

## Asociación Argentina de Economía Agraria

# ¿Cómo se asocian los descriptores de etiqueta de vinos en diferentes segmentos de precios? Un análisis empírico para vinos argentinos en el mercado doméstico minorista

Rodrigo Garúa Arancibia \*

Instituto de Economía Aplicada Litoral (IECAL-FCE-UNL) & CONICET, Santa Fe, Argentina.  
(Contacto: rgarcia@fce.unl.edu.ar)

Mariano N. Coronel

Instituto de Economía Aplicada Litoral (IECAL-FCE-UNL), Santa Fe, Argentina.  
(Contacto: mcoronel@fce.unl.edu.ar)

Juliana Baissetto

Instituto de Economía Aplicada Litoral (IECAL-FCE-UNL), Santa Fe, Argentina.  
(Contacto: )

### Resumen

Los descriptores de etiqueta en vinos buscan mostrar ciertos atributos del bien que pueden ser muy relevantes tanto desde el punto de vista del productor como del consumidor. Por el lado del consumidor, aportan información sobre el vino, posibilitando disminuir los costos de búsqueda y facilitar el proceso de compra. Por el lado del productor, revelan diferentes estrategias de diferenciación vertical y horizontal que realizan las bodegas para competir en diferentes mercados. Por su parte, existen varios segmentos de vinos, donde los tipos de descriptores de etiqueta pueden cambiar sustancialmente. Teniendo en cuenta su importancia, el objetivo de este trabajo es conocer las relaciones de dependencia condicional entre los diferentes descriptores de etiqueta en función de los distintos segmentos de precios. Para ello se especifica un modelo Ising con el objeto de representar la elección de descriptores como función de los precios, estimándose un grafo condicional que cuantifica las relaciones de dependencia buscadas. El modelo se estima vía pseudo-máxima verosimilitud, penalizando a fin de detectar interacciones no relevantes. Los resultados muestran que la decisión de poner un cierto descriptor en la etiqueta de un vino depende del segmento del mercado, a la vez de que existe un claro patrón de asociación entre pares de descriptores de acuerdo a dichos segmentos.

Palabras Clave : \* Información de etiqueta \* Dependencia Condicional \* Modelo Ising \* Vinos de Precios Bajos \* Vinos de Precios Altos.

---

\* Este trabajo se enmarca dentro del proyecto CAI+D 2016 (Cod. 50020150100094LI ) de la Universidad Nacional del Litoral.

## ¿Cómo se asocian los descriptores de etiqueta de vinos en diferentes segmentos de precios? Un análisis empírico para vinos argentinos en el mercado doméstico minorista

### Resumen

Los descriptores de etiqueta en vinos buscan mostrar ciertos atributos del bien que pueden ser muy relevantes tanto desde el punto de vista del productor como del consumidor. Por el lado del consumidor, aportan información sobre el vino, posibilitando disminuir los costos de búsqueda y facilitar el proceso de compra. Por el lado del productor, revelan diferentes estrategias de diferenciación vertical y horizontal que realizan las bodegas para competir en diferentes mercados. Por su parte, existen varios segmentos de vinos, donde los tipos de descriptores de etiqueta pueden cambiar sustancialmente. Teniendo en cuenta su importancia, el objetivo de este trabajo es conocer las relaciones de dependencia condicional entre los diferentes descriptores de etiqueta en función de los distintos segmentos de precios. Para ello se especifica un modelo Ising con el objeto de representar la elección de descriptores como función de los precios, estimándose un grafo condicional que cuantifica las relaciones de dependencia buscadas. El modelo se estima vía pseudo-máxima verosimilitud, penalizando a fin de detectar interacciones no relevantes. Los resultados muestran que la decisión de poner un cierto descriptor en la etiqueta de un vino depende del segmento del mercado, a la vez de que existe un claro patrón de asociación entre pares de descriptores de acuerdo a dichos segmentos.

**Palabras Clave:**\* Información de etiqueta \*Dependencia Condicional \* Modelo Ising \*Vinos de Precios Bajos \*Vinos de Precios Altos.

# How are wine label descriptors associated in different price segments? An empirical analysis for Argentine wines in the domestic retail market

## Abstract

Wine label descriptors seek to show certain attributes of the good that can be very relevant both from the point of view of the producer and the consumer. On the consumer side, they provide information about wine, making it possible to reduce search costs and facilitate the purchase process. On the producer side, they reveal different vertical and horizontal differentiation strategies carried out by the wineries to compete in different markets. On the other hand, there are several wine segments, where the types of label descriptors can change substantially. Taking into account its importance, the objective of this work is to know the relations of conditional dependence between the different label descriptors according to the different price segments. For this, an Ising model is specified in order to represent the choice of descriptors as a function of prices, estimating a conditional graph that quantifies the dependency relationships sought. The model is estimated via pseudo-maximum likelihood, penalizing in order to detect non-relevant interactions. The results show that the decision to put a certain descriptor on the label of a wine depends on the market segment, while there is a clear pattern of association between pairs of descriptors according to these segments.

**Keywords:** \* Label information \*Conditional Dependency \* Ising Model \*Higher Priced wines \*Cheaper Priced Wines.

## 1. Introducción

El vino es un producto con un alto grado de diferenciación vertical y horizontal, y en muchos mercados, como el de Argentina, la oferta está atomizada en una gran cantidad de marcas para los diferentes segmentos de precios. Por ello, las bodegas deben competir en cada uno de estos submercados con estrategias de promoción y marketing; y el etiquetado, tanto en su diseño visual como en la información contenida, constituye una de las estrategias más relevantes para atraer diferentes tipos de consumidores (Thomas & Pickering 2003, Lockshin & Corsi 2012, Corduas, Cinquanta & Ievoli 2013).

