. Estimación no paramétrica . . . . .	18
3.2. Oferta . . . . .	21
3.2.1. Escenario 1: Base . . . . .	25
3.2.2. Escenario 2: Excluye principales productos de arroz largo ancho de origen nacional . . . . .	25
3.2.3. Escenario 3: Excluye los productos de arroz nacional, largo ancho grado 1, de marcas principales y propias . . . . .	25
3.2.4. Escenario 4: Excluye los productos de arroz nacional, largo ancho grado 1, de marcas principales . . . . .	26
3.2.5. Escenario 5: Excluye los productos de arroz importado, largo delgado grado 2, de marcas principales y propias . . . . .	26
<b>4. Resultados</b>	<b>27</b>
4.1. Estimaciones modelo logit . . . . .	27
4.2. Estimaciones modelo mixed logit . . . . .	30
4.2.1. Distribución de probabilidad de las preferencias . . . . .	30
4.2.2. Elasticidades y patrones de sustitución . . . . .	36
4.2.3. Resultados de escenarios contrafactuales . . . . .	37
<b>Conclusión</b>	<b>43</b>
<b>Bibliografía</b>	<b>45</b>



# Introducción

Este trabajo tiene como objetivo analizar los efectos en las ganancias de la comercialización de productos de arroz en situaciones de equilibrio de oferta y demanda, que se derivan de las preferencias heterogéneas de los consumidores por productos diferenciados del mercado de este producto en Chile.

La industria de arroz a nivel nacional está conformada por dos productores de gran tamaño, y que según reportes de mercado poseen importante identificación y conocimiento en los consumidores [21]. Estas empresas son prácticamente los exclusivos proveedores de una variedad específica de este cereal desarrollada en Chile, que se cultiva en una zona acotada del territorio.

La variedad cultivada en el país presenta características que suponen alta valoración de parte de los consumidores por razones idiosincráticas, como lo destaca [4], y es objetivo testear en este trabajo. Asimismo, las empresas *national brands* también participan activamente en las importaciones de otras variedades de producto y compiten así en todos los segmentos del mercado. El retail en tanto, representa el principal canal de distribución de los productos en esta industria.

En los últimos veinte años, las estimaciones de demanda han servido de fundamento metodológico clave en la literatura de organización industrial enfocada en investigar poder de mercado, fusiones, valoración de nuevos productos o marcas y, en general, en el análisis de la relación vertical entre agentes en industrias con productos diferenciados. Ejemplos de sus aplicaciones de mayor relevancia son los trabajos de [6] y [5] en la industria automotriz, [14] en la valoración de nuevos bienes y su efecto sobre ganancias de bienestar, [22, 24] que examina la competencia en precios y fusiones en la industria de cereales para el desayuno y [9] en la industria de cines. En Chile, una aplicación al análisis de la relación vertical estratégica entre proveedores y el retail basada en una estimación de demanda estructural, se puede encontrar en los trabajos de [25, 26] sobre productos de café.

Con la utilización de técnicas no paramétricas aplicadas a la estimación de demanda de productos diferenciados, de desarrollo reciente en la literatura [12, 13], este trabajo aporta con la aplicación de un método que representa un avance importante en este campo, al no imponer una distribución paramétrica (no testeable) a la heterogeneidad de preferencias.

La investigación posee la ventaja de contar con datos de transacciones individuales efectuadas en el retail, junto con los precios mayoristas pagados al productor, que raramente suelen estar disponibles. La estimación de la demanda permite conocer la importancia relativa de

ciertos atributos y tipos de productos, y analizar escenarios contrafactuales de equilibrio asumiendo un modelo de oferta. En ese contexto, y a partir de la estimación estructural objeto de este trabajo, se pueden iniciar nuevos análisis a las estrategias que el retail adopta para enfrentar la interacción con estos proveedores, incrementando su poder de negociación mediante la entrada de marcas propias y su asociación con alguno de los dos mayores productores en esa industria.

A continuación, en el capítulo 1, se presentan los antecedentes necesarios para comprender la organización de esta industria en Chile. El capítulo 2 describe los datos disponibles para este estudio, y las decisiones relativas a la utilización de esta información en la modelación. El capítulo 3 presenta el marco analítico para la estimación de demanda y el modelo de oferta, junto con la discusión de la literatura para comprender su idoneidad metodológica para los objetivos de la investigación. En el capítulo 4 se discuten los resultados obtenidos y el capítulo final presenta las conclusiones.

# Capítulo 1

## Organización de la industria

### 1.1. Producción primaria

El arroz representa un 20 % de la ingesta calórica a nivel mundial [29], situándose como uno de los principales cereales en la dieta humana. Una gran proporción del arroz producido globalmente corresponde a dos variedades: *índica*, que se caracteriza por tener granos largos y delgados, y *japónica*, de granos más anchos. A estas variedades principales se agregan otras específicas, como el aromático *basmati*, cultivado en Pakistán y el norte de India, o el *javánica*, proveniente de algunas regiones de Indonesia y Madagascar.

La producción está concentrada en Asia, siendo el comercio mundial muy reducido, con un 8 % a 9 % de la disponibilidad total [11]<sup>1</sup>. Al tratarse de una planta de origen tropical, esta se debe adaptar al cultivo durante la estación de verano en los climas templados. Mientras en zonas tropicales y subtropicales se pueden obtener varias cosechas al año, en las templadas solo es posible obtener una, generalmente utilizando sistemas de cultivo irrigado y técnicas de manejo que, sin embargo, permiten obtener un mejor rendimiento por hectárea [28].

El arroz se produce en Chile desde el año 1925, sin embargo, comenzó a tener importancia comercial varias décadas más tarde. El cultivo se establece en las regiones de Maule y Biobío, principalmente en áreas con suelos de tipo arcilloso, que en gran parte se cultivan solo con arroz<sup>2</sup>.

El cultivo en Chile ha contado con el apoyo de un programa de mejoramiento genético desarrollado por el Instituto de Investigaciones Silvoagropecuarias (INIA), que en los años

---

<sup>1</sup>India, Tailandia y Vietnam son líderes en el comercio, con un 26 %, 19 % y 19 % de las exportaciones entre 2012-2014, respectivamente, según cifras reportadas en [11]. Este comercio está constituido por productos de baja calidad y menores precios, que representan excedentes de la producción destinada a los mercados internos de estos países

<sup>2</sup>De acuerdo con [1], los suelos arroceros tienen un alto contenido de arcillas (16 % a 42 % en estrata superficial), bajo contenido de materia orgánica y problemas de drenaje que dificultan el establecimiento de otros cultivos. Como indica [2] respecto a los inicios del rubro en Chile, la introducción de este cultivo permitió que una significativa superficie de suelos considerados marginales – porque no tenían otra alternativa de uso agrícola – pudieran usarse en forma más intensiva y tuvieran una mejor opción económica.



sesenta generó una variedad de arroz de grano corto que en la década siguiente representó casi el 100 % de la superficie sembrada. A raíz de esta experiencia, se fueron introduciendo y evaluando germoplasmas de otras variedades, derivando en los últimos años en variedades de arroz *japónica* con granos de tipo “largo ancho”. De las 8 variedades de arroz lanzadas al mercado por el INIA desde el inicio de este programa, existen 6 variedades aún presentes en la producción nacional [27].

## 1.2. Industria procesadora

En cuanto a la manufactura o elaboración de productos, los procesos industriales son relativamente simples, existiendo innovaciones en sistemas de clasificación de granos de acuerdo a su tamaño, o en el proceso de cocción usado en un tipo particular de producto, denominado arroz parbolizado, que sin embargo no es elaborado por las empresas nacionales.

Los procesos de las empresas de la industria local comprenden las etapas de acopio, secado, elaboración, almacenamiento y distribución. La industria procesadora adquiere directamente las cosechas de arroz con cáscara a los agricultores locales sin la participación de otras empresas o de agentes intermediarios.

El Ministerio de Agricultura de Chile genera indicadores semanales de costos alternativos de importación o de valor de mercado internacional en los principales cereales usados por la industria alimentaria, es decir, existen indicadores para trigo candeal, trigo panadero, maíz y arroz. Este último, creado en 2010, se construye con las cotizaciones que las empresas informan por las importaciones de arroz que posteriormente realizan. Un análisis a la consistencia de este indicador, su capacidad de representar el costo marginal de la industria local y a su relevancia como señal de mercado para la compra de la materia prima a los agricultores nacionales, se presenta en el trabajo de [10].

El análisis allí realizado permite sostener que las empresas que adquieren la producción nacional utilizan la información provista por el indicador para fijar una base para su política de compra. La coordinación se realiza mediante contratos o en transacciones *spot*, y existe una relación de largo plazo entre las empresas y los agricultores. Junto con esto, también se advierten esfuerzos de la industria de mantener la producción a nivel local, mediante la participación en conjunto con INIA en proyectos de investigación y desarrollo, adaptación de nuevas variedades, y favoreciendo el uso de contratos y semillas certificadas.

La industria local se conforma por dos grandes empresas manufactureras (procesadoras), que concentran una alta proporción de ventas en el mercado [21], y que interactúan con una industria del retail donde también existe una presencia reducida de actores. De acuerdo con antecedentes de las empresas citados en [10], el canal retail es la principal vía de distribución de su oferta, con cerca de un 65 al 70 % del total de la producción, siendo la fracción restante destinada al canal *foodservice* (hoteles, restaurantes, casinos) y una menor al canal tradicional (distribuidores de pequeños almacenes y mercados). Asimismo, estas empresas manufactureras nacionales también son las principales importadoras de este producto.

## 1.3. Productos

A diferencia de otros cereales, el arroz se consume principalmente como grano entero después de procesos de elaboración que básicamente consisten en remover total o parcialmente su cáscara y recubrimientos vegetales, lo que permite que el producto primario sea parte importante del costo del producto final [16].

Los productos disponibles en el mercado local abarcan en primer lugar al arroz blanco convencional, en la variedad de grano largo ancho de origen nacional. La característica culinaria principal de estos productos es su gran resistencia a la sobrecocción y, por su alto contenido de almidón, poseen una textura cremosa.

El segundo grupo importante de productos de arroz blanco está constituido por productos de variedades con granos de tipo largo delgado, que tienen su origen en Argentina o Paraguay y en países asiáticos como Tailandia o Vietnam. La preparación de este arroz genera un resultado más “graneado”, porque los granos se endurecen después de la cocción y tienen una textura más sólida y seca.

Junto con estos, que constituyen los segmentos más importantes, existe una menor proporción de productos de arroz parbolizado – también denominado pregraneado o precocido – que es totalmente importado.

De irrupción más reciente en las líneas de producto de cada marca, también se cuenta con productos de arroz integral, con granos que conservan parte de las capas vegetales de alto contenido nutricional, y que es principalmente elaborado con variedades de origen local.

Otros productos elaborados corresponden a los arroces con ingredientes o instantáneos, que incorporan verduras u otros ingredientes adicionales y son de rápida preparación, puesto que se elaboran con arroz pregraneado. Asimismo existen productos de “nicho” en el mercado, como los arroces de variedades de orígenes muy específicos, como el arroz *basmati* o *carnaroli*.

El principal atributo de calidad utilizado en Chile en la industria de elaboración de arroz es el rendimiento de grano entero [8, 15]. Este se define como el porcentaje de granos enteros que se obtiene después del proceso de elaboración, y que además es un parámetro utilizado para valorizar la compra de la producción a los agricultores.

La norma Nch. 1359 of. 2003, que constituye el estándar industrial usado en Chile, clasifica a los productos de arroz elaborado en grado 1, cuando el contenido es de hasta un 5% de granos partidos, y en grado 2, cuando el contenido es de hasta un 20% de granos partidos<sup>3</sup>. La misma norma establece además una clasificación de acuerdo con las dimensiones del grano que, tal como se ha señalado, se explica por su origen y variedad e influye en el resultado culinario obtenido luego de su preparación<sup>4</sup>.

---

<sup>3</sup>Estos estándares son similares a los establecidos en otros países y son acordes a la clasificación de productos transados en el comercio internacional. Por ejemplo, en Estados Unidos, el *Long Grain Milled US N°1* tiene un contenido máximo de 4% de granos partidos y además exige que no incorpore granos de otra variedad; a diferencia del *Long Grain Milled US N°5*, con una tolerancia de 35% de granos partidos y de un 10% de otras variedades [33].

<sup>4</sup>La norma Nch. 1359 of. 2003 establece que el tipo largo ancho es aquel con una relación longitud-ancho

En suma, puede deducirse que los productos de arroz disponibles al consumidor en Chile presentan diferenciación en términos de variedades, orígenes, características, usos culinarios y calidad.

## 1.4. Distribución: retail y consumidores

El consumo per cápita de este cereal en Chile se sitúa entre los 10 a 11 kilos de arroz elaborado al año, con una frecuencia de preparación de tres a cuatro veces por semana. Este nivel de consumo se asemeja al de países de Europa y Estados Unidos – que promedian menos de 10 kilos por persona al año – y se encuentra por debajo de Colombia y Brasil, que registran un consumo anual per cápita de casi 50 kilos [4].

La figura 1.1, construida con estadísticas de fuentes oficiales, refleja que el volumen de producción nacional es siempre menor al de las importaciones, que en la última década han representado cerca del 57 % de la disponibilidad aparente para el consumo interno<sup>5</sup>. Los volúmenes de importaciones reportados en la tabla 1.1, muestran que las empresas adquieren sobre un 80 % del volumen en productos de arroz elaborado, que no requieren procesos industriales adicionales salvo los asociados a la distribución. El porcentaje restante se atribuye a importaciones de arroz partido, que no se destinan al consumo como producto de arroz<sup>6</sup>. Se advierte que la industria local no realiza importaciones de arroz *paddy*, que es el cereal en su estado de materia prima previo al proceso de elaboración, y tampoco adquiere externamente volúmenes relevantes de arroz integral.

Tabla 1.1: Importaciones de arroz según clasificación aduanera de Chile, 2005, 2010, 2014-2016 (toneladas)

Tipo de producto en clasificación aduanera	2005	2010	2014	2015	2016
Arroz con cáscara ( <i>paddy</i> )	33	270	0	0	0
Arroz descascarillado (integral)	16	135	83	165	245
Arroz elaborado <5 % grano partido	16,010	29,713	37,927	33,427	32,469
Arroz elaborado ≥5 % y ≤15 % grano partido	53,008	62,445	50,009	79,502	63,200
Arroz elaborado >15 % grano partido	6,570	6,253	2,240	5,746	8,110
Arroz partido	17,589	25,106	19,488	23,404	25,158
Total	93,226	123,922	109,748	142,245	129,181

Fuente: [10], con datos del Servicio Nacional de Aduanas.

Una desagregación del tipo de productos de arroz elaborado importado, presentada en la figura 1.2, revela que en los últimos años cerca del 65 % del volumen corresponde a productos de arroz blanco que contienen entre un 5 a 15 % de granos partidos, es decir, se consideran en

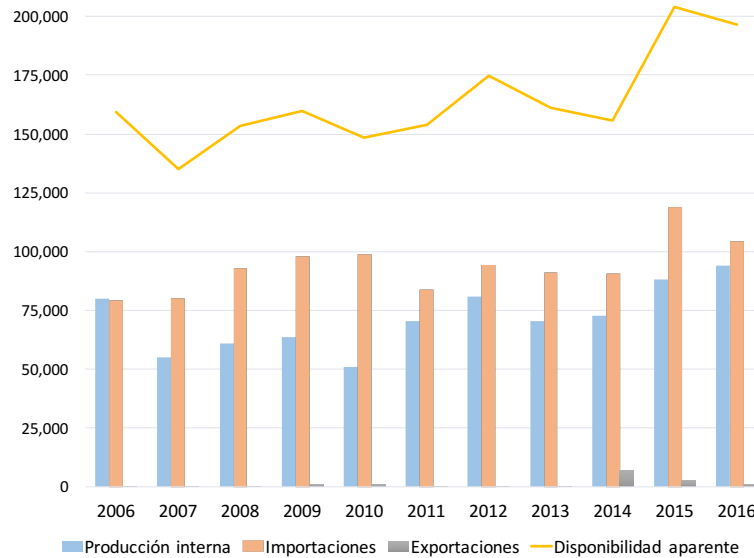
---

menor a 3 y largo delgado (también denominado fino o angosto), el que posee una relación mayor o igual a 3, mientras que los arroces de grano corto presentan una relación menor a 2.