Los descriptores de etiqueta en vinos buscan mostrar ciertos atributos del bien que pueden ser muy relevantes tanto desde el punto de vista del productor como del consumidor. Por el lado del consumidor, aportan información sobre el vino, posibilitando disminuir los costos de búsqueda y facilitar el proceso de compra (Thomas & Pickering 2003). Por el lado del productor, revelan diferentes estrategias de diferenciación vertical y horizontal que realizan las bodegas para competir en diferentes mercados (Saak 2011, Mueller Loose, Lockshin, Saltman & Blanford 2010). Siguiendo a Garcia et al. (2015), los descriptores de etiqueta pueden dividirse en aquellos que informan sobre características objetivas del vino, como ser la graduación alcohólica, el varietal, la región, la cosecha, el potencial de guarda y su paso por roble, entre otros; y aquellos descriptores que aportan una información adicional (o no necesariamente objetiva) como ser las recomendaciones de temperatura y el maridaje, las características sensoriales, la descripción del proceso de elaboración, entre otros. Los estudios económicos sobre las etiquetas, por lo general, están centrados en la evaluación que realizan los consumidores de algunos de estos descriptores. Su relevancia está dada por la información que aportan tanto para el diseño de estrategias competitivas por parte de las firmas como para el análisis de bienestar de los consumidores.

Existe una vasta literatura que estudia el rol de los diferentes descriptores de etiqueta tanto en las decisiones de compra como en los precios del mercado (Barber, Almanza & Donovan 2006, Charters, Lockshin & Unwin 1999, Mueller Loose, Lockshin, Louviere, Francis & Osidacz Williamson 2009, Lockshin, Mueller Loose, Louviere, Francis & Osidacz Williamson 2009, Mueller, Lockshin, Saltman & Blanford 2010, Mora & Livat 2013). Por ejemplo, Barber et al. (2006) evaluaron la influencia de los envases de vino y de los descriptores de etiquetas en la decisión de compra de los consumidores, incorporando descriptores de sabor, historia de la región del vino y proceso de elaboración del vino, entre otros. Charters et al. (1999) estudiaron las respuestas de los consumidores australianos, descubriendo que los descriptores sensoriales eran los atributos más valorados entre otra información contenida en la etiqueta. Mueller et al. (2009) muestran que el contenido de la etiqueta podría ser crucial para el consumidor en el proceso de elección del vino en góndola, debido a que más información reduce el riesgo de compra percibido por el comprador. Mueller et al. (2010) examinaron qué descriptores contenidos en la parte posterior de la etiqueta de un vino son los más valorados para los consumidores australianos mediante un experimento de elección discreta. Incluyeron descriptores de sabor, declaraciones relacionadas con la fabricación o elaboración, la historia del vino o de la bodega, consejos de temperatura de consumo, maridaje, ingredientes, características ambientales y el sitio web. Encontraron que la presencia de estos descriptores de información adicional resultan ser muy importantes para explicar la probabilidad de elegir un vino. Por su parte, Mora y Livat (2013) estudiaron el papel que tiene una narración en torno a los viñedos como un medio de comunicación corporativa. Pero esta información se proporcionó en una guía, no en la etiqueta. Según Mora y Livat (2013) en productos

orientados a la experiencia, como el vino, la imagen y la reputación están involucrados en las decisiones de fijación de precios, por lo que la información adicional podría explicar las diferencias de precios. La narración de historias permite al consumidor integrar la historia de una marca de vino o una finca vinícola. Otros trabajos muestran la relevancia en la elección y disposición a pagar que tiene los descriptores relacionados con las formas de producción sustentables (Schäufele & Hamm 2017). Lockshin et al. (2009), a su vez, descubrieron que era posible identificar diferentes segmentos de consumidores impulsados por atributos particulares del vino especificados en las etiquetas. Por lo tanto, se esperaba que no todos los consumidores reaccionaran de la misma manera a los cambios en el etiquetado. Mientras que tal heterogeneidad de preferencia podría observarse a nivel individual, sería posible que estas diferencias se cancelen entre sí en el mercado total. De esta manera, podría decirse que no es posible saber si el impacto previamente observado del etiquetado del vino en las preferencias individuales de los consumidores se tradujera en diferentes precios de mercado (Mueller Loose & Szolnoki 2010).

Para el caso del mercado de vinos Argentinos, existen investigaciones que estudian el rol de estos descriptores sobre los precios de mercado, tanto en góndolas del mercado doméstico (García Arancibia, Rossini & Depetris Guiguet 2015) como en el mercado on-line (García Arancibia, Guiguet & Rossini 2016). En ambos estudios se encuentra que la información contenida en los descriptores juegan un papel relevante en los precios de los vinos. Además, se encuentran algunas particularidades de los precios implícitos de los descriptores de etiqueta, según distintos segmentos de los vinos. El presente trabajo busca complementar estas investigaciones, pero en lugar de analizar el rol de los descriptores de etiqueta sobre el precio, se busca caracterizar las relaciones de dependencia o asociación entre dichos descriptores en función de los diferentes segmentos de mercado que definen ciertos rangos de precios. Específicamente se propone un modelo donde la variable respuesta es un vector de descriptores, asociados en pares, cuya relaciones de aparición y dependencia dependen de los precios de los vinos, o bien de los segmentos definidos a partir de estos. Con la estimación de este modelo se busca conocer cómo el mercado configura un cierto conjunto de información contenido en la etiqueta, y cómo ello depende del segmento que se mire. En base a los trabajos anteriores sobre el impacto de los descriptores en los precios, la hipótesis principal del presente trabajo es que existen claros patrones en las relaciones de dependencia entre descriptores y, a su vez, estas varían de acuerdo al segmento de precios de los vinos.

Lo que resta del presente trabajo se estructura de la siguiente manera: En la Proxima sección se presenta el modelo gráfico propuesto, el procedimiento de estimación, los datos utilizados y las variables consideradas. Luego se presentan los resultados, finalizando el trabajo con unas breves conclusiones.

## 2. Metodología

### 2.1. Modelo y Estimación

Vamos a suponer que en una situación de equilibrio de mercado, los descriptores de etiqueta con información adicional del vino configuran una red de asociación o dependencia entre los mismos. Al modelar la presencia/ausencia de cada uno de los  $q$  descriptores, se tiene que un vector de variables dicotómicas  $\mathbf{D}$  sintetiza la información contenida en la etiqueta de un cierto vino. Dada esta característica de  $\mathbf{D}$  es natural suponer una distribución Bernoulli multivariada para la misma. Debido a que esta distribución involucra todas las interacciones de segundo orden y superiores entre las variables,

para obtener una representación más parsimoniosa de la distribución conjunta de variables Bernoulli es usual tomar el caso especial del denominado modelo Ising (Dai, Ding & Wahba 2013, Dai 2013, Cheng, Levina, Wang & Zhu 2014), el que incorpora sólo interacciones de a pares. A su vez suponemos que la dependencia entre las variables contenidas en  $\mathbf{D}$  varía de acuerdo al segmento de mercado de los vinos; en particular, diremos que es función de los diferentes segmentos de precios definidos a partir de una variable  $Y$  que puede ser el precio mismo del vino o de varias variables categóricas que indicarían el segmento específico. Por ello, supondremos que  $Y \in \mathbb{R}^r$  con  $r \geq 1$ . Siguiendo a Cheng et al. (2014), condicionando en  $Y$ , el vector de las  $q$  variables binarias  $\mathbf{D}^T = (D_1, D_2, \dots, D_q)$  tienen la siguiente función de densidad conjunta

$$\begin{aligned} P(\mathbf{D}|Y = y) &= P(D_1, \dots, D_q|Y = y) \\ &= \frac{1}{G(\Theta^y)} \exp \left( \sum_{j=1}^q \theta_{jj}^y D_j + \sum_{1 \leq j < j' \leq q} \theta_{jj'}^y D_j D_{j'} \right), \end{aligned} \quad (1)$$

donde los  $\theta$ 's son los parámetros del modelo que dependen del valor asumido por las covariables  $Y$  y  $G(\Theta^y)$  asegura que las  $2^q$  probabilidades sumen 1. Para  $j, j' = 1, \dots, q$  se tiene que

$$\begin{aligned} \theta_{jj}^y &= \log \left( \frac{P(D_j = 1 | \mathbf{D}_{-j} = 0, y)}{1 - P(D_j = 1 | \mathbf{D}_{-j} = 0, y)} \right), \\ \theta_{jj'}^y &= \log \frac{P(D_j = 1, D_{j'} = 1 | \mathbf{D}_{-j, -j'} = 0, y) P(D_j = 0, D_{j'} = 0 | \mathbf{D}_{-j, -j'} = 0, y)}{P(D_j = 1, D_{j'} = 0 | \mathbf{D}_{-j, -j'} = 0, y) P(D_j = 0, D_{j'} = 1 | \mathbf{D}_{-j, -j'} = 0, y)}, \end{aligned}$$

$$G(\Theta^y) = \sum_{j=1}^q \sum_{\{D_j=0,1\}} \exp \left( \sum_{j=1}^q \theta_{jj}^y D_j + \sum_{1 \leq j < j' \leq q} \theta_{jj'}^y D_j D_{j'} \right),$$

con  $\mathbf{D}_{-j} = (D_1, \dots, D_{j-1}, D_{j+1}, \dots, D_q)$ ,  $\mathbf{D}_{-j, -j'} = (D_1, \dots, D_{j-1}, D_{j+1}, \dots, D_{j'-1}, D_{j'+1}, \dots, D_q)$  y  $\Theta^y$  es una matriz simétrica cuyo elementos vienen dados por  $[\Theta^y]_{ij} = \theta_{ij}^y$ . Siguiendo a (Cheng, Li, Levina & Zhu 2016), cada  $\theta_{ij}^y$  es modelado como función lineal de la función de  $Y$ . Específicamente, supondremos que

$$\theta_{ij}^y = \theta_{ij,0}^* + \theta_{ij}^T Y, \quad i, j = 1, \dots, q \quad (2)$$

donde  $\theta_{ij}^T = (\theta_{ij,1}, \dots, \theta_{ij,r})$  es un vector de parámetros (independientes de  $Y$ ) y  $\theta_{ij,0}^*$  es el intercepto para cada  $(i, j)$ .

Usando esta parametrización (2), la distribución conjunta para el modelo Ising (1) puede escribirse de

la forma

$$P(\mathbf{D}|y) = \exp \left( \sum_{j=1}^q \theta_{jj0}^* D_j + \sum_{j=1}^q \boldsymbol{\theta}_{jj}^T y D_j + \sum_{1 \leq j < j' \leq q} \theta_{jj'0}^* D_j D_{j'} + \sum_{1 \leq j < j' \leq q} \boldsymbol{\theta}_{jj'}^T y D_j D_{j'} \right) \frac{1}{G(\boldsymbol{\Theta}_y)} \quad (3)$$

Si consideramos una variable dicotómica particular  $j$  y condicionamos sobre las restantes ( $\mathbf{D}_{-j} \equiv (D_1, \dots, D_{j-1}, D_{j+1}, \dots, D_q)$ ), obtenemos

$$\log \frac{P(D_j = 1 | \mathbf{D}_{-j}, Y)}{P(D_j = 0 | \mathbf{D}_{-j}, Y)} = \theta_{jj0}^* + \boldsymbol{\theta}_{jj}^T Y + \sum_{j \neq j'} \theta_{jj'0}^* D_{j'} + \sum_{j < j'} \boldsymbol{\theta}_{jj'}^T Y D_{j'} \quad (4)$$

De esta manera se obtiene que el log-odds condicional de una determinada variable dicotómica  $H_j$  es lineal en los parámetros, de forma tal que los estimadores máximos verosímiles (condicionados) de dichos parámetros pueden obtenerse mediante una regresión logística tomando  $D_j$  como respuesta con los predictores  $(Y, \mathbf{D}_{-j}, Y \mathbf{D}_{-j})$ . De esta manera, ajustando  $q$  modelos logísticos univariados se obtienen estimadores para  $\boldsymbol{\theta}_0$  y  $\boldsymbol{\theta}$ . En particular, considerando una muestra de  $V$  vinos en el mercado, con siguientes los puntos muestrales  $(\mathbf{d}_v, y_v) \equiv (d_{v1}, \dots, d_{vq}, y_v)$  con  $v = 1, \dots, V$ , se tiene que para cada descriptor  $j$ , con  $j = 1, \dots, q$ , la función log-verosimilitud, vendrá dada por

$$\ell_j(\boldsymbol{\theta}_0, \boldsymbol{\theta}; \mathbf{d}_v, y_v) = \frac{1}{V} \sum_{v=1}^V \log P(d_{vj} | \mathbf{d}_{v,-j}, y_v) = \frac{1}{V} \sum_{v=1}^V (d_{vj} \epsilon_{vj} - \log(1 + \exp(\epsilon_{vj}))) \quad (5)$$

con

$$\epsilon_{vj} = \log \frac{P(d_{vj} = 1 | \mathbf{d}_{v,-j}, y_v)}{P(d_{vj} = 0 | \mathbf{d}_{v,-j}, y_v)} = \theta_{jj0}^* + \boldsymbol{\theta}_{jj}^T y_v + \sum_{j \neq j'} \theta_{jj'0}^* d_{j'} + \sum_{j \neq j'} \boldsymbol{\theta}_{jj'}^T y_v d_{j'}.$$