<sup>5</sup>El cálculo de disponibilidad aparente de este producto no considera inventarios, y se obtiene al sumar el volumen de producción interna con las importaciones y restar las exportaciones.

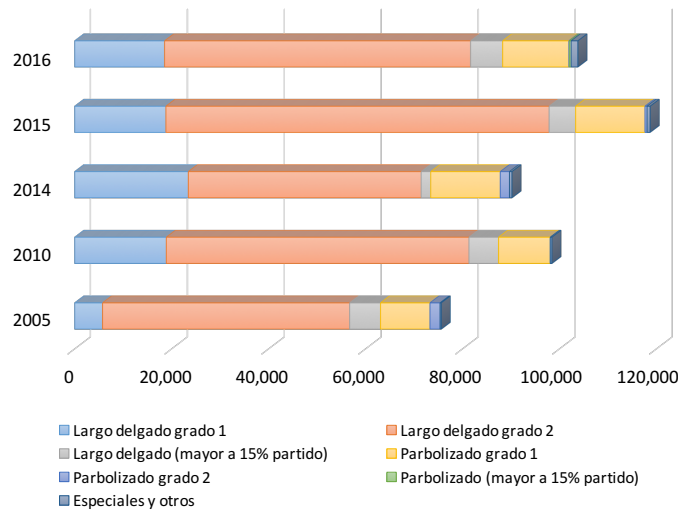
<sup>6</sup>[10] constatan que las importaciones de este tipo de producto son realizadas exclusivamente por empresas de la industria de elaboración de cerveza.

Figura 1.1: Producción interna, importaciones, exportaciones y disponibilidad aparente de arroz en Chile, 2006-2016 (toneladas).



Fuente: [10], con datos de INE, ODEPA y Servicio Nacional de Aduanas.

Figura 1.2: Importaciones de arroz elaborado, según tipo de producto: 2005, 2010 y 2014-2016 (toneladas).



Fuente: [10], con datos del Servicio Nacional de Aduanas.

el estándar de grado 2, mientras que un 19% del volumen de importaciones corresponde al estándar grado 1. Todo este volumen de arroz de origen importado posee granos de tipo largo delgado. Una proporción aproximada de un 15% del volumen en dicho período corresponde

a arroz de tipo parbolizado o pregraneado y a otras variedades más específicas.

El trabajo de [10] advierte que en los últimos años se ha producido la irrupción de empresas dedicadas a la importación que comienzan a ganar volumen en desmedro de las empresas procesadoras, que son las que principalmente han efectuado la importación de este producto en Chile. Los antecedentes indican que las empresas del retail prácticamente no participan en la importación directa de productos, lo que denota que prefieren adquirirlo a las empresas importadoras distribuidoras, o bien a las empresas manufactureras nacionales que también disponen de esos productos en su oferta.

Sin embargo, el retail en los últimos años ha ingresado al mercado mediante productos de marca propia, en los segmentos de variedades de origen nacional e importado. Como muestra el análisis desarrollado por [18], esto ha generado efectos importantes en el *assortment* de productos ofrecidos por el retail, en desmedro de productos de empresas importadoras o de procesadores locales menores.

# Capítulo 2

## Datos disponibles

Los datos provienen de transacciones individuales diarias, realizadas en 64 salas de supermercado distintas de la Región Metropolitana, en un período que abarca desde octubre de 2009 hasta julio de 2010, es decir, 10 meses.

Estos registros pertenecen a dos cadenas de retail diferentes – que en este estudio se denominan QUALITY y VALUE – de propiedad de un mismo holding matriz. Las dos cadenas en conjunto poseen un 28 % del total de participación del mercado retail en Chile en el período respectivo, medido por sus ventas, situándose en un segundo lugar en esta industria según cifras reportadas en [31]<sup>1</sup>.

El tipo de salas tiene características más homogéneas en el retailer QUALITY, con 11 tiendas de gran tamaño – hipermercados – ubicadas en comunas de mayor población y mayor nivel relativo de ingresos en el área urbana del Gran Santiago. En el retailer VALUE, el grupo es más heterogéneo, con 53 tiendas de distintos tamaños, acorde a estrategia y modelo de negocios de esta cadena.

Los registros de compras diarias en cada local se organizan para disponer de datos semanales en base a la combinación o tupla *individuo-producto-tienda-semana*. Esto implica consolidar registros dejando una observación cuando esta combinación aparece más de una vez, lo que ocurre cuando un cliente repite la compra de un mismo ítem de producto en otro día de la misma semana en el mismo supermercado. En estos casos, se considera una sola compra por ese producto y se toma la mediana de su precio<sup>2</sup>.

Los registros de transacciones poseen el número de identificación (RUN) de la persona que adquiere los productos cuando voluntariamente esta entrega esa información en caja al momento de pagar, con el fin de obtener beneficios y/o descuentos del retailer respectivo. Si bien la base de datos también registra transacciones de individuos que no informan su número de identificación (que aproximadamente corresponden a un tercio de las transacciones), en

---

<sup>1</sup>De acuerdo con el cuadro 1 en [31], el índice C3 de la industria de retail en 2010 correspondió a un 80 %.

<sup>2</sup>Se puede notar que si en una misma semana un individuo adquiere productos distintos, o bien compra el mismo producto pero en locales distintos, para los efectos de esta base de datos se considerará como una observación diferente, puesto que en esas situaciones no se produce una repetición de la combinación que define al dato individual en la tupla *individuo-producto-tienda-semana*.

este estudio solo se utilizarán registros de consumidores que dieron a conocer su número de RUN.

Es el número de identificación el que permite reconocer las visitas de los individuos a las tiendas, y hace posible determinar las ocasiones en que estos van al supermercado y no adquieren productos de arroz. Este hecho es considerado en la definición del *outside good*, que para efectos de la modelación corresponderá a la alternativa de no compra de arroz por parte del individuo  $i$ .

Los precios de venta de los productos corresponden a precios observados por el comprador, por lo tanto, ya incluyen todos los descuentos que podrían haberse promocionado por parte del retailer a nivel de cada local. Respecto a cantidades compradas, todas las unidades del producto  $j$  que adquiere un individuo  $i$  en cada local  $n$  en una semana  $w$  se transforman en la variable:

$$y_{ijnw} = 1[q_{ijnw} > 0]$$

Lo anterior se debe a que en este modelo la decisión de compra es binomial, 1 o 0, para múltiples productos, o multinomial. La estructura de la base de datos permite que la cantidad de opciones de producto disponible en cada tienda y semana pueda variar. Cuando un producto no está a la venta en una semana y local cualquiera, la probabilidad de adquirirlo es cero, lo que en el modelo a estimar es considerado en la fórmula del logit multinomial, ya que la utilidad individual y la probabilidad de compra es computada conociendo las alternativas – o choice-set – que estuvo efectivamente disponible en esa *tienda-semana* para el individuo  $i$  (ver capítulo 3).

También para efectos de la modelación de la utilidad individual y la posterior estimación, los otros productos disponibles en el choice-set y no comprados por el individuo  $i$  presentan un precio que corresponde a su valor mediano semanal, es decir, se aplica la misma lógica del precio mediano tanto para productos comprados como para los no comprados por  $i$ .

El número de productos disponibles en ambos retailers es de aproximadamente 200 ítems distintos. Al tratarse de información transaccional, estos se encuentran identificados con un número SKU único y la base de datos incluye un campo con la descripción de su marca y características. Debido a la gran cantidad de tipos de producto en la oferta, se generan *ranking* de ventas de los productos más importantes en ambos retailers. El espacio de productos se fija en los 33 principales ítems en términos de ingreso por ventas para cada cadena de retail. En el retailer QUALITY, estos 33 principales SKU representan un 74% de los ingresos por ventas de arroz, mientras que en VALUE representan el 85%. El resto de SKU's se agrupan como un producto #34 en el choice-set ofrecido por cada retailer.

En los productos principales se cuenta con ítems de arroz largo ancho grado 1 y grado 2, de origen nacional, arroces importados de tipo largo delgado grado 1 y grado 2, más algunos parbolizados e integrales. Dado que el criterio de inclusión en el *inside good* corresponde al de ventas en ambas cadenas de retail, estos productos pertenecen a las principales empresas (marcas) del mercado, como Carozzi (Miraflores, Rizzo) y Tucapel (Banquete), además de las marcas propias del retailer y también otros de productores menores. El resto de productos que se agrupan como opción #34, considera principalmente a los arroces instantáneos con ingredientes y variedades especiales, con menores ventas y participaciones de mercado.

Tabla 2.1: Participaciones de mercado para grupos de productos, sobre el total de transacciones o el total de compras de arroz en cada retailer

Marca y tipo de producto	Total de transacciones		Total de compras arroz	
	QUALITY	VALUE	QUALITY	VALUE
Carozzi ancho grado 1	3.1 %	1.1 %	19.1 %	11.7 %
Tucapel ancho grado 1	1.8 %	0.7 %	11.0 %	7.9 %
M. Propia ancho grado 1	0.7 %	0.3 %	4.2 %	2.6 %
Carozzi ancho grado 2	0.6 %	0.5 %	3.4 %	4.8 %
Tucapel ancho grado 2	1.8 %	1.3 %	10.8 %	13.7 %
M. Propia ancho grado 2	0.3 %	0.2 %	1.7 %	2.0 %
Carozzi delgado grado 2	0.2 %	0.3 %	1.2 %	2.7 %
Tucapel delgado grado 2	0.5 %	0.6 %	2.9 %	6.3 %
M. Propia delgado grado 2	n/d	0.2 %	n/d	2.0 %
Carozzi otro tipo	0.4 %	0.1 %	2.1 %	1.1 %
Tucapel otro tipo	0.5 %	0.2 %	2.8 %	2.1 %
M. Propia otro tipo	0.2 %	0.1 %	1.3 %	1.0 %
Otras marcas	1.6 %	2.3 %	9.6 %	24.5 %
Resto de otro tipo	4.9 %	1.7 %	29.8 %	17.5 %
Todos los productos	16.5 %	9.5 %	100 %	100 %

*Nota:* Muestra de  $N=100,000$  transacciones de cada retailer. n/d: No disponible en la oferta.

Las participaciones de mercado se presentan en la tabla 2.1, tanto para una definición que consiste en el share respecto al total de transacciones de una muestra aleatoria, como para aquella obtenida al considerar solo las compras de arroz. Por simplicidad se presentan precios de grupos de productos según productor (o marca) y tipo. Al sumar shares de grupos de productos se aprecia que Carozzi y Tucapel son los fabricantes más relevantes, con participaciones de 26 % y 28 % respectivamente en QUALITY y de 20 % y 30 % en VALUE, al medirlas sobre las compras de arroz. Las marcas propias, en tanto, representan un 7 % de preferencias en QUALITY y 8 % en VALUE, respectivamente.

En cuanto a la participación de otras marcas, la tabla 2.1 revela que en VALUE esta se sitúa en torno al 25 % del total de compras de arroz, mientras que en QUALITY estas suman un 10 %. Es en este retailer en donde los productos de otro tipo de variedades especiales, arroces instantáneos, parbolizados de otros fabricantes y el resto de productos tiene una mayor preferencia, con casi un 30 % del total, lo que refleja las diferencias en las alternativas y tipos de producto ofrecidos. Con todo, es importante destacar que en una muestra aleatoria de 100,000 transacciones en cada retailer, en QUALITY un 16.5 % de estas registran compras de arroz, mientras en VALUE esta cifra solo alcanza a un 9.5 % de las transacciones.

La tabla 2.2 presenta estadísticas descriptivas de los precios de los productos de arroz que venden ambos retailers. El coeficiente de variación de los precios de venta denota que estos no varían sustantivamente en su oferta semanal. Sin embargo, al aplicar un contraste a las medias de precio, se observan diferencias significativas en la gran mayoría de productos. Los productos de origen nacional y mejor estándar (ancho grado 1) tienen precios más elevados en QUALITY con excepción de aquellos de marca propia.



Tabla 2.2: Estadísticas descriptivas de precios de productos de arroz en ambos retailers

Marca y tipo de producto	p10 (\$)	p50 (\$)	p90 (\$)	Desv. Estd. (\$)	Coef. Var. (%)	Precio medio QUALITY (\$)	Precio medio VALUE (\$)	Diferencia Q-V (%)	Test t
Carozzi ag1	769	849	899	68	7.9	850	845	0.59	**
Tucapel ag1	809	899	999	96	10.6	907	868	4.43	***
M. Propia ag1	695	799	890	73	9.2	785	800	-1.85	**
Carozzi ag2	579	779	899	102	13.2	771	761	1.31	
Tucapel ag2	633	749	929	115	14.8	765	775	-1.26	**
M. Propia ag2	665	749	799	55	7.5	725	746	-2.79	***
Carozzi dg2	679	799	799	72	9.7	784	705	10.68	***
Tucapel dg2	543	579	599	57	9.7	570	595	-4.20	***
M. Propia dg2	499	499	539	17	3.5	n/d	503	n/d	
Carozzi otro	867	909	1,139	99	10.5	928	981	-5.62	***
Tucapel otro	949	1,095	1,119	100	9.4	1,085	1,001	8.03	***
M. Propia otro	860	909	1,009	76	8.1	918	976	-6.06	***
Otras marcas	499	599	879	386	53.7	713	644	10.19	***
Resto otro tipo	679	999	2,851	1,210	80.6	1,221	1,238	-1.38	
Todos	589	835	1,199	705	72.2	915	804	12.92	***

*Nota:* Muestra de  $N=100,000$  transacciones de cada retailer. En última columna se presenta el nivel de significancia del contraste de diferencia de medias ( $H_0 = 0$ ) para muestras de distinta varianza, con  $\Pr(|T| > |t|)$  menor al 1% (\*\*\*), 5% (\*\*) y 10% (\*). n/d: No disponible en la oferta.

Sorprendentemente, en los productos de origen nacional y menor estándar (ancho grado 2), el contraste entre retailers posiciona a Tucapel y a la marca propia con precios mayores en VALUE respecto a QUALITY, sin diferencias significativas en productos de Carozzi. En los productos de origen importado, de tipo largo delgado grado 2, QUALITY tiene una oferta a un precio un 10% más alto en marcas de Carozzi, mientras que los de Tucapel se encuentran un 4% más baratos que en VALUE. Productos de otras marcas tienen un precio un 10% mayor en QUALITY respecto a VALUE, mientras que el grupo que representa el resto de variedades especiales y otros tipos no tienen un precio medio estadísticamente distinto entre retailers.

Las diferencias de precios observadas en las muestras de transacciones de ambos retailers, y las distintas participaciones de mercado que se registran para productos de diferente variedad o tipo, estándar de calidad y fabricante, dan cuenta de una heterogeneidad de preferencias que requiere ser explorada en el análisis de la demanda de los productos, cuestión de la que se ocupa el capítulo 3 de esta tesis.

Por último, es necesario indicar que la modelación de estos datos asume que no existe un peso muy significativo de los productos de arroz en el presupuesto para alimentación, que induzca a que los consumidores requieran cambiarse a otra tienda o retailer cuando el producto buscado no está disponible. Esto se debe a que los datos no capturan situaciones de cambio de local explicadas por escenarios donde se alteren las marcas y tipos de productos disponibles para el consumidor en una *tienda-semana* específica. Para sustentar este supuesto, la ponderación que tiene el arroz en la canasta de bienes y servicios del índice de precios del consumidor (IPC) es de 0.2%, que corresponde a un 1.2% del grupo de alimentos, de acuerdo a lo establecido en la metodología del INE para el IPC<sup>3</sup>.