De esta manera puede obtenerse un estimador para  $(\boldsymbol{\theta}_0, \boldsymbol{\theta})$  de forma tal que para cada  $j$  se maximicen (5). Por simetría,  $\theta_{jj'}^y = \theta_{j'j}^y$ , por lo que para tal parámetro tenemos dos estimadores, i.e. el de la regresión logística de  $D_j$  sobre  $(Y, \mathbf{D}_{-j}, Y \mathbf{D}_{-j})$  y el de  $D_{j'}$  sobre  $(Y, \mathbf{D}_{-j'}, Y \mathbf{D}_{-j'})$ , pues  $D_{j'} \in \mathbf{D}_{-j}$  y  $D_j \in \mathbf{D}_{-j'}$ , y como nada nos garantiza que de la estimación obtengamos  $\theta_{jj'0}^* = \theta_{j'j0}^*$  y  $\boldsymbol{\theta}_{jj'} = \boldsymbol{\theta}_{j'j}$ , debemos seleccionar algún criterio para que se cumpla la simetría (Cheng et al. 2014).

Adicionalmente, Cheng et al. (2014) proponen dos algoritmos de selección de variables para el modelo Ising con covariables. Específicamente, utilizando regularización con la norma  $\ell_1$ , proponen dos métodos para maximizar (5). En ambos sólo se penaliza  $\boldsymbol{\theta}$ , pero uno lo hace de forma separada para cada  $j$ , mientras que el otro lo hace de manera conjunta. La regularización para regresiones logísticas de forma separada la realiza vía

$$\min_{\boldsymbol{\theta}_j \in \mathbb{R}^{(r+1)q}} -\ell_j(\boldsymbol{\theta}_0, \boldsymbol{\theta}; \mathbf{d}_v, y_v) + \lambda \|\boldsymbol{\theta}_{j \setminus 0}\|_1, \quad (6)$$

donde  $\boldsymbol{\theta}_j \doteq (\theta_{j10}, \boldsymbol{\theta}_{j1}^T, \dots, \theta_{jq0}, \boldsymbol{\theta}_{jq}^T)$  y  $\boldsymbol{\theta}_{j \setminus 0} \doteq \boldsymbol{\theta}_j \setminus \{\theta_{jj0}\}$ . Para imponer la simetría, la propuesta de (Cheng et al. 2014), en base a (Meinshausen & Bühlmann 2006), sugiere la comparación de las

magnitudes estimadas (pre-regularización) para  $\theta_{jkl}$  y  $\theta_{kjl}$  (con  $j, k = 1, \dots, q$  y  $l = 1, \dots, r$ ), eligiendo el mínimo o el máximo de ellos, dependiendo si se busca un método más o menos conservativo (mayor o menor cantidad de ceros estimados).

Por otra parte, la regularización conjunta termina estimando  $(\boldsymbol{\theta}_0, \boldsymbol{\theta})$  de forma conjunta de la siguiente manera

$$\min_{(\boldsymbol{\theta}_0, \boldsymbol{\theta})} \sum_{j=1}^q -\ell_j(\boldsymbol{\theta}_0, \boldsymbol{\theta}; \mathbf{d}_j, y_j) + \lambda \|\boldsymbol{\theta}\|_1. \quad (7)$$

Una ventaja de esta regularización conjunta es que la simetrización se impone automáticamente cuando se resuelve (7), sin embargo el costo computacional de este método es mayor. En el presente trabajo, al no tener un  $\mathbf{D}$  de dimensiones muy elevadas ( $q = 5$ ), se utiliza este último método, usando el código para Matlab provisto por los autores<sup>1</sup>. Estimados los parámetros con este código, se utiliza el paquete 'qgraph' del software R para graficar los grafos de dependencia condicional.

Para la selección del parámetro de penalización  $\lambda$  se realiza un experimento de 10-fold cross-validation, seleccionado de acuerdo al poder predictivo del mismo en cada submuestra de validación.

## 2.2. Variables y Datos

Los datos utilizados son los usados en el trabajo de García et al. (2015). Estos provienen de una muestra de 1.015 vinos varietales producidos por 49 bodegas argentinas, en el segmento de botellas de vidrio de 750 mililitros (ml) . Se recolectaron en tres supermercados, tres hipermercados y dos tiendas especializadas de vino en la ciudad de Santa Fe, todas ellas ofreciendo una amplia variedad de vinos en sus estantes. La recolección ocurrió en abril de 2011 en supermercados y tiendas especializadas, y en mayo en hipermercados. El “vino” incluido corresponde al producto definido por el Artículo 17 de la Ley Nacional del Vino No. 14.878, legalmente considerado vino de uva genuino. Cada observación contenía variables que caracterizan el tipo de vino (color, año de cosecha, variedad de uva, graduación alcohólica, etc.), la región de origen e *información adicional* incluida en la etiqueta (García Arancibia et al. 2015).

Sobre la base de dicha *información adicional*, vamos a considerar 5 descriptores de etiqueta definidos a partir de las siguientes variables binarias: 1) *Maridaje*, que vale 1 cuando la etiqueta presenta recomendaciones sobre platos de comidas que acompañarían adecuadamente le vino, y 0 en caso contrario. 2) *Descripción Sensorial*: vale 1 cuando existe información enológica sensorial, tal como el aspecto, color, aroma, entre otras; 3) *Temperatura*: indica si la etiqueta contiene recomendación sobre la temperatura que debe servirse el vino; 4) *Elaboración*: con valor igual a 1 si en la etiqueta contienen alguna descripción del proceso de elaboración del vino; y 5) *Medioambiente*: con valor 1 si la etiqueta presenta alguna información geológica o ambiental de la producción de los viñedos utilizados para la fabricación del vino, tal como el tipo de suelo, clima, altitud u otras características geológicas. De esta manera, en términos de la notación utilizada para el modelo se tiene  $q = 5$ , con  $\mathbf{D}^T = (D_1, \dots, D_5)$ , donde  $D_1 = \textit{Maridaje}$ ,  $D_2 = \textit{Sensorial}$ ,  $D_3 = \textit{Temperatura}$ ,  $T_4 = \textit{Elaboracion}$  y  $D_5 = \textit{Medioambiente}$ .