<sup>3</sup>Este valor es distante del pan, con un 2.1%, que representa un 12.3% del total de la canasta de alimentos, o de la carne de vacuno, que posee una ponderación de 1.7% y equivale al 10.1% del grupo de alimentos en

Además de lo anterior, se debe señalar que la forma de modelar la demanda a partir de estos datos individuales de transacciones asume que esta será estática, es decir, sin acumulación de inventarios o compras con razones intertemporales. Esto implica que cada transacción se considera independiente a través del tiempo. Este supuesto se puede sostener en la constatación de que los precios en los principales productos no presentan coeficientes de variación muy elevados (por lo general menor a 10 % y como máximo un 15 %, ver tabla 2.2), lo que refleja que los retailers no tienen períodos de promociones de precio de gran magnitud o duración. La importancia y frecuencia de consumo de este producto (mencionada anteriormente), y el formato y tamaño de envase comúnmente usado en Chile (1 kg), permiten argumentar que en este mercado las compras son acordes con una demanda de características estáticas.

---

el IPC.

# Capítulo 3

## Marco analítico

### 3.1. Demanda

#### 3.1.1. Preferencias homogéneas: Logit

En este capítulo, se describe el modelo de utilidad lineal y el enfoque econométrico usado para estimar los parámetros de “gustos” por las características que presentan los productos, como una primera aproximación basada en preferencias homogéneas, o que describen a un único agente representativo.

En los datos a disposición de este estudio se observan registros de transacciones que fueron individualizados en la combinación *individuo-producto-tienda-semana* para cada cadena de retail por separado.

Se asume entonces un modelo logit condicional en alternativas específicas, que tiene sus fundamentos en el trabajo pionero de [19], donde la utilidad indirecta de individuos  $i = \{1, \dots, I\}$  por productos  $j = \{1, \dots, J\}$  en la tienda  $n = \{1, \dots, N\}$  en la semana  $w = \{1, \dots, W\}$  está dada por:

$$u_{ijnw} = \alpha(y_i - p_{jnw}) + \beta'x_j + \xi_{jnw} + \varepsilon_{ijnw} \quad (3.1)$$

donde  $y_i$  es el ingreso del consumidor,  $p_{jnw}$  es el precio del producto  $j$  y  $x_j$  son dummies que identifican a su marca y características, que en este setting específico no varían entre semana o tienda<sup>1</sup>. El término  $\xi_{jnw}$  agrupa a factores no observables para el econometrista (pero sí percibidos por el consumidor), que pueden afectar la demanda del producto  $j$  y son específicos a la tienda  $n$  y semana  $w$ . Como es usual en esta literatura, se asume que el término aleatorio  $\varepsilon_{ijnw}$  es iid y posee una función de distribución Type I extreme value.

En cuanto a los parámetros del modelo,  $\alpha$  captura la utilidad marginal del ingreso, por

---

<sup>1</sup>Como se indicó en el capítulo 2, los productos en cada semana y tienda pueden ser distintos, sin embargo, las dummies que denotan sus características y marcas permanecen invariantes y son específicas para cada  $j$ .

lo tanto, refleja cuanta desutilidad se explica por un precio que disminuye el ingreso disponible, mientras  $\beta$  agrupa a los coeficientes de “gustos” por las características y marcas de los productos. Se define entonces a  $\theta = (\alpha, \beta)$  el vector que contiene todos los parámetros, que como se ha señalado, no varían entre individuos.

De acuerdo a lo que discute [23], la especificación dada por la ecuación (3.1) deja implícito algunos elementos. Primero, esta forma de utilidad indirecta puede ser derivada de una función de utilidad cuasilineal, que está libre de efecto-ingreso. Para algunos productos este supuesto es razonable, como ejemplifica este autor con los cereales para el desayuno, y que puede extenderse a los productos de arroz; mientras que para otros, como los automóviles, puede no serlo<sup>2</sup>.

Un segundo aspecto, es que la ecuación (3.1) modela a individuos que observan precios específicos a la *tienda-semana* por ese producto, lo que a diferencia de un modelo con datos agregados, previene del problema de sesgo por error de medida que podría generarse si en cambio se usara un precio de lista o un promedio. Sin embargo, y aun cuando con datos individuales se tiene esta ventaja, las estimaciones de demanda deben hacerse cargo del potencial problema de endogeneidad que podrían seguir presentando los precios observados, en caso de existir una correlación con los factores no observables contenidos en  $\xi_{jnw}$ . En el setting definido, para que esto ocurra el retailer en cada *local – semana* tendría que fijar precios basado en esos factores no capturados por el econométrista en las características y marcas de los productos. Es en esa situación donde se podría presentar sesgo en el parámetro del precio, y aunque por lo señalado anteriormente este sería menor que si se emplearan datos agregados, resulta plausible suponer que las posibilidades de variación que aportan a este modelo los datos individuales previenen de que ese efecto sea sistemático.

Se denota ahora a  $y_{inw} = \{1, \dots, J\}$  cuando el individuo  $i$  elige el producto  $j$  en la tienda  $n$  y semana  $w$ . Así, la probabilidad de adquirir ese producto  $j$  será la integral sobre los shocks  $\varepsilon$  que aseguran que dicho producto  $j$  maximiza la utilidad dado el choice-set de productos disponibles en ese mercado. Usando la distribución de  $\varepsilon$ , se obtiene una solución de forma cerrada para la probabilidad individual  $s_{ijnw}$  de adquirir el producto  $j$ . Formalmente, esto equivale a:

$$s_{ijnw}(\theta) \equiv \mathbb{P}(y_{inw} = j \mid \theta) = \mathbb{P}(u_{ijnw} > u_{iknw}, \forall k \neq j) = \frac{\exp(-\alpha p_{jnw} + \beta' x_j + \xi_{jnw})}{\sum_{h=1}^J \exp(-\alpha p_{hnw} + \beta' x_h + \xi_{hnw})} \quad (3.2)$$

La alternativa del *outside good* se denota  $j_0$ , y como ya se indicó representa la opción de no adquirir los productos de arroz en la transacción efectuada por el individuo  $i$ . La utilidad dada por la ecuación (3.1) para la opción del *outside good* se normaliza en su término aleatorio, y entonces se tiene que para esta alternativa su utilidad corresponde a  $u_{ij_0nw} = \varepsilon_{ij_0nw}$ , lo que implica que  $-\alpha p_{j_0nw} + \beta' x_{j_0} + \xi_{j_0nw} = 0$ <sup>3</sup>.

---

<sup>2</sup>Es por esta razón que el modelo desarrollado en [5] para estimar demanda de productos diferenciados en la industria automotriz, se construye sobre una función de utilidad Cobb-Douglas, que deriva una utilidad indirecta en función de  $\log(y_i - p_{jt})$ .

<sup>3</sup>Con esta normalización para la opción del *outside good*, y considerando que el individuo enfrenta un

La normalización de las utilidades individuales es válida, puesto que como se desprende de la ecuación (3.2), la probabilidad de compra resulta de una comparación de utilidad de distintos productos para un mismo consumidor, y entonces lo que importa es el orden de utilidad que determina su preferencia. Por esta misma razón, variables explicativas que no son específicas a los productos, como es el caso del ingreso del individuo dado por  $y_i$ , o incluso efectos-fijo para cada tienda  $n$  en caso de haberse definido, se restan en la comparación y no tienen influencia en la probabilidad individual que decide la compra.

Se estima el modelo usando información detallada de una muestra de transacciones, limitada a 100,000 observaciones en cada retailer por razones computacionales. La estimación es realizada vía máxima verosimilitud (MLE), encontrando los parámetros  $\theta$  que resuelven:

$$\hat{\theta}_{MLE} = \arg \max_{\theta \in \Theta} \ln L(\theta) = \arg \max_{\theta \in \Theta} \sum_{i=1}^I \sum_{j=1}^J \sum_{n=1}^N \sum_{w=1}^W y_{ijnw} \ln(s_{ijnw}(\theta)) \quad (3.3)$$

donde la dummy  $y_{ijnw}$  es uno si el individuo  $i$  escoge el producto  $j$  en la tienda  $n$  en la semana  $w$ , y cero en otro caso.

La popularidad que exhibe el modelo logit se debe fundamentalmente a que resuelve el problema de dimensionalidad que presentaría la estimación de un sistema de ecuaciones de demanda para productos diferenciados, con tantos parámetros a estimar como el cuadrado del número de productos<sup>4</sup>. Al proyectar los productos a un espacio de características y atributos valorados en la utilidad, es el tamaño de este espacio el que determina los parámetros de interés. Si bien los supuestos permiten obtener una forma cerrada para la integral de las probabilidades de compra en la ecuación (3.2), su simpleza tiene costos en términos de las capacidades analíticas que posee.

Debido a la forma como es restringida la heterogeneidad, los patrones de sustitución entre los productos se explican por sus shares de mercado y no por cuán similares sean. Esta poco atractiva propiedad del modelo logit es denominada Independencia de Alternativas Irrelevantes (IIA) y se explica porque los shocks son iid no solo entre individuos y/o mercados, sino también entre productos, lo que repercute en predicciones sobre la sustitución muy cuestionables<sup>5</sup>.

Conscientes de este problema, desarrollos metodológicos a partir del trabajo de [20] con los modelos extreme-value generalizado (GEV) y nested logit, inducen una mayor similitud entre las opciones disponibles, permitiendo que  $\varepsilon$  esté correlacionado entre productos y no independientemente distribuido. Sin embargo, la definición a priori de grupos, y que entre

---

choice-set que depende de la disponibilidad de productos en la *tienda-semana* respectiva, que se denota  $a_{jnw}$ , la expresión para la probabilidad individual de compra a utilizar es:

$$s_{ijnw} = \frac{a_{jnw} \exp(-\alpha p_{jnw} + \beta' x_j + \xi_{jnw})}{1 + \sum_{h=1}^J a_{hnw} \exp(-\alpha p_{hnw} + \beta' x_h + \xi_{hnw})}, \forall h \neq j_0.$$

<sup>4</sup>Un mercado con 20 productos implicaría estimar al menos 400 parámetros, repartidos en 20 ecuaciones, una para cada producto, con 20 precios en cada una.

<sup>5</sup>Para ver esto, se puede notar que al derivar los shares de mercado generados por la ecuación (3.2) respecto al precio se obtiene  $\partial s_j / \partial p_j = -\alpha s_j (1 - s_j)$  y  $\partial s_j / \partial p_k = \alpha s_j s_k$ . Esto implica que la elasticidad precio propio de un producto  $j$  estará dada por la expresión  $\eta_j = -\alpha p_j (1 - s_j)$ , creciente en el precio. Esto no es realista teniendo en cuenta que quienes adquieren productos más costosos podrían ser menos sensibles al precio. Por otro lado, la elasticidad precio cruzada entre dos productos  $j$  y  $k$  estará dada por  $\eta_{jk} = \alpha p_k s_k$ , que depende exclusivamente del precio y share del producto  $k$  y no de la similitud que pueda tener con el producto  $j$ .

estos no se alteren los patrones de sustitución, hace que el problema de IIA siga presente. Adicionalmente, como puntualiza [23], este modelo no resuelve el problema de que la elasticidad precio propio siga dependiendo de la forma funcional.

Es a partir de la contribución de [6] y principalmente [5] que se logra desarrollar un modelo más general de la utilidad indirecta, que “libera” el espacio de preferencias a distribuciones de parámetros específicos para cada individuo, que incluso pueden depender de características demográficas, y logra así generar la correlación entre productos similares al ser estos preferidos por individuos con similares gustos.

Este modelo, que en su aplicación específica a la presente tesis se describe en la siguiente sección, ha recibido mucha atención en la literatura de organización industrial, como fue mencionado en la Introducción. Posee varias denominaciones, siendo las más comunes modelo BLP, discrete-choice random coefficients o mixed logit.

### 3.1.2. Preferencias heterogéneas: Mixed Logit

Como se ha señalado anteriormente, la necesidad de disponer de una mayor flexibilidad para capturar en forma realista los patrones de sustitución entre productos, convierten al modelo mixed logit en apropiado a este fin. Al tratarse de un modelo estructural, además, logra capturar preferencias que permiten simular situaciones hipotéticas del mercado, con entrada o salida de productos. Estos escenarios contrafactuales, se analizarán en detalle con un modelo de oferta descrito en la sección siguiente, que como condición exige estimar previamente la demanda. Entre otras cosas, por ejemplo, hace factible conocer el efecto en la demanda cuando desaparece un proveedor o un tipo particular de productos, algo que no es observado en los datos del período bajo análisis, y que por lo tanto no sería posible racionalizar con estimaciones de forma reducida.

Se tiene entonces que este modelo se define en base a la siguiente expresión para la utilidad individual, que considera para el individuo  $i$ , producto  $j$ , tienda  $n$  y semana  $w$  un valor dado por:

$$u_{ijnw} = \alpha_i(y_i - p_{jnw}) + \beta_i'x_j + \xi_{jnw} + \varepsilon_{ijnw} \quad (3.4)$$

A diferencia del modelo logit tradicional, aquí los parámetros de sensibilidad al precio dados por  $\alpha_i$ , son específicos para cada individuo, y en forma análoga los parámetros asociados a gustos por características o marcas de los productos, contenidos en  $\beta_i$ .

Estos parámetros pueden descomponerse en una parte constante (p.ej. su promedio) y otra aleatoria, en la forma:

$$\begin{pmatrix} \alpha_i \\ \beta_i \end{pmatrix} = \begin{pmatrix} \bar{\alpha} \\ \bar{\beta} \end{pmatrix} + \Gamma D_i + \Sigma v_i \quad , \quad (3.5)$$

donde  $D_i$  es un vector de variables demográficas y  $v_i$  un término que captura la heterogeneidad no observable en la sensibilidad al precio y los gustos de los individuos. Las variables demográficas podrían tener una distribución conocida y la discusión principal de este modelo no se centra en ellas<sup>6</sup>. Pero en las variables no observables esta distribución es claramente desconocida, y el modelo mixed logit requiere establecer una forma paramétrica sobre ella o estimarla para encontrar los shares de mercado de cada producto.

En concreto, definiendo a  $\theta = (\alpha, \beta, \Gamma, \Sigma)$  el vector que contiene todos los parámetros y omitiendo los subíndices de tienda y semana por simplicidad,  $s_j$  se obtiene integrando sobre la masa de individuos que escogen el producto  $j$  ( $A_j$ ), y que depende de variables aleatorias  $\varepsilon_{ij} = (\varepsilon_{i0}, \dots, \varepsilon_{iJ})$  y un shock individual  $v_i$ . Formalmente,

$$s_j(x, p, \xi; \theta) = \int_{A_j} dF_\varepsilon(\varepsilon|v_i)d\Phi(v_i) = \int_{A_j} s_{ij}d\Phi(v_i) \quad (3.6)$$

Al tratarse de un modelo logit,  $\varepsilon$  es iid con una distribución Type I extreme value, y entonces el término  $s_{ij}$  tendrá una forma cerrada similar a a la ecuación (3.2), con la salvedad de que aquí contiene parámetros individuales<sup>7</sup>. Como explica detalladamente [23], el modelo propuesto por [5] asume que la distribución  $\Phi(v_i)$  es normal multivariada estándar, y es la matriz  $\Sigma$  la que permite que cada componente  $v_i$  influya con diferente varianza en los parámetros individuales y provoque correlación en las características preferidas por  $i$ , necesaria para encontrar patrones de sustitución que respondan a la similitud de los productos.

[6] y [5] proponen un método iterativo para resolver la integral definida en (3.6) mediante simulaciones. Con esto obtienen shares de mercado que se contrastan con los datos observados, y así encuentran los parámetros  $\theta$  que minimizan el término  $\xi_j$ , que en este modelo opera como residuo. Asimismo, proponen instrumentos para resolver la endogeneidad que pueda presentar el precio.

En esta tesis en cambio, se adoptará un enfoque no paramétrico de desarrollo reciente en la literatura, que a diferencia del modelo BLP no establece una forma a la distribución de los shocks  $v_i$  que determinan la heterogeneidad de los parámetros individuales.