<sup>1</sup><http://onlinelibrary.wiley.com/doi/10.1111/biom.12202/>

Puesto que una de las hipótesis del trabajo es que las relaciones de dependencias varían de acuerdo al segmento de precios de los vinos, como co-variable  $Y$  del modelo gráfico consideramos, en primer lugar, el precio del vino como predictor continuo. Luego se incluye a  $Y$  como categórica del segmento de precios que representa, siguiendo a García et al. (2015). Específicamente, se definieron dos segmentos según el precio: a) segmento de vinos varietales con precios inferiores a 6.14 dólares por botella, nombrados aquí con vinos de *Precios Bajos*; b) mayores 6.14 dólares por botella, de aquí en adelante vinos de *Precios Altos*, el que comprende las siguientes subcategorías: b.1) “Premium” (entre 6.14 y 8.6 dólares; b.2) “Super Premiums” (De US\$ 8.6 a US \$ 12.3 por botella); b.3) “Ultra Premium” (de US\$ 12.3 a US \$ 18.5); b.4) “Icon” (de US \$ 12.3 a US \$ 30.7 inclusive) y finalmente b.5) “Uber” (mayor a US \$ 30.7) (Área del Vino, 2011). Todos los precios (en pesos argentinos) fueron deflactados a abril de 2011 utilizando el Índice de Precios al Consumidor de Santa Fe publicado por la Unidad de Estadística de Santa Fe (IPEC) y luego transformado a dólares estadounidenses<sup>2</sup>. El precio promedio es de 6.4 dólares por botella, con una desviación estándar de 3.9 dólares. Además del precio como co-variable, se incluye si una *dummy* que indica si el vino pasó o no por roble, a los efectos de contrastar si existe algún efecto roble que segmente el mercado por este tipo de diferenciación.

### 3. Resultados

Antes de analizar las relaciones de dependencia estimadas entre los 5 descriptores de etiqueta, el Cuadro 1 muestra la distribución de frecuencia de los descriptores para la muestra total de vinos y diferenciando según los dos segmentos de precios definidos anteriormente. En primer lugar se puede observar que la descripción sensorial se encuentra en la mayoría de los vinos, tanto para la muestra total como en cada uno de los segmentos. Al comparar la frecuencia por segmento, se observa que las recomendaciones de maridaje y temperatura predominan entre los vinos de precios bajos, mientras que la información sobre el proceso de elaboración del vino y las características medioambientales tienen una mayor frecuencia en los vinos de precios más altos. Por ende, esta simple medida de frecuencia relativa da información relevante sobre las características de cada segmento: En el de segmento de precios bajos, la información contenida en la etiqueta está más orientada a recomendar al consumidor para una buena experimentación del vino, mientras que en el segmento de precios altos, la información está más especializada en cuestiones técnicas (manufacturación y características de los viñedos), que podrían ser mejor apreciadas por consumidores más experimentados o especializados. CITAAAA.

Previamente a controlar por los segmentos de precios, en la Figura 1 se muestra el grafo de dependencia entre los descriptores (sintetizado en el  $\theta_0$ ). El grosor de las aristas muestra el grado de asociación, mientras que el color indica si esta es negativa (en rojo) o positiva (en verde). Se puede observar una alta dependencia positiva entre la descripción sensorial y la recomendación de temperatura. Es decir, si está uno de ellos en la etiqueta, con alta probabilidad estará el otro descriptor. Por el contrario, se observa una alta dependencia negativa entre la descripción de la tierra donde se ubican los viñedos (denominada *Medioambiente*) con el maridaje y la descripción sensorial. Es decir que cuando en una botella, la etiqueta contiene la descripción medioambiental, es poco probable que se encuentre la del maridaje o la sensorial. Entre estas dos, la dependencia es positiva. Otras dependencia positiva se observan entre la recomendación de temperatura y el maridaje. Por su parte, entre el medioambiente

<sup>2</sup>El Tipo de Cambio a Abril de 2011 fue de 1 U\$S = 4.07 pesos Argentinos.

Cuadro 1: Frecuencia muestral (en %) de los Descriptores de Etiqueta según Segmentos de Precios de Vinos

	Muestra Total	Segmentos	
		<i>Precios Bajos</i>	<i>Precios Altos</i>
Maridaje	41	50.8	23
Sensorial	83	89.5	70.2
Temperatura	57	67.5	37.7
Elaboración	36	34	40.7
Medioambiente	17	14	21.3

y la temperatura, la asociación es positiva. Respecto a la información sobre la elaboración del vino, se observa poca asociación con el resto de los descriptores, mostrando ser negativa con la de medioambiente y con la de temperatura, y levemente positiva con el maridaje.

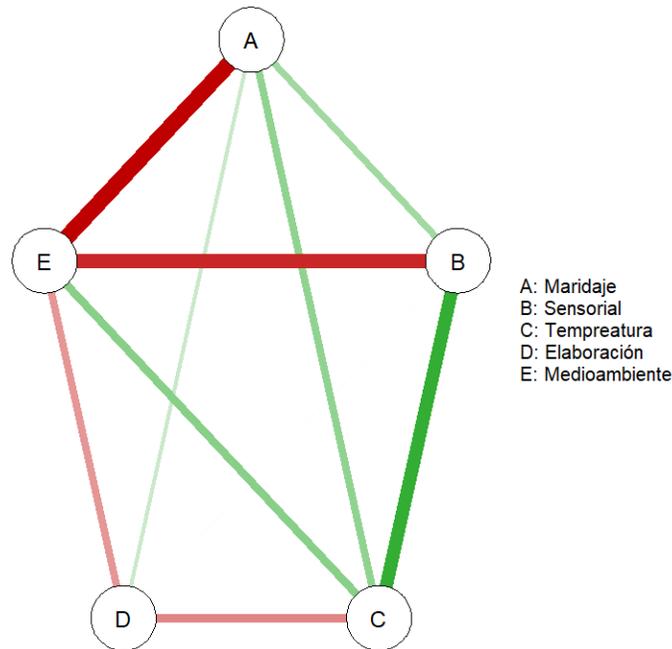


Figura 1: Grafo de Dependencia de Descriptores de Etiqueta

Cuando se incorpora la co-variable  $Y$  binaria de los segmentos de *Precios Bajos* y de *Precios Altos*, las relaciones de dependencia cambian como se muestra en la figura 2. En primer lugar se observa que al controlar por segmento, la asociación entre el maridaje y la descripción medioambiental desaparece (i.e. el método de regularización fijó en cero esta interacción doble), mientras que otras asociaciones se