### 3.1.3. Estimación no paramétrica

Esta sección se basa en los trabajos de [3], [12] y [13], quienes desarrollan un estimador no paramétrico de mixtures de distribuciones conjuntas, aplicado a parámetros heterogéneos en modelos estructurales. Anteriormente, [17] había presentado el uso de una grilla discreta para la obtención de histogramas de “puntos ideales” para un espacio predefinido, aunque en

<sup>6</sup>Como señala [23], esta puede ser una distribución no paramétrica conocida desde una fuente de datos demográficos, o una distribución paramétrica estimada en un modelo aparte.

<sup>7</sup>Se puede notar que al reemplazar  $(\alpha_i, \beta_i)$  en términos de los parámetros comunes  $(\alpha, \beta, \Gamma, \Sigma)$  y asumiendo por simplicidad que  $\Gamma = 0$  se tendrá que:  $-\alpha_i p_j + \beta'_i x_j + \xi_j = -\alpha p_j + \beta' x_j + \xi_j + [-p_j, x_j] \Sigma v_i$ . En este modelo, esta ecuación con términos que en el caso general pueden ser no lineales, es la que corresponde incluir en la forma cerrada para la probabilidad individual  $s_{ij}$  descrita por la ecuación (3.2). Se puede advertir que esto complica la obtención de un valor analítico para la integral definida en la ecuación (3.6).

el contexto de preferencias asociadas a atributos psicofísicos en campos distintos a economía. La ventaja del enfoque, asegura, es tener una mayor flexibilidad para encontrar distribuciones que no necesariamente sean unimodales.

Para simplicidad en la notación se asume que los parámetros que contienen la heterogeneidad individual  $(\alpha_i, \beta_i)$  estarán contenidos en un único vector  $\beta$ . Con esto, el objetivo será estimar la distribución de estos parámetros heterogéneos,  $F(\beta)$ , en un modelo del tipo:

$$P_j(x) = \int g_j(x, \beta) dF(\beta) \quad (3.7)$$

Se puede notar la similitud entre esta expresión y la ecuación (3.6). En este contexto,  $j$  denota el índice de  $J$  valores finitos que puede tomar un resultado,  $x$  es un vector de variables explicativas observadas, y  $g_j(x, \beta)$  es la probabilidad de que el resultado  $j$ -ésimo ocurra para una observación con parámetros heterogéneos  $\beta$  y variables  $x$ . Con esta estructura,  $P_j$  es una probabilidad de observar ese resultado  $j$  en un corte transversal cualquiera cuando las variables explicativas son  $x$ .

En el modelo logit, el resultado corresponde a  $y_{ij}$ , que es igual a uno cuando  $y_i = j$  y cero en otro caso, lo que denota la elección observada en  $i$ . Agregando la variable  $y_{ij}$  en ambos lados de la ecuación (3.7), y moviendo  $P_j$  al lado derecho, para la observación  $i$  se tendrá la siguiente expresión:

$$y_{ij} = \int g_j(x, \beta) dF(\beta) + (y_{ij} - P_j(x)) \quad (3.8)$$

[13] argumentan que este es un modelo de probabilidad lineal que tiene en  $F(\beta)$  un “parámetro” de dimensión infinita. Incluso, indican que sería posible trabajar directamente con esta ecuación si fuera computacionalmente simple estimar este parámetro de dimensión infinita, restringiéndolo a que sea una función de distribución acumulada (CDF) válida. En cambio, proponen una estrategia empírica basada en un espacio de dimensión finita para hacer una aproximación a  $F$ , explotando la forma lineal como entra en la ecuación (3.8).

En particular, definen a  $R(N)$  el número total de puntos en todas las dimensiones, en una grilla  $\mathcal{B}_{R(N)} = (\beta^1, \dots, \beta^{R(N)})$ , donde cada punto es un vector dado que  $\beta$  es un vector. El economista debe escoger los puntos en  $\mathcal{B}_{R(N)}$  y, dada esa elección, se estima  $\theta = (\theta^1, \dots, \theta^{R(N)})$ , que equivalen a los “pesos” sobre cada uno de los puntos correspondientes de la grilla elegida. Con esta aproximación, la ecuación (3.8) se convierte en:

$$y_{ij} \approx \sum_{r=1}^{R(N)} \theta^r g_j(x_i, \beta^r) + (y_{ij} - P_j(x)) \quad (3.9)$$

Como cada  $\theta^r$  entra linealmente en  $y_{ij}$ , la estimación de  $(\theta^1, \dots, \theta^{R(N)})$  se realiza mediante una regresión del modelo de probabilidad lineal de  $y_{ij}$  en los  $R$  “regresores”  $z_{ij}^r = g_j(x_i, \beta^r)$ . Dado que se requiere que la distribución sea una CDF válida, se establecen las restricciones



$\theta^r \geq 0 \forall r = 1, \dots, R(N)$  y  $\sum_{r=1}^{R(N)} \theta^r = 1$ . Sujeto a estas restricciones, el estimador propuesto se define:

$$\hat{\theta}_{FKRB} = \arg \min_{\theta} \frac{1}{NJ} \sum_{i=1}^N \sum_{j=1}^J \left( y_{ij} - \sum_{r=1}^{R(N)} \theta^r z_{ij}^r \right)^2 \quad (3.10)$$

Se puede advertir que existirán  $J$  “observaciones de la regresión” del modelo de probabilidad lineal para cada observación estadística  $(y_i, x_i)$ . La minimización propuesta por el estimador  $FKRB$  se puede implementar como un problema de minimización cuadrática sujeto a restricciones de desigualdad<sup>8</sup>.

Luego de estimar los parámetros  $\theta$ , se puede construir la función de distribución acumulada de los parámetros heterogéneos  $\beta$  en base a:

$$\hat{F}_N(\beta) = \sum_{r=1}^{R(N)} \hat{\theta}^r \mathbf{1}[\beta^r \leq \beta] \quad (3.11)$$

Dado que se trata de un enfoque no paramétrico, siempre existirá la desventaja de elegir “puntos de partida” en la sintonización de los parámetros. [13] señalan que en este caso, la elección de una grilla de puntos es una calibración de dimensión muy amplia, a diferencia de otros enfoques no paramétricos.

[12] discuten sobre el número de puntos en la grilla, el rango de soporte de los puntos, y sus valores. En particular para este trabajo se adoptará lo allí sugerido en relación a que los puntos elegidos corresponderán a valores de un múltiplo de los intervalos de confianza de los parámetros encontrados en las estimaciones del modelo logit tradicional. Asimismo, el número  $R$  estará acotado por razones computacionales a un máximo de 200, dado que se consideran hasta 100,000 observaciones para la estimación, y a 6 variables explicativas además del precio<sup>9</sup>. [12] presentan resultados de ejercicios de Montecarlo que demuestran excelentes resultados en términos del ajuste de los estimadores en muestras pequeñas, y además confirman que este estimador tiene una ventaja importante en cuanto a tiempos de procesamiento<sup>10</sup>.

---

<sup>8</sup>En este trabajo, se implementa con la rutina “lsqlin” de Matlab. [13] también definen un estimador basado en el criterio de máxima verosimilitud siguiendo a [32]. Asimismo, verifican las condiciones de convergencia a la verdadera distribución para el modelo logit y para otras aplicaciones.

<sup>9</sup>Como explica [12], el estimador, como cualquier otro, posee un trade-off sesgo-varianza: elegir un  $R$  más alto hace que la grilla sea más flexible y por ende reduce el sesgo, pero implica un mayor número de parámetros a estimar e incrementa la varianza. En este trabajo se presentan estimaciones con distintos tamaños para  $R$ , y también distintos rangos de soporte de los puntos de la grilla.

<sup>10</sup>En estos ejercicios [12] comparan el estimador con una distribución normal bivariada, y consideran un número de  $R \leq 100$  y  $N \leq 10,000$ , y entre 2 a 6 mixtures de distribuciones de parámetros para una estimación de demanda de  $J = 10$  productos.

## 3.2. Oferta

Con el objetivo de modelar precios de equilibrio de mercado y estimar ganancias por la venta de los productos en distintos escenarios contrafactuales, se estima un modelo de oferta. Este se desarrolla a partir del marco analítico propuesto en los trabajos de [22, 24], [30] y [34], y supone la existencia de firmas – distribuidoras y productoras – que maximizan utilidades y compiten en un mercado de productos diferenciados, con precios de equilibrio que se derivan de estrategias acordes al paradigma Nash-Bertrand.

Estos artículos son ampliamente citados en la literatura de análisis de relaciones verticales, específicamente en la determinación del valor que se genera como superávit de distribuidores y productores. En base a la estimación de modelos estructurales de demanda, proponen métodos para identificar los márgenes precio-costo del contrato vertical que mejor se ajuste a los datos, asumiendo implícitamente un modelo de conducta competitiva de los agentes y sin negociación bilateral.

A diferencia de lo que comúnmente utiliza esa literatura, este trabajo cuenta con la ventaja de disponer de precios mayoristas efectivos para cada producto. Por lo tanto, se conoce la parte más relevante del costo marginal para el retailer: el precio mayorista negociado con el productor por cada ítem. Esto implica que no se requiere establecer supuestos ni estimaciones adicionales sobre los componentes del costo marginal, para contrastar con la predicción de las estimaciones del modelo estructural.

Asimismo, en el caso de las empresas productoras, el costo marginal de producción de arroz que se utilizará para calcular las ganancias de estos agentes se basará en el indicador de costo de internación de arroz que el Ministerio de Agricultura de Chile publica semanalmente para efectos de transparentar la comercialización interna de este cereal. Este índice está basado en el costo de las importaciones de este producto, por ende, en el precio de mercados internacionales, y considera coeficientes técnicos de transformación del cereal desde su estado de materia prima a arroz elaborado con diferentes estándares de calidad (grano entero). Como ya fue señalado, un completo análisis a la construcción de este indicador y su capacidad de representar el costo de producción de arroz en Chile se encuentra en el artículo de [10].

De este modo, en el análisis tampoco se hace necesario establecer supuestos sobre la formación del precio mayorista, puesto que se utilizará este indicador de costos de público conocimiento.

Esta ventaja en términos de la información disponible, permite simplificar el foco del análisis, evaluando la influencia que sobre la generación de utilidades de retailers y productores se explica por la demanda y las características de los productos. En base a distintos escenarios contrafactuales de la oferta de productos de parte de los agentes, el objetivo es examinar el resultado económico en situaciones hipotéticas considerando la conducta que se espera de estos en un contexto de competencia de productos diferenciados. La literatura antes citada, en particular [34], genera estos valores de equilibrio mediante modelos de doble marginalización, tanto en distribuidores del retail como en productores, y será también el enfoque del modelo de oferta a emplear en el siguiente análisis.

La información de precios mayoristas o precios negociados refleja el resultado de un acuerdo contractual entre privados y, por lo general, son datos que no están disponibles para fines de investigación. Similar a lo que ocurre con los costos, se trata de información estratégicamente sensible para productores y distribuidores. Sin embargo, en los últimos años, al ampliarse las posibilidades de transparentar y disponer de información en transacciones B2B, se han identificado elementos que inciden en el resultado observado, que en general se definen y reconocen como componentes del “poder de negociación” que poseen los agentes que intervienen en esas transacciones, aspectos que no podrán ser explorados en el modelo de oferta aquí desarrollado<sup>11</sup>.

A continuación se presentan las definiciones necesarias para el análisis a las rentas generadas en la venta minorista y también las que corresponden a las empresas productoras, asumiendo una conducta de competencia Nash-Bertrand de los agentes y sin considerar el enfoque de negociación bilateral. Si bien el modelo puede extenderse a conductas no competitivas y también asumiendo un rol de integración vertical del productor-distribuidor, el contar con información efectiva de precios negociados y de indicadores de costos de producción permite que el chequeo de robustez no requiera ser efectuado entre distintos modelos y supuestos, sino frente a los datos de márgenes precio-costo efectivamente disponibles.

Se define entonces a  $F$  cadenas de retail que comercializan un conjunto  $\mathcal{J}$  de  $j = 1, \dots, J$  diferentes tipos de producto. La utilidad por la venta de estos distintos ítemes para el retailer  $f$  en el mercado  $t$  estará dada por la expresión:

$$\Pi_{ft} = \sum_{j \in \mathcal{J}} (p_{jt} - p_{jt}^w) T s_{jt}(p) - C_f \quad (3.12)$$

La dimensión que define el escalar  $T$  para el tamaño total del mercado, permite una cuantificación en términos de unidades, que se obtiene a partir del respectivo market share  $s_{jt}(p)$  de cada producto  $j$  en ese mercado  $t$ . En el contexto de bienes diferenciados que compiten por la preferencia de los consumidores, este market share se encuentra en función del precio propio, pero también del precio de los demás productos que se presentan como opciones a elegir, lo que implica que cada market share depende de todos los precios que se agrupan en el vector  $p$ . Los ingresos por ventas se explican por el diferencial entre el precio ofertado  $p_{jt}$  y el precio mayorista  $p_{jt}^w$ , que corresponde al valor que el retailer paga a la empresa productora para cada uno de los productos  $j \in \mathcal{J}$  que vende. Este precio mayorista es equivalente al costo para el retailer en términos unitarios por ese producto, mientras que la expresión  $C_f$  corresponde a un costo fijo de comercialización.

En el setting aquí descrito, el retailer  $f$  toma la decisión de oferta y es quien establece precios para cada producto a comercializar. Asumiendo la existencia de un equilibrio Nash-Bertrand de estrategias puras en precios, y que estos son estrictamente positivos, el precio óptimo para el producto  $j$ ,  $p_{jt}$ , debe satisfacer la siguiente condición de primer orden:

---

<sup>11</sup>Esta nueva literatura se hace cargo de limitaciones que un modelo de doble marginalización posee para capturar elementos como aversión al riesgo, impaciencia, y también las propias habilidades de negociación de las partes que inciden en la determinación de precios negociados, en transacciones de agentes que interactúan en formas de relación vertical en una cadena de suministro. Los trabajos de [25, 26] incorporan modelos de negociación al análisis estructural.

$$s_{jt}(p) + \sum_{m \in \mathcal{J}} (p_{mt} - p_{mt}^w) \frac{\partial s_{mt}}{\partial p_{jt}} = 0 \quad \forall j \in \mathcal{J} \quad (3.13)$$

Luego, se define a  $O_f$  como una “matriz de oferta” con el elemento general  $O_f(i, j)$  igual a 1 cuando los productos  $i$  y  $j$  son vendidos por el retailer  $f$ , e igual a 0 en otro caso. Junto con esto, se establece a  $\Delta_{ft}$  como la “matriz de respuesta” del retailer  $f$  en el mercado  $t$ , que contiene las derivadas parciales de todos los shares de los productos con respecto a sus precios, con el elemento  $\Delta_{ft}(i, j) = \frac{\partial s_{jt}}{\partial p_{it}}$ .

Se puede advertir que las condiciones de primer orden dadas por la ecuación (3.13) para el precio de todos los productos, pueden apilarse una a una conformando vectores columna con sus respectivos términos. Usando la notación descrita en el párrafo anterior y reordenando elementos, se obtiene la siguiente expresión vectorial para los márgenes de precio:

$$p_t - p_t^w = -(O_f * \Delta_{ft})^{-1} s_t(p) \quad (3.14)$$

De este modo, se tiene que el vector de márgenes de precio al lado izquierdo de esta ecuación se encuentra en función de la oferta que posee el retailer y de variables de la demanda en cada mercado  $t$  en el lado derecho, donde  $O_f * \Delta_{ft}$  corresponde a la multiplicación elemento por elemento de las dos matrices. Si el equilibrio es único, esta ecuación permite al retailer encontrar los precios óptimos para vender los productos y maximizar sus utilidades.

Por parte de las empresas productoras, cada una de estas, denominadas  $w$ , maximizarán sus utilidades eligiendo los precios mayoristas  $p_j^w$  de los productos que elaboran sabiendo que los retailers se comportan en forma óptima de acuerdo con la ecuación (3.14). Al asumir que estas fijan precios acordes con estrategias puras del equilibrio Nash-Bertrand, las condiciones de primer orden para los precios mayoristas que presentarán los productores se expresan en forma vectorial de la siguiente manera:

$$p_t^w - c_t^w = -(O_w * \Delta_{wt})^{-1} s_t(p), \quad (3.15)$$

donde  $c_t^w$  corresponde al costo marginal del productor. En esta expresión se hace uso de la matriz  $O_w$ , con el elemento general  $O_w(i, j)$  que es igual a 1 cuando los productos  $i$  y  $j$  son manufacturados por el productor  $w$  y 0 en otro caso, motivo por el que en esta literatura se la denomina "matriz de propiedad". Por otra parte, la matriz  $\Delta_{wt}$  en esta ecuación contiene también la respuesta en los shares de mercado de los productos que ofrece la firma  $w$ , pero en este caso, frente a los cambios en los precios mayoristas que el productor establece para su oferta, es decir, contiene un elemento general  $\Delta_{wt}(i, j) = \frac{\partial s_{jt}}{\partial p_{it}^w}$ .

Como advierte [34], encontrar una expresión cerrada para estas derivadas es difícil incluso para el caso logit. Sin embargo, la autora hace notar que  $\Delta_{wt} = \Delta_{pt}' \Delta_{ft}$ , por lo tanto, conociendo  $\Delta_{ft}$  solo se requiere obtener los elementos que conforman la matriz  $\Delta_{pt}$  que representa el *pass-through* de los precios mayoristas a los precios del retail, y para esto,

propone una derivación a partir de la diferenciación total de las condiciones de primer orden (3.14)<sup>12</sup>.

Con los datos disponibles para los precios mayoristas, y una vez que ya han sido estimados los parámetros de la demanda mediante el modelo estructural descrito en la sección anterior, las condiciones de primer orden para retailers y productores permiten simular escenarios contrafactuales donde se comercializan un subconjunto de los productos disponibles. En esos casos, que se explican y describen a continuación, se modifica la oferta de productos y los agentes establecen precios óptimos en base a las ecuaciones (3.14) y (3.15). Los consumidores eligen ahora sobre los productos disponibles produciéndose reacciones de acuerdo con lo que predicen ambas “matrices de respuesta”, que al contener las derivadas de los market share respecto a precios del retail y precios mayoristas, queda en función de parámetros de la demanda.

En el caso del retailer  $f$ , la utilidad obtenida en la situación de escenario contrafactual  $c$ , estará dada por la siguiente expresión, donde ahora los productos comercializados se encuentran en un subconjunto restringido  $\mathcal{D}$  de ítemes disponibles, con precios y cantidades para cada producto en el nuevo escenario  $c$ .

$$\Pi_f^c = \sum_{j \in \{\mathcal{J} \cap \mathcal{D}\}} (p_j^c - p_j^{w*}) q_j^c \quad (3.16)$$

En forma análoga para la firma productora  $w$ , la utilidad generada en el nuevo escenario  $c$  está dada por (3.17), que contiene los precios y cantidades que resultan del nuevo equilibrio  $(p_j^c, p_j^{w*}, q_j^c)$  y reflejan el nuevo “contrato vertical óptimo” entre productor y retail.

$$\Pi_w^c = \sum_{j \in \{\mathcal{J} \cap \mathcal{D}\}} (p_j^{w*} - c_j^w) q_j^c \quad (3.17)$$

Dado que el foco de la presente tesis es el análisis de las ganancias explicadas por la valoración de los productos de arroz de variedad local, es necesario señalar que los escenarios contrafactuales sobre los que se analizan los nuevos equilibrios de mercado consideran la posibilidad de que la variedad “largo ancho” no se encuentre disponible en el mercado, y por ende, deba ser reemplazada por productos sustitutos, especialmente variedad de arroz de origen importado “largo delgado” o bien por el resto de productos de arroz disponibles. De este modo, se estudiará si la sustitución explicada por las preferencias de los consumidores, y su efecto sobre el tamaño del mercado, generan repercusiones a las ganancias de los agentes frente a estas nuevas situaciones, tanto para retailers como para las empresas productoras.

Junto con lo anterior, y considerando que se tiene conocimiento de márgenes precio-costo efectivos tanto para el retail como para la industria productora, la simulación de un equilibrio de mercado acorde al paradigma Nash-Bertrand permitirá chequear si los resultados económicos actuales son consistentes con ese comportamiento competitivo.

---

<sup>12</sup>La referencia con ejemplos y la definición de las expresiones en una forma idónea para programar y obtener los elementos de estas matrices en los casos del modelo mixed logit y también del logit convencional, se presentan en el suplemento online del paper de [34].

### 3.2.1. Escenario 1: Base

El escenario base corresponde al cómputo de las utilidades de retailers y productores por la venta de los productos, empleando las ecuaciones (3.16) y (3.17) sin restringir la oferta de productos, usando los datos disponibles de precios de venta, precios mayoristas efectivos y costos marginales de producción del período bajo análisis junto con las participaciones de mercado. Aquí es importante advertir que no se cuenta con información de los costos fijos u otros costos de comercialización del retailer, y tampoco con otros costos de las firmas productoras distintos al de su procesamiento de arroz desde su estado de materia prima, por lo tanto, estas ganancias corresponderían a márgenes brutos.

### 3.2.2. Escenario 2: Excluye principales productos de arroz largo ancho de origen nacional

En este escenario, se obtienen precios de oferta en cada retailer para los productos de arroz en una situación donde se decide excluir ítemes de las principales variedades de origen nacional de sus estantes. Dado que estos productos, de arroz variedad largo ancho grados 1 y 2, de las marcas nacionales más importantes y también de sus marcas propias, poseen las mayores participaciones de mercado, se presupone que este escenario ejerce presión a la demanda de los consumidores, quienes en base a sus preferencias deberán optar entre productos que no forman parte de sus compras habituales.

Los parámetros de utilidad asociados a las características, atributos y marcas de los productos cobran relevancia, puesto que ahora muchos de esos ítemes preferidos no están disponibles y deben sutituirse por otros. En este plano, el retailer maximizador de utilidad puede cobrar precios más altos por productos que están recibiendo una mayor demanda de parte de consumidores que optan por estos para reemplazar a los ítemes que no se encuentran disponibles. Algo similar debe ocurrir en la respuesta óptima por parte de las firmas productoras.

Esta lógica de interacción en la oferta y la demanda es la que también se espera encontrar en el resto de los escenarios a analizar, y tiene consecuencias para el resultado económico que la venta de los productos de arroz generan a los agentes. Sin embargo, dado que los productos presentan diferencias en base a sus atributos, la sustitución en muchos consumidores no será una alternativa y optarán por la opción neutra para su utilidad que representa el *outside good*, es decir, no comprar arroz, y por lo tanto, el tamaño del mercado se verá afectado y también repercutirá en las ganancias.

### 3.2.3. Escenario 3: Excluye los productos de arroz nacional, largo ancho grado 1, de marcas principales y propias

En este caso, se descartan de la oferta de cada retailer aquellos productos de origen nacional variedad largo ancho grado 1, que se encuentran en el segmento de mayor valor de

cada marca. El objetivo es analizar si la similitud que presentan otros productos en cuanto a la variedad de arroz que contienen y su origen, es capaz de solventar la falta de los ítemes del mejor segmento con otros que tienen un grado menor de calidad, considerando la demanda que estos últimos presentan en cada retailer y utilizando los patrones de sustitución generados por el modelo estructural.

#### **3.2.4. Escenario 4: Excluye los productos de arroz nacional, largo ancho grado 1, de marcas principales**

El objetivo de este escenario es analizar el resultado para el retailer de dejar solo a sus productos de marca propia en el segmento de mayor valor, excluyendo de este a los ítemes de arroz largo ancho grado 1 fabricados por las dos principales empresas manufactureras nacionales, Carozzi y Tucapel.

#### **3.2.5. Escenario 5: Excluye los productos de arroz importado, largo delgado grado 2, de marcas principales y propias**

Finalmente, este escenario permite conocer lo que sucedería en caso de excluir a los productos de arroz largo delgado grado 2, de origen importado, y que se posicionan como los productos del segmento de menor valor en el mercado. Es importante señalar que tanto Carozzi como Tucapel son participantes relevantes de este segmento de productos, no así las marcas propias, que se concentran en la variedad local (ver tabla 2.1).

# Capítulo 4

## Resultados

### 4.1. Estimaciones modelo logit

Los resultados de la estimación vía máxima verosimilitud del modelo logit se presentan en las tablas 4.1 y 4.2, para ambos retailers.

El parámetro de sensibilidad al precio de los productos es estadísticamente significativo en todas las especificaciones para ambos retailers y es además negativo, acorde a lo esperado. Por la homogeneidad en preferencias impuesta en el modelo logit, este resultado se interpreta como un cliente representativo que experimenta desutilidad frente al nivel de precio de los productos.

Las estimaciones indican que existen diferencias entre retailers en el nivel de significancia de las otras variables explicativas además del precio.

Se advierte que en los clientes del retailer QUALITY (ver tabla 4.1), los parámetros en la utilidad que corresponden a variables dummy del fabricante de la marca son estadísticamente significativos en las especificaciones que las incluyen. Esto no ocurre con la estimación logit para el retailer VALUE (ver tabla 4.2), que solo presenta resultados estadísticamente significativos en el parámetro asociado al precio en todas las especificaciones. En cuanto a los parámetros que capturan los atributos de variedad (tipo de arroz) y calidad, solo en algunos casos en las especificaciones del retailer QUALITY se obtienen estimaciones estadísticamente significativas.

La variable dummy “estacionalidad”, que identifica transacciones realizadas en meses de primavera o verano, y que fue incluida para capturar efectos en la demanda atribuibles a la época del año, no es estadísticamente significativa. Lo mismo ocurre con otras variables dummy de control incluidas en la especificación (4) en ambos retailers. En esta se modeló la utilidad con una variable que refleja si el consumidor repite la compra de un producto de la misma marca o tipo que en la compra inmediatamente anterior, como una forma de controlar la “lealtad” del consumidor en su función de utilidad. Sin embargo, su parámetro asociado no resultó estadísticamente significativo en ninguno de los retailers. Similar resultado se observa



Tabla 4.1: Resultados estimación modelo logit, retailer QUALITY

VARIABLES	(1)	(2)	(3)	(4)
Precio	-0.63 (0.02) [-28.00]	-0.70 (0.03) [-20.88]	-0.51 (0.02) [-22.90]	-0.61 (0.03) [-20.53]
Ancho grado 1		-0.33 (0.27) [-1.24]	0.28 (0.24) [1.16]	0.27 (0.26) [1.02]
Ancho grado 2		-1.03 (0.35) [-2.93]	-0.50 (0.33) [-1.50]	-0.53 (0.35) [-1.53]
Delgado grado 2		-1.13 (0.49) [-2.29]	-0.98 (0.48) [-2.04]	-1.01 (0.49) [-2.07]
Estacionalidad		22.18 (1881.22) [0.01]		
Carozzi			-0.77 (0.28) [-2.77]	-0.76 (0.29) [-2.61]
Tucapel			-1.05 (0.27) [-3.86]	-1.03 (0.29) [-3.60]
Marca propia			-1.29 (0.52) [-2.47]	-1.37 (0.53) [-2.60]
Repite tipo o marca				19.29 (878.71) [0.02]
Comuna sector oriente				-0.16 (0.27) [-0.58]

*Nota:* Especificaciones con  $N=100,000$ . Bajo cada parámetro se presenta el valor de su (desviación estándar) y su respectivo [estadígrafo t].

Tabla 4.2: Resultados estimación modelo logit, retailer VALUE

Variabes	(1)	(2)	(3)	(4)
Precio	-0.53 (0.03) [-18.03]	-0.62 (0.04) [-14.17]	-0.49 (0.03) [-16.81]	-0.57 (0.04) [-15.59]
Ancho grado 1		-0.29 (0.44) [-0.65]	-0.17 (0.39) [-0.43]	-0.16 (0.42) [-0.39]
Ancho grado 2		-0.35 (0.39) [-0.92]	-0.25 (0.33) [-0.74]	-0.28 (0.36) [-0.79]
Delgado grado 2		-0.49 (0.46) [-1.07]	-0.41 (0.41) [-0.99]	-0.45 (0.43) [-1.04]
Estacionalidad		17.27 (275.14) [0.06]		
Carozzi			-0.04 (0.41) [-0.09]	-0.07 (0.44) [-0.16]
Tucapel			-0.05 (0.34) [-0.16]	-0.06 (0.37) [-0.17]
Marca propia			-0.21 (0.67) [-0.32]	-0.28 (0.70) [-0.40]
Repite tipo o marca				18.8 (982.79) [0.02]
Comuna sector oriente				-0.06 (0.58) [-0.11]

*Nota:* Especificaciones con  $N=100,000$ . Bajo cada parámetro se presenta el valor de su (desviación estándar) y su respectivo [estadígrafo t].

para la variable dummy que identifica compras realizadas en establecimientos ubicados en comunas del sector oriente de Santiago, para controlar por aspectos socioeconómicos y su asociación con la ubicación del supermercado en comunas de mayores ingresos.

Estos resultados son indicativos de una demanda en donde, además del precio, las marcas de los productos jugarían un papel en las preferencias de algunos consumidores. Asimismo, existiría diferenciación en el mercado objetivo que acude a las tiendas de cada retailer, explicado por distintas valoraciones en variables de atributos obtenidas en las estimaciones. Esta heterogeneidad de preferencias en el mercado de productos de arroz que abastece cada retailer no puede capturarse adecuadamente con un consumidor representativo, y hacia esa mayor estructura avanza el modelo mixed logit.

## 4.2. Estimaciones modelo mixed logit

### 4.2.1. Distribución de probabilidad de las preferencias

Como fue señalado en el capítulo 3, la estimación del modelo mixed logit permite reconocer la heterogeneidad en las preferencias de los consumidores por las características de los productos. A esto se suma que la información disponible hace posible identificar distintos parámetros de la demanda en cada retailer, que como se pudo apreciar con la estimación del modelo logit tradicional, arrojó resultados distintos para la demanda en los mercados abastecidos por sus respectivas tiendas.

Como el enfoque de estimación utilizado en este trabajo es no paramétrico, las distribuciones de probabilidad de estos coeficientes tendrán como soporte la grilla de puntos  $\beta^r$  de dimensión  $R$  que sean generados por simulación en las especificaciones que se definan.

[12] destacan la importancia de tener una cobertura de la grilla en el área relevante del espacio de parámetros, y recomiendan al economista centrar esa búsqueda en las estimaciones iniciales del modelo logit. Afirman que una distribución Uniforme posee superiores resultados de convergencia que una aleatoria, porque en presencia de limitaciones computacionales y de datos asegura una mejor cobertura de puntos en el soporte definido.

Por estas razones, el soporte de puntos usados en las estimaciones se genera desde una distribución Uniforme, en el rango dado por múltiplos de las desviaciones estándar de las estimaciones logit. El rango de dominio de los coeficientes se presenta en detalle para cada retailer en la tabla 4.3. Las especificaciones a estimar consideran las variables usadas en el modelo logit tradicional de la especificación (3), sin las variables dummies asociadas a estacionalidad, lealtad o ubicación, ya que estas no reportaron ser estadísticamente significativas (ver tablas 4.1 y 4.2).