fortalecen. Ambos grafos son complementarios en los colores puesto que en el (a) se toma a los vinos de precios altos como base, mientras que en el (b) se toman a los vinos de precios bajos como base. Se presentan ambos dado que dan una mejor descripción de las relaciones de dependencia en cada segmento. En el segmento de precios bajos se observa que si en la etiqueta se pone información del maridaje, es muy probable que también contenga información sensorial, de la temperatura y del proceso de elaboración. Por su parte, cuando existe información sobre el medioambiente, es muy probable que también contenga información sobre la elaboración. Por otra parte, para el segmento de precios altos se observa que cuando existe recomendación de temperatura es muy probable que contenga información de la elaboración y de la descripción sensorial. La mayor asociación negativa con los otros descriptores muestra que en general en el segmento de vinos de precios altos, las etiquetas están menos cargadas de información, y en su mayoría funcionan de forma sustituta, mientras que en el de vinos bajos existe una mayor complementación entre las mismas.

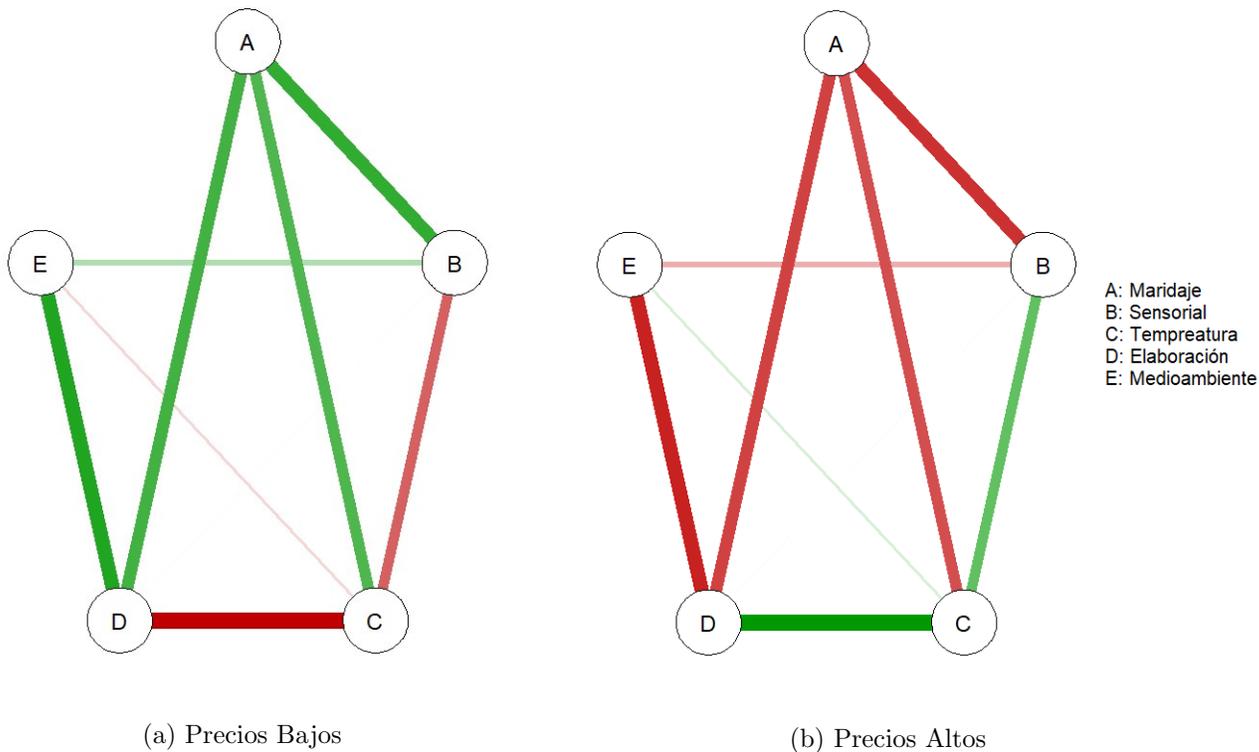


Figura 2: Dependencia Condicional de Descriptores de Etiquetas en dos Segmentos de Precios

El efecto del controlar por el precio en las relaciones de dependencia entre descriptores de etiqueta también se observa claramente en la Figura 3, donde el precio entra como co-variable continua. Cuando condicionamos al precio, varias relaciones de dependencia desaparecen como ser entre el medioambiente con el maridaje, la descripción sensorial y la temperatura. El grafo es similar al que mostramos en la Figura 2 (b) cuando condicionamos al segmento de precios altos, pero en este caso las aristas más sua-

ves desaparecen, mostrando nula asociación entre esos descriptores. Nuevamente, a precios mayores, los únicos descriptores que parecen complementarse son el de elaboración con el de la recomendación de temperatura. De esta manera, en el segmento de vinos de *Precios Altos* un descriptor técnico como el del proceso de elaboración (e.g. maceración y crianza del vino) que podría considerarse como un atributo de *búsqueda*, es más probable que esté acompañado de sólo un atributo de *experiencia* del consumidor, cómo lo es la temperatura.

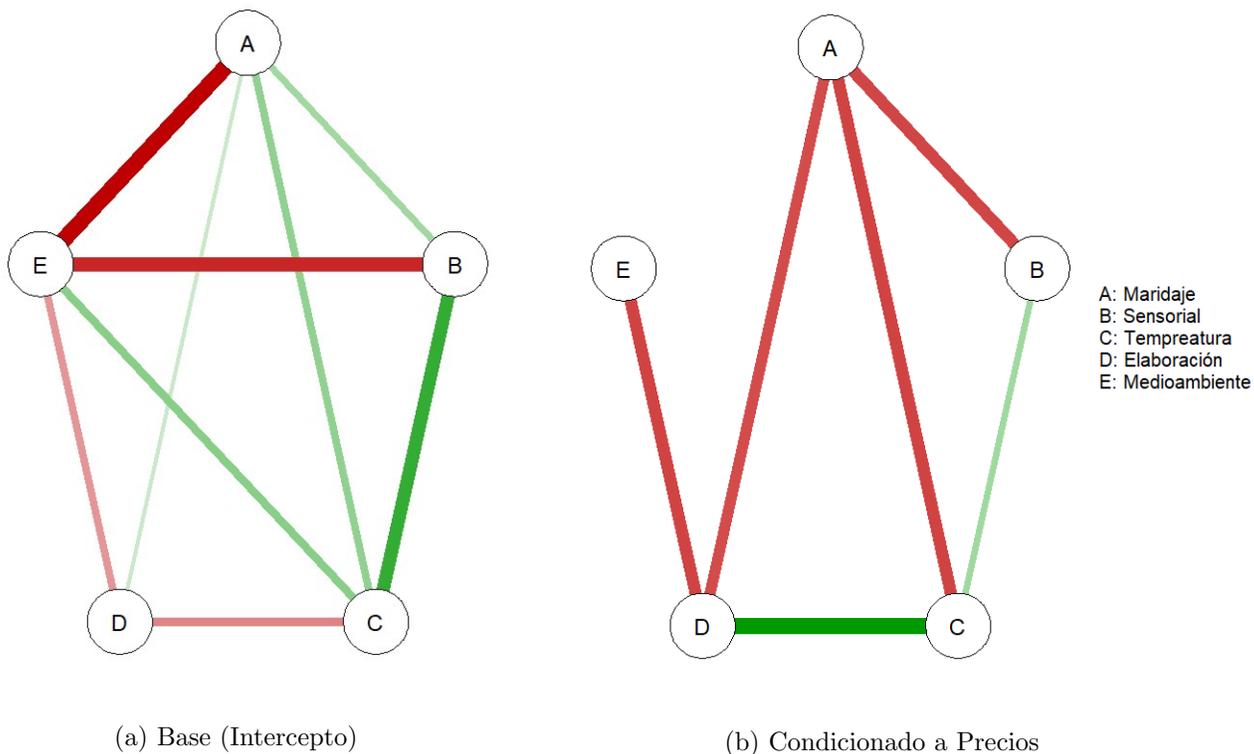


Figura 3: Grafo de Dependencia Condicional de Descriptores de Etiquetas. Efecto de considerar los precios (como covariable continua)

En el Cuadro 2 se presentan los coeficientes  $\theta$  estimados para ambos modelos (i.e. con  $Y$  discreta y continua). Para el caso de  $Y$  discreta, los resultados reportados toman al segmento de *Precios Bajos* como base. Analizando la diagonal (que no se refleja en los grafos), se observa que a pesar de que en el segmento de vinos de precio alto hay una menor proporción de vinos con maridaje en la etiqueta (en relación a los de precios bajos), los resultados muestran que si no están los otros descriptores, la chance de que este el maridaje es un 15 por ciento superior respecto al segmento de vinos más baratos (que tampoco tienen los otros descriptores). Para la característica sensorial, cuya presencia en las etiquetas es alta (más del 70 por ciento de los vinos) para ambos segmentos, se observa de que la chance de que aparezca solamente este descriptor es menor en el segmento de vinos de precios altos. Algo similar ocurre con la recomendación de temperatura, que es más frecuente en el segmento

Cuadro 2: Log Odds Ratios estimados de los Modelos gráficos con respuesta continua (precio) y discreta (categoría del segmento de vinos)

		$\theta_{ij}(Y \text{ discreta}; \hat{\lambda}=0.015)$				
$(i, j)$	Maridaje	Sensorial	Temperatura	Elaboración	Medioambiente	
Maridaje	0.154	-0.439	-0.377	-0.402	0	
Sensorial		-0.271	0.336	0.012	-0.172	
Temperatura			-0.757	0.543	0.09	
Elaboración				0	-0.473	
Medioambiente					0.195	
		$\theta_{ij}(Y \text{ continua}; \hat{\lambda}=0.20)$				
$(i, j)$	Maridaje	Sensorial	Temperatura	Elaboración	Medioambiente	
Maridaje	0	-0.331	-0.335	-0.319	0	
Sensorial		-0.254	0.169	0	-0.005	
Temperatura			-0.573	0.455	0	
Elaboración				0	-0.335	
Medioambiente					0.049	

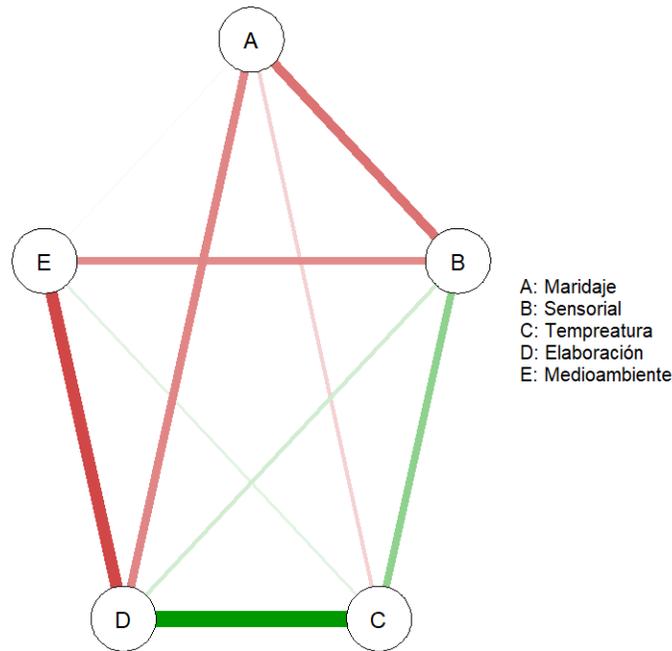


Figura 4: Grafo de Dependencia de Descriptores de Etiquetas condicionado al Segmento de Vinos de Precios Altos con Roble

de vinos de precios bajos. Para la descripción de elaboraciónse tiene que las chances de que aparezca

sola no cambia por segmento del vino. Para el caso de la descripción del medioambiente, que tiene mayor presencia en las etiquetas de los vinos de precio alto, la chance de aparecer, dada la ausencia de los otros descriptores, es casi un 20 por ciento superior en dicho segmento. Cuando se considera a  $Y$  continua, los signos de los (logaritmos) de los cociente de chance, como es de esperarse, coinciden en ambos modelos, solamente que cuando se toman los precios como co-variable continua, el método regularizado da más ceros (con un parámetro  $\lambda$  de regularización óptimo que es levemente mayor), confirmando mas relaciones de independencia y haciendo con ello más fácil la interpretación de los resultados. En este caso se tiene que las chances de la aparición sola del maridaje no depende de los precios, y además de mostrar independencia entre los descriptores de medioambiente y maridaje, también lo hace entre el medioambiente y temperatura, y entre la descripción sensorial y de elaboración.