Las especificaciones usadas en la estimación consideran muestras de 50,000 y 100,000 observaciones y un número  $R$  de 30 a 200 simulaciones para cada parámetro. Como el coeficiente  $\alpha$  es fundamental para el cálculo de la elasticidad precio, es necesario analizar la sensibilidad del estimador  $FKRB$  frente a distintas especificaciones. Para esto, en las tablas 4.4 y 4.5 se

Tabla 4.3: Soporte de distribución Uniforme para grillas de  $(\alpha^r, \beta^r)$  en estimaciones no paramétricas del modelo mixed logit para QUALITY y VALUE

Coeficientes ( $K=7$ )	4 D.E. desde logit				8 D.E. desde logit				12 D.E. desde logit			
	QUALITY		VALUE		QUALITY		VALUE		QUALITY		VALUE	
	<i>min</i>	<i>max</i>	<i>min</i>	<i>max</i>	<i>min</i>	<i>max</i>	<i>min</i>	<i>max</i>	<i>min</i>	<i>max</i>	<i>min</i>	<i>max</i>
Precio	-0.6	-0.4	-0.6	-0.4	-0.7	-0.3	-0.7	-0.3	-0.8	-0.2	-0.8	-0.1
Ancho grado 1	-0.7	1.3	-1.7	1.4	-1.7	2.2	-3.3	3.0	-2.6	3.2	-4.8	4.5
Ancho grado 2	-1.8	0.8	-1.6	1.1	-3.2	2.2	-2.9	2.4	-4.5	3.5	-4.2	3.7
Delgado grado 2	-2.9	0.9	-2.0	1.2	-4.8	2.9	-3.7	2.9	-6.7	4.8	-5.3	4.5
Carozzi	-1.9	0.3	-1.7	1.6	-3.0	1.5	-3.3	3.3	-4.1	2.6	-5.0	4.9
Tucapel	-2.1	0.0	-1.4	1.3	-3.2	1.1	-2.8	2.7	-4.3	2.2	-4.2	4.1
Marca propia	-3.4	0.8	-2.9	2.5	-5.5	2.9	-5.6	5.1	-7.6	5.0	-8.2	7.8

*Nota:* Soporte basado en múltiplos de desviaciones estándar de los coeficientes logit de especificación (3) en cada retailer, presentados en tablas 4.1 y 4.2, respectivamente.

presentan la media y desviación estándar del parámetro  $\alpha$  estimado con diferentes tamaños de muestra de transacciones  $N$ , número de simulaciones  $R$  y distinta amplitud de la grilla de puntos utilizados en la estimación, en cada retailer respectivamente.

Los momentos indican que la estimación no paramétrica es robusta frente a la amplitud de la grilla, más si se dispone de un suficiente número de puntos  $R$ . Tal como lo discute [12], en los resultados obtenidos se puede apreciar el trade-off sesgo-varianza, pues si bien una grilla amplia y con suficiente densidad de puntos ofrece flexibilidad para encontrar el valor poblacional, también existe un mayor número de parámetros a estimar y se incrementa la desviación estándar del estimador.

En las variables que identifican características de marca y variedad de los productos, las tablas 4.6 y 4.7 presentan la media y desviación estándar de los coeficientes. Estos se obtienen para las diferentes grillas utilizadas, pero se escogen solo las especificaciones (2) y (5) para  $N$  y  $R$  por considerarse que equilibran adecuadamente el trade-off de acuerdo con lo analizado para el parámetro de sensibilidad al precio. Los momentos de estos coeficientes confirman lo obtenido en el modelo logit tradicional para el retailer QUALITY, con variables asociadas a fabricantes de los productos que tienen magnitudes significativas sobre la utilidad. En el retailer VALUE, en cambio, las estimaciones de estos coeficientes presentan una varianza más elevada, y por lo tanto, no se puede hacer esa inferencia.

Al obtenerse un valor para la probabilidad del vector  $(\alpha^r, \beta^r)$ , la estimación no paramétrica del modelo permite encontrar una distribución de probabilidad conjunta. Como la cantidad de variables explicativas ( $K = 7$ ) impide la representación de esta distribución, se pueden observar formas bivariadas de esta probabilidad al graficar el coeficiente de una variable de características junto con el del precio (ver figuras 4.1 y 4.3), o simplemente en dos dimensiones cuando se examina la distribución de probabilidad de cada parámetro por separado (ver figuras 4.2 y 4.4).

Los resultados revelan heterogeneidad en la distribución de parámetros de la demanda en

Tabla 4.4: Media y desviación estándar de estimador  $FKRB$  para el coeficiente de sensibilidad al precio, bajo distintas especificaciones y grillas de  $\alpha^r$  en retailer QUALITY

Especificación	$N$	$R$	(a)		(b)		(c)	
			4 D.E.		8 D.E.		12 D.E.	
			Media	D.E.	Media	D.E.	Media	D.E.
(1)	50,000	30	-0.49	0.06	-0.48	0.09	-0.60	0.20
(2)	50,000	50	-0.49	0.06	-0.51	0.10	-0.56	0.17
(3)	100,000	30	-0.53	0.04	-0.53	0.07	-0.60	0.16
(4)	100,000	50	-0.53	0.08	-0.58	0.12	-0.63	0.21
(5)	100,000	100	-0.53	0.06	-0.53	0.09	-0.58	0.17
(6)	100,000	200	-0.54	0.06	-0.58	0.17	-0.64	0.21

*Nota:* El número de variables explicativas es  $K=7$  en todas las especificaciones.

Tabla 4.5: Media y desviación estándar de estimador  $FKRB$  para el coeficiente de sensibilidad al precio, bajo distintas especificaciones y grillas de  $\alpha^r$  en retailer VALUE

Especificación	$N$	$R$	(a)		(b)		(c)	
			4 D.E.		8 D.E.		12 D.E.	
			Media	D.E.	Media	D.E.	Media	D.E.
(1)	50,000	30	-0.45	0.02	-0.48	0.14	-0.61	0.24
(2)	50,000	50	-0.44	0.03	-0.46	0.14	-0.52	0.19
(3)	100,000	30	-0.53	0.08	-0.56	0.13	-0.58	0.14
(4)	100,000	50	-0.53	0.07	-0.55	0.10	-0.64	0.18
(5)	100,000	100	-0.51	0.06	-0.55	0.14	-0.55	0.13
(6)	100,000	200	-0.51	0.07	-0.55	0.14	-0.65	0.20

*Nota:* El número de variables explicativas es  $K=7$  en todas las especificaciones.

ambos grupos de consumidores, aunque esta es acotada ya que no se identifica un número elevado de sub-poblaciones con preferencias muy distintas. No obstante, se advierte que las masas de probabilidad son más altas, es decir, tienen su moda, en valores que no necesariamente corresponden a la media de las distribuciones. Se puede además señalar que las distribuciones en la mayoría de los casos distan de asemejarse a una distribución normal.

En particular para los parámetros de sensibilidad al precio,  $\alpha^r$ , se observa en la comparación de las figuras 4.2 y 4.4 que las masas y grupos de mayor probabilidad se sitúan en zonas distintas en ambos retailers, aún cuando la media entre ellos sea similar (ver tablas 4.4 y 4.5). En la estimación para QUALITY, existen grupos de consumidores que presentan menor sensibilidad al precio, mientras en el caso del grupo que adquiere productos en el retailer VALUE se acumula una masa de probabilidad en valores elevados de disgusto por el nivel del precio.

Tabla 4.6: Media y desviación estándar de estimador  $FKRB$  para coeficientes de características, bajo distintas grillas de  $\beta^r$  y especificaciones (2) y (5) en retailer QUALITY

Variables	(a)		(b)		(c)	
	4 D.E.		8 D.E.		12 D.E.	
	(5)	(2)	(5)	(2)	(5)	(2)
Ancho grado 1	-0.1 (0.5)	-0.1 (0.6)	-0.9 (0.9)	0.0 (1.2)	-1.1 (1.1)	-0.6 (1.7)
Ancho grado 2	-1.0 (0.6)	-0.9 (0.7)	-1.7 (1.3)	-1.5 (1.2)	-1.9 (1.3)	-2.5 (1.4)
Delgado grado 2	-2.1 (1.1)	-1.7 (1.0)	-3.0 (1.1)	-3.1 (1.8)	-4.5 (2.3)	-4.4 (2.2)
Carozzi	-1.0 (0.6)	-1.2 (0.7)	-2.0 (1.3)	-1.4 (1.2)	-1.8 (1.3)	-2.5 (1.4)
Tucapel	-1.3 (0.5)	-1.0 (0.5)	-2.1 (1.0)	-1.3 (1.2)	-2.0 (1.1)	-2.3 (1.5)
Marca propia	-1.5 (0.5)	-1.1 (1.2)	-1.3 (1.0)	-3.8 (1.8)	-2.3 (1.5)	-4.0 (2.6)

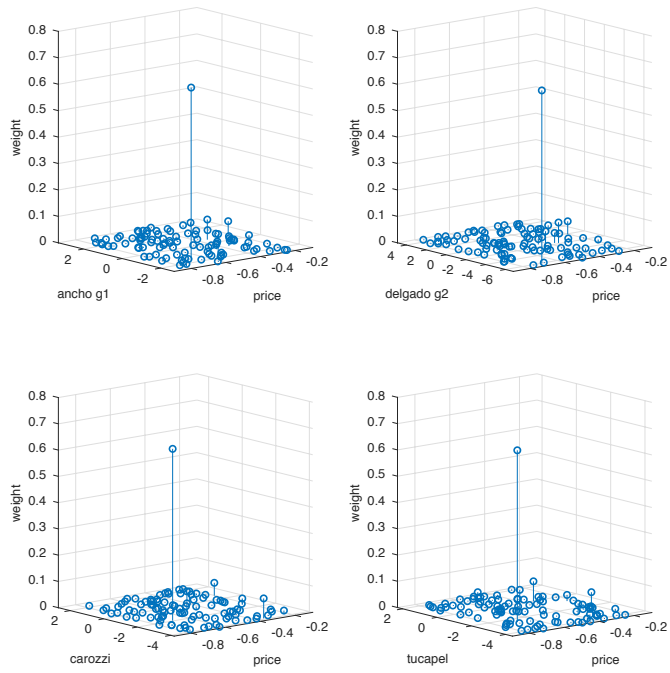
*Nota:* Desviación estándar entre paréntesis.

Tabla 4.7: Media y desviación estándar de estimador  $FKRB$  para coeficientes de características, bajo distintas grillas de  $\beta^r$  y especificaciones (2) y (5) en retailer VALUE

Variables	(a)		(b)		(c)	
	4 D.E.		8 D.E.		12 D.E.	
	(5)	(2)	(5)	(2)	(5)	(2)
Ancho grado 1	-0.3 (0.5)	-0.1 (0.9)	-0.8 (0.9)	-0.3 (1.7)	-2.1 (2.4)	-0.9 (1.5)
Ancho grado 2	-0.6 (0.7)	-0.3 (0.6)	-0.5 (0.8)	-1.0 (1.3)	-1.6 (2.4)	-2.2 (2.1)
Delgado grado 2	-1.1 (1.0)	-0.4 (0.9)	-1.5 (1.6)	-1.2 (1.2)	-1.5 (1.9)	-2.4 (2.3)
Carozzi	-1.0 (0.7)	-0.2 (0.6)	-1.9 (0.8)	-1.1 (1.3)	-1.2 (2.4)	-1.6 (2.1)
Tucapel	-0.3 (0.5)	-0.3 (0.8)	-0.9 (1.2)	-0.7 (0.9)	-1.0 (1.4)	-2.1 (2.3)
Marca propia	-0.6 (1.1)	-1.0 (1.9)	-2.0 (1.7)	-2.8 (2.3)	-1.6 (2.9)	-4.5 (3.2)

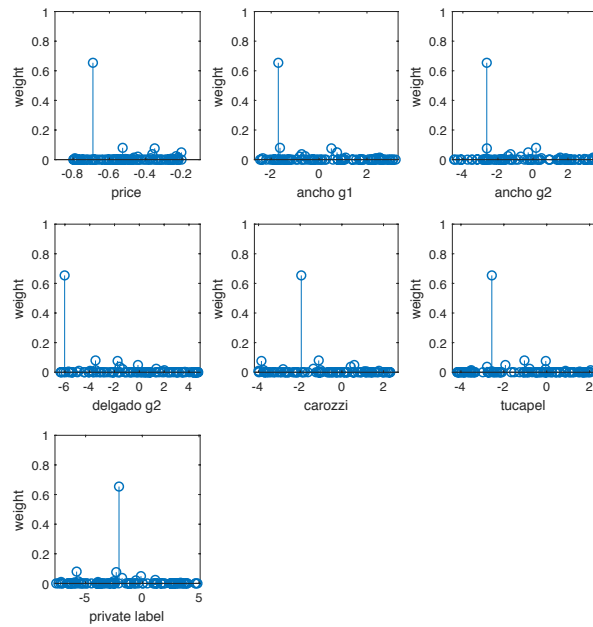
*Nota:* Desviación estándar entre paréntesis.

Figura 4.1: Distribución conjunta en tuplas elegidas de  $(\alpha^r, \beta^r, \hat{\theta}^r)$ , retailer QUALITY



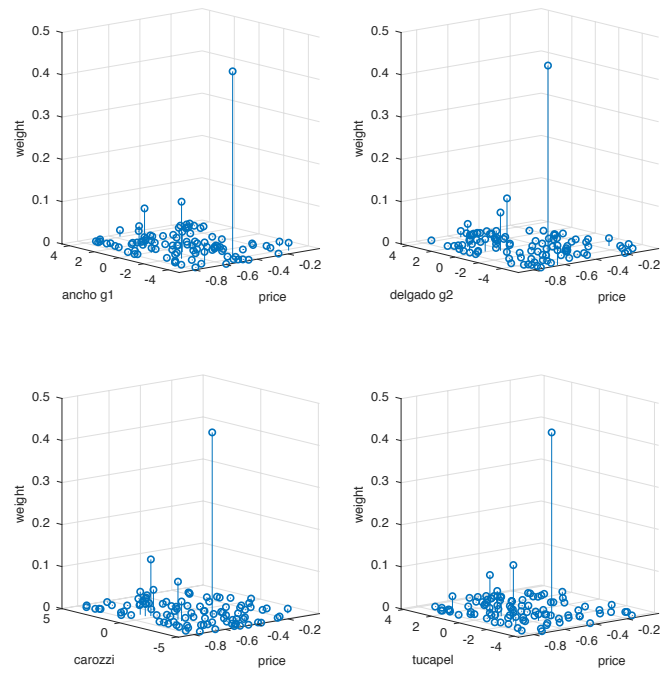
*Nota:* Especificación (5.c), con  $N=100,000$  y  $R=100$ .

Figura 4.2: Plot en 2-D para la probabilidad  $\hat{\theta}^r$  estimada para  $\alpha^r$  y distintos  $\beta^r$ , retailer QUALITY



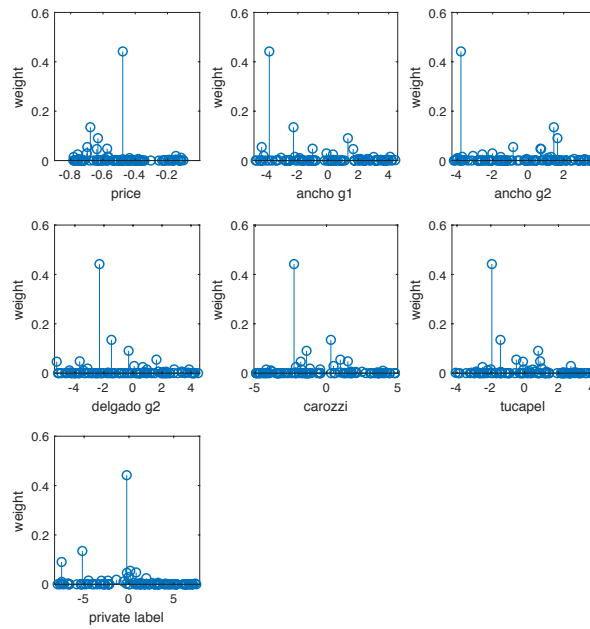
*Nota:* Especificación (5.c), con  $N=100,000$  y  $R=100$ .

Figura 4.3: Distribución conjunta en tuplas elegidas de  $(\alpha^r, \beta^r, \hat{\theta}^r)$ , retailer VALUE



*Nota:* Especificación (5.c), con  $N=100,000$  y  $R=100$ .

Figura 4.4: Plot en 2-D para la probabilidad  $\hat{\theta}^r$  estimada para  $\alpha^r$  y distintos  $\beta^r$ , retailer VALUE



*Nota:* Especificación (5.c), con  $N=100,000$  y  $R=100$ .



## 4.2.2. Elasticidades y patrones de sustitución

De acuerdo con lo señalado en el capítulo 3, la estimación de un modelo mixed logit enriquece el análisis de la demanda en un modelo estructural, porque los parámetros heterogéneos son capaces de generar la correlación en el espacio de características que es necesaria para encontrar patrones de sustitución acordes a la similitud de los productos. Los resultados de los cálculos de elasticidad precio propio y precio cruzado se presentan en las tablas 4.8 y 4.9, para ambos retailers.

Para el desarrollo de estos cálculos, y para facilitar la exposición de los resultados, se realizó una agrupación de productos de acuerdo a su marca y tipo, idéntica a la utilizada para presentar la estadística descriptiva de los datos provenientes de las muestras usadas. La obtención de las elasticidades es posible gracias a la fórmula cerrada del modelo logit, y al conocimiento de la distribución de parámetros que es resultado de la estimación de la demanda. La referencia se encuentra en [23].

Tabla 4.8: Elasticidades modelo mixed logit, retailer QUALITY

	Carozzi ag1	Tucapel ag1	M. Propia ag1	Carozzi ag2	Tucapel ag2	M. Propia ag2	Carozzi dg2	Tucapel dg2	Carozzi otro	Tucapel otro	M. Propia otro	Otras marcas	Resto otro tipo
Carozzi ag1	-1.05	0.04	0.05	0.10	0.02	0.02	0.04	0.01	0.07	0.01	0.01	0.14	0.06
Tucapel a1	0.07	-2.23	0.03	0.02	0.18	0.02	0.01	0.05	0.01	0.09	0.01	0.19	0.12
M. Propia ag1	0.24	0.08	-1.63	0.07	0.10	0.09	0.03	0.03	0.05	0.03	0.04	0.21	0.08
Carozzi ag2	0.55	0.06	0.09	-2.45	0.22	0.09	0.03	0.01	0.08	0.02	0.02	0.17	0.09
Tucapel ag2	0.04	0.19	0.04	0.07	-1.75	0.08	0.01	0.11	0.01	0.04	0.01	0.14	0.03
M. Propia ag2	0.18	0.10	0.22	0.18	0.53	-2.03	0.02	0.02	0.03	0.03	0.05	0.16	0.05
Carozzi dg2	0.66	0.12	0.10	0.08	0.05	0.03	-2.30	0.12	0.09	0.02	0.04	0.30	0.10
Tucapel dg2	0.08	0.21	0.05	0.02	0.43	0.01	0.05	-1.59	0.01	0.04	0.03	0.23	0.06
Carozzi otro	0.65	0.06	0.09	0.13	0.04	0.03	0.05	0.02	-4.00	0.02	0.03	0.29	0.16
Tucapel otro	0.05	0.34	0.04	0.02	0.14	0.01	0.01	0.04	0.02	-3.47	0.02	0.23	0.14
M. Propia otro	0.19	0.10	0.14	0.06	0.08	0.07	0.04	0.07	0.05	0.05	-2.83	0.38	0.13
Otras marcas	0.27	0.23	0.09	0.06	0.16	0.03	0.04	0.07	0.07	0.07	0.05	-5.09	0.20
Resto otro tipo	0.04	0.04	0.01	0.01	0.01	0.00	0.00	0.01	0.01	0.01	0.01	0.06	-1.14

Nota: Cálculos basados en estimación de especificación (5.c).

Tabla 4.9: Elasticidades modelo mixed logit, retailer VALUE

	Carozzi ag1	Tucapel ag1	M. Propia ag1	Carozzi ag2	Tucapel ag2	M. Propia ag2	Carozzi dg2	Tucapel dg2	M. Propia dg2	Carozzi otro	Tucapel otro	M. Propia otro	Otras marcas	Resto otro tipo
Carozzi ag1	-2.27	0.08	0.01	0.08	0.06	0.00	0.04	0.06	0.01	0.02	0.01	0.00	0.16	0.06
Tucapel a1	0.12	-2.99	0.05	0.02	0.15	0.02	0.03	0.08	0.02	0.01	0.02	0.01	0.12	0.05
M. Propia ag1	0.05	0.16	-2.85	0.01	0.10	0.04	0.01	0.07	0.06	0.00	0.01	0.03	0.14	0.04
Carozzi ag2	0.20	0.04	0.00	-2.40	0.10	0.01	0.05	0.03	0.01	0.02	0.01	0.00	0.20	0.06
Tucapel ag2	0.05	0.09	0.02	0.03	-2.97	0.03	0.01	0.07	0.01	0.01	0.02	0.01	0.12	0.05
M. Propia ag2	0.02	0.08	0.06	0.02	0.19	-2.87	0.01	0.07	0.04	0.00	0.02	0.02	0.15	0.05
Carozzi dg2	0.20	0.08	0.01	0.08	0.08	0.01	-2.55	0.13	0.02	0.02	0.02	0.01	0.23	0.06
Tucapel dg2	0.11	0.11	0.03	0.03	0.15	0.02	0.06	-2.67	0.04	0.01	0.03	0.01	0.17	0.05
M. Propia dg2	0.04	0.09	0.09	0.02	0.11	0.04	0.02	0.12	-2.53	0.01	0.02	0.02	0.20	0.06
Carozzi otro	0.16	0.04	0.01	0.08	0.06	0.01	0.04	0.05	0.01	-2.55	0.01	0.01	0.13	0.06
Tucapel otro	0.06	0.08	0.02	0.02	0.16	0.02	0.02	0.07	0.02	0.01	-3.23	0.01	0.11	0.06
M. Propia otro	0.03	0.06	0.07	0.02	0.11	0.04	0.01	0.06	0.04	0.01	0.02	-2.77	0.13	0.06
Otras marcas	0.08	0.04	0.02	0.04	0.07	0.01	0.02	0.04	0.02	0.01	0.01	0.01	-2.84	0.06
Resto otro tipo	0.04	0.02	0.01	0.02	0.04	0.01	0.01	0.02	0.01	0.00	0.01	0.00	0.08	-2.59

Nota: Cálculos basados en estimación de especificación (5.c).

En relación a valores de elasticidades encontrados en otros estudios de demanda de productos alimenticios comercializados por el retail en Chile, los resultados son indicativos de una menor elasticidad precio propio si se lo compara con productos como café [25], y bastante similar a lo encontrado en el mercado de aceites, en cadenas de retail que cuentan también con marcas propias [7]. En el mercado del arroz aquí analizado, el resultado indica que es

relativamente menor en algunos grupos de productos, en particular para la variedad largo ancho de origen nacional<sup>1</sup>, lo que es más pronunciado en consumidores del retailer QUALITY.

Las elasticidades de precio entre productos – o “cruzada” – muestran un patrón interesante en el caso del arroz largo ancho grado 1 de marcas de Carozzi. En estos productos revelan que en caso de alzas de precio los consumidores tienden a sustituirlos con otros del mismo fabricante, lo que se observa en la demanda de ambos retailers. Este patrón de sustitución no se presenta con la misma intensidad en el caso del producto ancho grado 1 de Tucapel.

En las marcas propias, se observa que en el retailer QUALITY estas desafían en una magnitud moderada a los productos de Carozzi (elasticidad cruzada de 0.24 en el arroz ancho grado 1), y en una magnitud mayor a los productos de Tucapel (hasta 0.53 en el arroz ancho grado 2). En VALUE se repite esta tendencia, con productos de marcas propias que suelen tener elasticidad cruzada más elevada (de hasta 0.19 en el segmento de arroz ancho grado 2) con los productos de Tucapel que con los de Carozzi.

Al examinar los productos de arroz largo delgado grado 2 de origen importado, las elasticidades cruzadas que presentan son mayores con productos del mismo segmento (de hasta 0.13 en QUALITY y 0.12 en VALUE) que respecto al resto de productos, lo que predice que las alzas de precio de estos productos tendría un efecto acotado en la demanda del resto de variedades.

Sin embargo, alzas de precio de los segmentos líderes de participación de mercado conformados por variedades de origen nacional, muestran ser desafiadas por el segmento de menores precios que integran los productos de variedad importada, con elasticidades cruzadas de hasta 0.66 en el caso del producto largo delgado grado 2 de Carozzi respecto al largo ancho grado 1 del mismo fabricante en QUALITY, y de 0.20 en VALUE en la elasticidad entre esas mismas variedades y fabricante.

### 4.2.3. Resultados de escenarios contrafactuales

Con los resultados de la estimación de demanda estructural, y en base al modelo de oferta definido, se simulan las participaciones de mercado y las ganancias de retailers y empresas productoras por la venta de los productos en escenarios contrafactuales, a partir de las expresiones dadas por las ecuaciones (3.16) y (3.17) presentadas en dicha sección.

Al modificarse el set de productos disponibles, los parámetros de preferencias estimados para la demanda de los productos y el modelo de oferta permiten generar precios de equilibrio en cada nueva situación de mercado<sup>2</sup>. Estos equilibrios implican nuevos shares para los productos, los que se presentan para cada escenario en las tablas 4.10 y 4.11.

En ambos retailers, la exclusión de productos de origen nacional (escenario 2), y del

---

<sup>1</sup>Estos productos poseen un share de 50.2% del mercado de arroz en el retailer QUALITY y un 42.7% en el retailer VALUE (ver tabla 2.1), por lo que son sin duda los más relevantes.

<sup>2</sup>La obtención de los precios de equilibrio implicados por las condiciones de primer orden de los nuevos escenarios se realiza sobre las mismas muestras de transacciones utilizadas en las estimaciones de demanda.

Tabla 4.10: Predicción de shares en escenarios contrafactuales, retailer QUALITY

Productos	Share efectivo base	Share modelo	Escenarios				Escenario N-B
			2	3	4	5	
Carozzi ag1	3.15 %	3.25 %				3.65 %	3.30 %
Tucapel ag1	1.82 %	2.02 %				2.32 %	2.14 %
M. Propia ag1	0.69 %	0.69 %			0.96 %	0.86 %	0.73 %
Carozzi ag2	0.57 %	0.69 %		1.23 %	1.16 %	0.80 %	0.74 %
Tucapel ag2	1.78 %	1.76 %		2.03 %	1.98 %	1.93 %	1.82 %
M. Propia ag2	0.28 %	0.28 %		0.39 %	0.35 %	0.33 %	0.30 %
Carozzi dg2	0.20 %	0.30 %	0.60 %	0.51 %	0.48 %		0.30 %
Tucapel dg2	0.48 %	0.43 %	1.00 %	0.54 %	0.52 %		0.44 %
Carozzi otro	0.36 %	0.66 %	1.28 %	1.07 %	1.01 %	0.31 %	0.70 %
Tucapel otro	0.46 %	0.62 %	0.98 %	0.81 %	0.79 %	0.71 %	0.66 %
M. Propia otro	0.21 %	0.30 %	0.52 %	0.41 %	0.38 %	0.39 %	0.32 %
Otras marcas	1.59 %	3.52 %	4.83 %	5.01 %	4.83 %	3.32 %	4.05 %
Resto otro tipo	4.93 %	1.94 %	2.70 %	2.20 %	2.12 %	1.92 %	1.80 %
Total <i>inside goods</i>	16.52 %	16.46 %	11.91 %	14.20 %	14.60 %	16.53 %	17.30 %
Diferencia respecto a base (%)		-0.4 %	-27.9 %	-14.0 %	-11.6 %	0.0 %	4.7 %

*Nota:* Escenarios basados en estimación de especificación (5.c), en muestra de  $N=100,000$  transacciones.

Tabla 4.11: Predicción de shares en escenarios contrafactuales, retailer VALUE

Productos	Share efectivo base	Share modelo	Escenarios				Escenario N-B
			2	3	4	5	
Carozzi ag1	1.11 %	1.09 %				1.24 %	1.15 %
Tucapel ag1	0.75 %	0.77 %				0.91 %	0.86 %
M. Propia ag1	0.25 %	0.25 %			0.28 %	0.28 %	0.26 %
Carozzi ag2	0.45 %	0.48 %		0.58 %	0.57 %	0.54 %	0.53 %
Tucapel ag2	1.31 %	1.31 %		1.55 %	1.54 %	1.50 %	1.43 %
M. Propia ag2	0.19 %	0.17 %		0.20 %	0.19 %	0.20 %	0.18 %
Carozzi dg2	0.26 %	0.26 %	0.37 %	0.34 %	0.33 %		0.28 %
Tucapel dg2	0.60 %	0.59 %	0.77 %	0.69 %	0.68 %		0.62 %
M. Propia dg2	0.19 %	0.20 %	0.23 %	0.22 %	0.21 %		0.20 %
Carozzi otro	0.10 %	0.10 %	0.14 %	0.12 %	0.12 %	0.11 %	0.11 %
Tucapel otro	0.20 %	0.21 %	0.27 %	0.24 %	0.24 %	0.24 %	0.23 %
M. Propia otro	0.10 %	0.12 %	0.15 %	0.14 %	0.14 %	0.14 %	0.13 %
Otras marcas	2.33 %	2.44 %	1.60 %	2.72 %	2.79 %	1.76 %	2.62 %
Resto otro tipo	1.67 %	1.43 %	1.73 %	1.61 %	1.60 %	1.60 %	1.56 %
Total <i>inside goods</i>	9.49 %	9.44 %	5.26 %	8.41 %	8.70 %	8.51 %	10.17 %
Diferencia respecto a base (%)		-0.6 %	-44.6 %	-11.4 %	-8.3 %	-10.3 %	7.1 %

*Nota:* Escenarios basados en estimación de especificación (5.c), en muestra de  $N=100,000$  transacciones.

subgrupo de mayor estándar de calidad (escenarios 3 y 4, que solo excluyen productos grado 1), genera una reducción del mercado. En el caso del retailer VALUE la reducción es casi de 45 % en el escenario que no tiene productos de variedad largo ancho (ver última fila de tabla 4.11). Los consumidores, frente a la nueva oferta disponible, realizan elecciones de compra sustituyendo de acuerdo a sus preferencias por productos de otras variedades, o bien optando por el *outside good* no comprando arroz, y los resultados indican que esta última alternativa es escogida frecuentemente cuando no hay productos de variedad nacional. Los escenarios 3 y 4 indican, además, que al no encontrarse los productos de mejor calidad del mercado, estos se sustituyen por otros de la misma variedad nacional largo ancho, pero de grado 2. Las marcas propias de arroz en el escenario 4 indican que son solo parcialmente capaces de hacer frente a las ausencias de los competidores de *national brands* más importantes de este mercado.

Lo observado para las variedades de origen nacional resulta muy distinto a lo que se produce con la exclusión de productos de la variedad importada de arroz largo delgado (escenario 5), de acuerdo con lo reportado en las tablas 4.10 y 4.11. Se puede apreciar que en el retailer QUALITY, los consumidores incrementan el consumo de la variedad nacional cuando los productos de arroz largo delgado no están disponibles, y así el tamaño del mercado se mantiene prácticamente inalterado. En el caso de VALUE, se produce un desvío de la demanda hacia los otros productos, pero este no es total y el tamaño del mercado experimenta una reducción del 10 %.

En cuanto a las utilidades o ganancias, el conocimiento que se tiene del precio de venta a consumidor y el precio mayorista, implica tener certeza de la parte más relevante del costo marginal que para el retailer significa vender cada producto. Sin embargo, al no conocerse otros componentes del costo de la distribución minorista que van también dentro del costo marginal, como por ejemplo el gasto de administración de tiendas del retailer, estas ganancias se deben interpretar como resultado de un margen de utilidad directo contable – bruto – por la venta de los productos. En forma análoga para las ganancias de la empresa productora, el costo marginal de producción de arroz se deriva exclusivamente del procesamiento del cereal, y por lo tanto, no incluye otros componentes.

Teniendo presente esta advertencia, las estimaciones de las ganancias de cada retailer y empresas productoras en el contrato óptimo entre ellos en el equilibrio de los escenarios contrafactuales, se presenta en las tablas 4.12 y 4.13, para los mercados abastecidos por QUALITY y VALUE respectivamente.