Por último nos preguntamos qué ocurre cuando incorporamos como co-variable al roble (i.e. dummy que indica si tiene o no paso por roble) con los precios. La figura 4 muestra un grafo similar al del segmento de precios altos. El único cambio que hay, es que en este caso ahora varias de las asociaciones negativas se debilitan, lo que daría indicio de etiquetas más cargadas en vinos con roble versus sin roble.

## 4. Conclusiones

El objetivo del presente trabajo ha sido estudiar las relaciones de dependencia o asociación entre los descriptores de etiqueta en base a un modelo gráfico de redes, tomando como co-variable el precio de los vinos o el segmento definido por un intervalo específico de precios.

Con los resultados estimados pudimos corroborar la hipótesis de que las relaciones de dependencia entre descriptores no son invariantes a los precios. De hecho se observa que a medida que nos movemos hacia segmentos de vinos con precios mas altos, existen mayores relaciones de dependencia negativa entre los descriptores. Este resultado concuerda con el encontrado por (García Arancibia et al. 2015), donde comprueban que el número de descriptores juega un papel positivo sobre los vinos de precios bajos, pero negativo en vinos de precios mayores.

Existen varias extensiones que pueden realizarse para enriquecer este breve trabajo descriptivo. En primer lugar, probar con incorporar otras características objetivas como co-variables o bien como variable respuesta, a fin de analizar la potencial asociación con estos descriptores de *información adicional*. Considerando a estos descriptores como variables endógenas, otra posible extensión consiste en especificar un sistema simultáneo de equilibrio donde además de la ecuación hedónica del precio, entren estos descriptores como variables respuestas. Esto permitiría modelar conjuntamente el equilibrio en precios y atributos endógenos, enriqueciendo así la estimación de precios implícitos.

## Referencias

Barber, N., Almanza, B. & Donovan, J. (2006), ‘Motivational factors of gender, income and age on selecting a bottle of wine’, *International Journal of Wine Marketing* **18**, 218–232.

- Charters, S., Lockshin, L. & Unwin, T. (1999), ‘Consumer responses to wine bottle back labels’, *Journal of Wine Research* **10**, 183–195.
- Cheng, J., Levina, E., Wang, P. & Zhu, J. (2014), ‘A sparse ising model with covariates’, *Biometrics* **70**(4), 943–953.
- Cheng, J., Li, T., Levina, E. & Zhu, J. (2016), ‘High-dimensional mixed graphical models’, *Journal of Computational and Graphical Statistics* (**in press**).
- Corduas, M., Cinquanta, L. & Ievoli, C. (2013), ‘The importance of wine attributes for purchase decisions: A study of italian consumers’ perception’, *Food Quality and Preference* **28**, 407–418.
- Dai, B. (2013), ‘Mvb: Multivariate bernoulli log-linear model’. R package version 1.1.
- Dai, B., Ding, S. & Wahba, G. (2013), ‘Multivariate Bernoulli distribution’, *Bernoulli* **19**, 1465–1483.
- García Arancibia, R., Guiguet, E. D. & Rossini, G. (2016), *Wine Price Determination in Online Specialized Stores: An Empirical Study for Argentina*, Palgrave Macmillan US, New York, pp. 133–153.
- García Arancibia, R., Rossini, G. & Depetris Guiguet, E. (2015), ‘Wine Label Descriptors and Shelf Price Paid by Argentine Consumers’, *Agricultural Economics Review* **16**(2), 1–17.
- Lockshin, L. & Corsi, A. M. (2012), ‘Consumer behaviour for wine 2.0: A review since 2003 and future directions’, *Wine Economics and Policy* **1**(1), 2 – 23.
- Lockshin, L., Mueller Loose, S., Louviere, J., Francis, I. & Osidacz Williamson, P. (2009), ‘Development of a new method to measure how consumers choose wine’, *The Australian and New Zealand Wine Industry Journal* **24**.
- Meinshausen, N. & Bühlmann, P. (2006), ‘High-dimensional graphs and variable selection with the lasso’, *The Annals of Statistics* **34**(3), 1436–1462.
- Mora, P. & Livat, F. (2013), ‘Does storytelling add value to fine bordeaux wines?’, *Wine Economics and Policy* **2**(1), 3 – 10.
- Mueller Loose, S., Lockshin, L., Louviere, J., Francis, I. & Osidacz Williamson, P. (2009), ‘How does shelf information influence consumers’ wine choice?’, *Australian and New Zealand Wine Industry Journal* **24**.
- Mueller Loose, S., Lockshin, L., Saltman, Y. & Blanford, J. (2010), ‘Message on a bottle: The relative influence of wine back label information on wine choice’, *Food Quality and Preference* **21**, 22–32.
- Mueller Loose, S. & Szolnoki, G. (2010), ‘Wine packaging and labelling-do they impact market price? a hedonic price analysis of us scanner data’, pp. 8–10.
- Mueller, S., Lockshin, L., Saltman, Y. & Blanford, J. (2010), ‘Message on a bottle: The relative influence of wine back label information on wine choice’, *Food Quality and Preference* **21**(1), 22 – 32.

- Saak, A. E. (2011), 'A model of labeling with horizontal differentiation and cost variability', *American Journal of Agricultural Economics* **93**(4), 1131–1150.
- Schäufele, I. & Hamm, U. (2017), 'Consumers' perceptions, preferences and willingness-to-pay for wine with sustainability characteristics: A review', *Journal of Cleaner Production* **147**, 379 – 394.
- Thomas, D. & Pickering, G. (2003), 'The importance of wine label information', *International Journal of Wine Marketing* **15**, 58–74.