En cada tabla, la primera fila toma como base de comparación la situación efectiva que informan los datos disponibles, y que se contrasta con las ganancias totales del retailer y de las empresas productoras en los escenarios contrafactuales simulados y, en la última columna, un escenario con el resultado de equilibrio con todos los productos de la oferta actual bajo el supuesto de competencia Nash-Bertrand (en adelante, escenario N-B). En la segunda fila de cada tabla se comparan las ganancias de los agentes tomando como base de contraste a las utilidades que se obtendrían en el escenario N-B con todos los productos. En esta comparación, la base de contraste es menor, puesto que el escenario N-B logra explicar parcialmente las ganancias obtenidas.

En efecto, para el retailer QUALITY, el escenario N-B es capaz de explicar el 49 % de las ganancias actuales del distribuidor y un 74 % de las ganancias actuales de las empresas

Tabla 4.12: Ganancias por la venta de productos de arroz en escenarios contrafactuales, retailer QUALITY

Ganancias		Base efectiva	(2)	Escenarios			Escenario N-B
				(3)	(4)	(5)	
Totales sobre base efectiva (en %)	Retailer	100 %	26.3 %	34.7 %	34.9 %	47.3 %	49.3 %
	Productores	100 %	33.6 %	60.6 %	60.3 %	74.3 %	74.0 %
Totales sobre escenario N-B (en %)	Retailer	202.9 %	53.3 %	70.5 %	70.9 %	95.9 %	100 %
	Productores	135.2 %	45.4 %	82.0 %	81.5 %	100.4 %	100 %
Unitarias (\$/kg)	Retailer	\$256	\$93	\$104	\$101	\$121	\$121
	Productores	\$185	\$86	\$131	\$126	\$138	\$131

*Nota:* Cálculos basados en estimaciones de especificación (5.c).

Tabla 4.13: Ganancias por la venta de productos de arroz en escenarios contrafactuales, retailer VALUE

Ganancias		Base efectiva	(2)	Escenarios			Escenario N-B
				(3)	(4)	(5)	
Totales sobre base efectiva (en %)	Retailer	100 %	37.0 %	54.1 %	55.8 %	59.9 %	66.1 %
	Productores	100 %	35.4 %	50.9 %	52.3 %	58.7 %	64.2 %
Totales sobre escenario N-B (en %)	Retailer	151.2 %	55.9 %	81.8 %	84.4 %	90.5 %	100 %
	Productores	155.8 %	55.2 %	79.4 %	81.6 %	91.4 %	100 %
Unitarias (\$/kg)	Retailer	\$228	\$152	\$139	\$139	\$152	\$141
	Productores	\$226	\$144	\$130	\$129	\$148	\$135

*Nota:* Cálculos basados en estimaciones de especificación (5.c).

productoras. Los escenarios contrafactuales que alteran la oferta de los productos revelan reducciones significativas en las ganancias para ambos tipos de agente. Al observar la última fila de la tabla, que presenta el cálculo de las ganancias unitarias en cada escenario, se refleja que la ausencia de productos de variedad largo ancho presiona a la baja el margen unitario que retailer y productores pueden obtener por los productos que quedan en el mercado. Si a esto se agrega que la demanda también se restringe (ver última fila de tabla 4.10), se logra así comprender la reducción de ganancias totales.

En el retailer VALUE, la capacidad de explicación de las ganancias con la conducta Nash-Bertrand es similar entre ambos agentes, con un 66 % de la ganancia actual que obtiene el distribuidor y un 64 % de la utilidad actual de las empresas productoras. Al igual que en el caso anterior, los escenarios contrafactuales revelan importantes pérdidas para retailer y productores. No obstante, y a diferencia del caso anterior, en VALUE no se produce una presión a la baja en los márgenes de ganancia unitaria por productos que quedan en la oferta, y lo que explica la disminución en utilidades es la fuerte reducción de la demanda en los nuevos escenarios (ver última fila de tabla 4.11).

En suma, para ambos retailers y también para las empresas productoras, las ganancias implicadas por los nuevos escenarios denotan la importancia que tiene la venta de algunos productos en este mercado, en particular aquellos de variedad de origen nacional.

Como indica [22], no conocer el costo marginal e imponerlo a partir de un modelo de conducta competitiva forma parte de la tradición de la nueva literatura de organización industrial empírica. En el análisis de poder de mercado que este autor presenta respecto a la industria de cereales para el desayuno [24], el equilibrio Nash-Bertrand resultó más consistente a datos contables que otros modelos de conducta no colusiva. En cambio, en el mercado aquí analizado, y advirtiendo que el margen conocido dado por la diferencia entre precio de venta y precio pagado al productor representa una medida incompleta, el paradigma Nash-Bertrand ofrece una respuesta parcial, especialmente en el retailer QUALITY, y motiva la búsqueda de otras explicaciones sobre la forma como el retailer y el productor negocian el precio mayorista.

Una explicación es la ofrecida por [34], quien señala que examinar la demanda en una única categoría de producto puede ser restrictivo, porque los fabricantes no solo comercian productos diferenciados en segmentos distintos de un mismo mercado, sino también en otras categorías con productos diferentes. Esto es válido para el mercado aquí analizado en el caso del productor Carozzi, que tiene fuerte presencia en una amplia gama de alimentos, no solo en arroz. Este hecho no solo afectaría la decisión del precio a negociar en ese producto, sino también mejoraría su flexibilidad negociadora con el retailer. Como indica esta autora, los retailers pueden también utilizar una forma estratégica de fijar precios de venta en algunas categorías de producto para potenciar los viajes a las tiendas, incrementando así las ventas globales. Desde ese punto de vista, un precio atractivo en productos de más importancia en la canasta de consumo familiar generarían una mayor señal comercial, y liberaría a otros productos – como el arroz – donde el margen del retailer puede ser mayor y podría actuar en compensación a las otras categorías.

La segunda explicación es más obvia, y pasa por reconocer que la parte del costo marginal que no está siendo considerada en estos datos puede tener una magnitud importante, sobre

todo para el retailer QUALITY, considerando que para VALUE la explicación del margen bruto alcanza el 66 % y para las empresas manufactureras una cifra equivalente o mayor. Condiciones de diferenciación de las tiendas como su tamaño, ubicación y la inclusión en ellas de otros servicios o su cercanía con otras tiendas de departamento, malls y economías de ámbito en general que se ofrecen al consumidor, implican costos para el retailer QUALITY que un análisis más completo de los márgenes que racionalizan el equilibrio de mercado debe contemplar.

# Conclusión

En Chile, el arroz no es un alimento de niveles elevados de consumo en la comparación internacional, sin embargo, los resultados de la estimación de demanda estructural presentada en este trabajo indican que en grupos de la población la decisión de compra semanal se revela algo más desconsiderada respecto de los precios y se manifiesta acorde a preferencias por variedades y marcas.

Utilizando datos de transacciones individuales efectuadas en dos importantes cadenas de retail, la estimación de demanda y el modelo de oferta en equilibrio competitivo reflejan que existe un significativo resultado económico para el distribuidor minorista y también en las empresas productoras. Asimismo, revelan fidelidad de los consumidores a los fabricantes de *national brands* y a las variedades de producto de origen local, de grano largo ancho. Al revisar las ganancias que obtienen tanto distribuidores como productores, estos resultados sugieren la existencia de un importante poder de negociación frente al retailer en las empresas productoras que poseen estas variedades en su oferta.

Este resultado económico es coherente con el interés de la industria nacional de mantener relaciones de largo plazo con los agricultores que producen la variedad de origen nacional, y su vinculación con organismos públicos que realizan investigación y desarrollo en este rubro. También el resultado es consistente con la entrada de marcas propias del retail ocurrida en el período bajo análisis y el interés por su posicionamiento en el mercado. El modelo de oferta basado en estrategias Nash-Bertrand explica del orden del 64-74 % de los datos de margen contable bruto para las empresas productoras y 66 % para el retailer VALUE. Sin embargo, en QUALITY la explicación del margen es cercana al 50 %. Las características diferenciadoras como tienda tipo “hipermercado” y elementos asociados a la negociación entre este retailer y productores, que no pueden abordarse con el modelo aquí desarrollado, motivan la búsqueda de otras respuestas que complementen el análisis a los márgenes precio-costos y a las utilidades totales.

Las preferencias por ciertos atributos develada por la estimación de demanda, también tiene implicancias “aguas arriba” en la cadena de suministro de este producto. La compra de arroz de variedad local a agricultores por parte de la industria procesadora, especialmente por parte de las dos empresas fabricantes de *national brands* con fuerte presencia en este mercado, en base a un indicador del Ministerio de Agricultura construido con los costos de importación, resulta cuestionable desde la perspectiva de la demanda. Esto se debe a que la variedad producida localmente presenta mayor preferencia por parte de los consumidores, y tener estos productos en su oferta genera importantes utilidades en la cadena agroindustrial.



En base a la estimación estructural desarrollada en este trabajo, se puede afirmar que excluir todos los productos de variedad local de la venta en cada retailer genera un muy acotado desplazamiento de la demanda hacia los productos de origen importado, y la respuesta en las preferencias de los consumidores generaría una reducción del tamaño del mercado de este producto. Además, en los escenarios estimados para uno de los retailer se evidencia que la demanda provoca una presión a la baja en márgenes unitarios de los productos restantes de la oferta, lo que se manifiesta tanto para el distribuidor como para las empresas productores. Todos estos factores permiten concluir que un escenario donde la variedad nacional de arroz largo ancho no se encuentre disponible en el mercado generaría importantes pérdidas de utilidades a ambos tipos de agente.

# Bibliografía

- [1] José Roberto Alvarado A. Sistema Integrado de Preparación de Suelos. In *Arroz: Manejo Tecnológico*, chapter 5. INIA, 2007.
- [2] José Roberto Alvarado A. and Pablo Grau B. El Cultivo de Arroz en Chile y sus expectativas. Technical report, INIA, 1986.
- [3] Patrick Bajari, Jeremy T. Fox, and Stephen P. Ryan. Linear regression estimation of discrete choice models with nonparametric distributions of random coefficients. *American Economic Review*, 97(2):459–463, May 2007.
- [4] Bárbara Barrios Aguire. Características del arroz chileno. Boletín de los consumidores, ODEPA, 2009.
- [5] Steven Berry, James Levinsohn, and Ariel Pakes. Automobile prices in market equilibrium. *Econometrica*, 63(4):841–890, 1995.
- [6] Steven T. Berry. Estimating discrete-choice models of product differentiation. *The RAND Journal of Economics*, 25(2):242–262, 1994.
- [7] M. Bosch, R. Montoya, and R. Inostroza. Estudio de los Efectos de la Introducción de un Producto de Marca Propia en una Cadena de Retail. Technical report, Departamento de Ingeniería Industrial, Universidad de Chile, 2001.
- [8] Karla Cordero L. and Fernando Saavedra B. Manual de Procedimientos para la Medición de la Calidad Industrial del Arroz en Chile. Boletín N 230, INIA. Centro Regional de Investigación Quilamapu, Chillán, Chile, 2011.
- [9] Peter Davis. Spatial competition in retail markets: Movie theaters. *The RAND Journal of Economics*, 37(4):964–982, 2006.
- [10] Álvaro Espinoza and Claudio Farías. La cadena del arroz en Chile. Artículos, Unidad de Transparencia de Mercados, ODEPA, 11 2017.
- [11] FAO. Rice Market Monitor. Volume XX 3, FAO, Trade and Markets Division, Food and Agriculture Organization of the United Nations, October 2017.
- [12] Jeremy T. Fox, Kyoo il Kim, Stephen P. Ryan, and Patrick Bajari. A simple estimator for the distribution of random coefficients. *Quantitative Economics*, 2(3):381–418, 2011.

- [13] Jeremy T. Fox, Kyoo il Kim, and Chenyu Yang. A simple nonparametric approach to estimating the distribution of random coefficients in structural models. *Journal of Econometrics*, 195(2):236 – 254, 2016.
- [14] Jerry A. Hausman. Valuation of new goods under perfect and imperfect competition. In Timothy F. Bresnahan and Robert J. Gordon, editors, *The Economics of New Goods*, NBER Book Series Studies in Income and Wealth, pages 207–248. University of Chicago Press, January 1996.
- [15] Santiago Hernaíz L. and José Roberto Alvarado A. Calidad Industrial del Arroz: Un factor importante en la modernización del cultivo. Technical report, Centro Regional de Investigación Quilamapu, INIA, Chillán, Chile., 2002.
- [16] Bienvenido O. Juliano and Donald. B. Bechtel. The Rice Grain and its Gross Composition. In Bienvenido O. Juliano, editor, *Rice Chemistry and Technology*, pages 17–57. American Association of Cereal Chemists, St. Paul, MN, USA, second edition, 1985.
- [17] Wagner A. Kamakura. Estimating flexible distributions of ideal points with external analysis of preferences. *Psychometrika*, 56(3):419–431, September 1991.
- [18] Nicolás Lara T. Inserción y crecimiento de las marcas propias en el mercado del arroz. Un análisis en forma reducida. Master’s thesis, Facultad de Ciencias Físicas y Matemáticas, Departamento de Ingeniería Industrial, Universidad de Chile, 2017.
- [19] D. L. McFadden. Conditional logit analysis of qualitative choice behavior. In P. Zarembka, editor, *Frontiers in Econometrics*, chapter 4, pages 105–142. Academic Press, New York, NY, 1973.
- [20] D. L. McFadden. Modelling the choice of residential location. In A. Karlqvist, L. Lundqvist, F. Snickars, and J. Weibull, editors, *Spatial Interaction Theory and Planning Models*, pages 75–96. North Holland, Amsterdam, 1978.
- [21] Mintel. Rice in Chile. Market Report, 2017.
- [22] Aviv Nevo. Mergers with differentiated products: The case of the ready-to-eat cereal industry. *The RAND Journal of Economics*, 31(3):395–421, 2000.
- [23] Aviv Nevo. A practitioner’s guide to estimation of random - coefficients logit models of demand. *Journal of Economics & Management Strategy*, 9(4):513–548, Winter 2000.
- [24] Aviv Nevo. Measuring market power in the ready-to-eat cereal industry. *Econometrica*, 69(2):307–342, 2001.
- [25] Carlos Noton and Andrés Elberg. Revealing bargaining power through actual wholesale prices. Working Papers 51, Facultad de Economía y Empresa, Universidad Diego Portales, April 2014.
- [26] Carlos Noton and Andrés Elberg. Are supermarkets squeezing small suppliers? Evidence from negotiated wholesale prices. *The Economic Journal*, 128(610):1304–1330, 2018.

- [27] Mario Paredes, Viviana Becerra, and Álvaro Vega. Variedades Tradicionales Desarrolladas y Recomendadas. Producción de Arroz: Buenas Prácticas Agrícolas (BPA), INIA, 2015.
- [28] Guillermo Scarlato. Trayectoria y Demandas Tecnológicas de las Cadenas Agroindustriales en el MERCOSUR Ampliado. Cereales: trigo, maíz y arroz. Serie Documentos 2, PROCISUD, BID, Montevideo, 2000.
- [29] B.D. Smith. *The Emergence of Agriculture*. Scientific American Library series. Scientific American Library, 1998.
- [30] K. Sudhir. Structural analysis of manufacturer pricing in the presence of a strategic retailer. *Marketing Science*, 20(3):244–264, 2018/06/22 2001.
- [31] TDLC. Resolución 43/2012. Consulta de SMU S.A. sobre los efectos en la libre competencia de la fusión de las sociedades SMU S.A. y Supermercados del Sur S.A., H. Tribunal de Defensa de la Libre Competencia, Santiago, Diciembre 2012.
- [32] Kenneth E. Train. EM Algorithms for nonparametric estimation of mixing distributions. *Journal of Choice Modelling*, 1(1):40–69, 2008.
- [33] USDA. Standards for Rice. Technical report, United States Department of Agriculture Grain Inspection, Packers and Stockyards Administration, Federal Grain Inspection Service, November 2009.
- [34] Sofia Berto Villas-Boas. Vertical relationships between manufacturers and retailers: Inference with limited data. *The Review of Economic Studies*, 74(2):625–652, 2007.