

Una estimación de la demanda nacional de vinos

Autores: Santiago Fernández
Mariano Grondona
Lucía Mauro
Anabella Sammarone

Cátedra: Econometría II

Docentes: Lic. Horacio Fuster, Lic. Natacha Liseras y Lic. Daniela Calá

**Facultad de Ciencias Económicas y Sociales
Universidad Nacional de Mar del Plata**

2006

Índice

Introducción	3
Marco teórico	4
Metodología aplicada	6
Modelo de Heckman	6
Modelo a estimar	7
Ajuste de precios	9
Análisis de la variable “Litros de vino consumidos por hogar por mes”	11
Análisis por región	12
Análisis por quintil de ingreso	1
Análisis sin considerar los casos nulos	2
Estimación del modelo	3
Estimación del modelo	4
Significatividad del modelo	4
Bondad del ajuste del modelo	4
Significatividad e interpretación de los coeficientes estimados	5
Evaluación de los supuestos del modelo MCO	1
Análisis de la independencia de los residuos	1
Análisis de multicolinealidad	1
Análisis de heteroscedasticidad	2
Análisis de normalidad de residuos	4
Análisis de influencia	5
Conclusiones	6
Conclusiones	6
Bibliografía	7

Introducción

El presente trabajo tiene por objetivo estimar la demanda nacional de litros de vino. Para ello, se utilizaron los datos de la Encuesta Nacional de Gastos de los Hogares realizada en 1996-1997 por el INDEC¹. De este modo, se busca establecer los principales determinantes de la demanda nacional de vinos. Este trabajo, en consecuencia, realiza un aporte al contrastar empíricamente algunos de los postulados de la teoría microeconómica clásica.

El informe se encuentra estructurado del siguiente modo. En el primer apartado se presenta el marco teórico en el que se encuadra el análisis. Se enuncian los principales elementos de la teoría económica en relación a funciones de demanda de bienes. En el segundo apartado se detalla la metodología aplicada y se formula el modelo econométrico. Luego, en el tercer apartado, se efectúa un análisis descriptivo de la variable que se desea explicar. Finalmente, se presenta el modelo estimado, se exponen los resultados de la investigación y las conclusiones.

¹ Se agradece la colaboración del grupo de investigación de Economía Agraria de la UNMdP por haber facilitado los datos de la ENGHO 1996.

Marco teórico

Las curvas de demanda son simplemente una representación gráfica de las preferencias por un bien particular. La forma de la curva de demanda dice al economista mucho acerca del grado de deseo del bien en cuestión (Gould y Lazear, 2000). Para cada sujeto existirá una curva de demanda determinada para cada bien, de modo que la demanda de mercado es simplemente la suma de demanda de todos los consumidores. Asimismo, es importante conocer que es lo que condiciona la función de demanda de los bienes. Por lo que, siguiendo a Lipsey (1977), podríamos decir que los determinantes de la cantidad demandada de un bien son:

- El precio del bien considerado: en la mayoría de los bienes sucede que al disminuir su precio, la cantidad demandada por los consumidores aumenta, permaneciendo constante los demás determinantes. Es decir que la relación entre las dos variables es inversa. A su vez, la disminución del precio, le permite al bien competir con sus sustitutos. Las variaciones el precio del bien inducen, en consecuencia, movimientos sobre la misma curva de demanda.

- El nivel de ingreso del consumidor: normalmente, es de esperar que un incremento en el ingreso provoque un aumento en la demanda del bien. Si se trata de un bien normal, un cambio en el ingreso del consumidor induce a un desplazamiento de la curva de demanda. Asimismo, existen ciertos bienes en los cuales un incremento en el ingreso deja inalterada la demanda. Incluso puede suceder que se reduzca la demanda de un bien al aumentar el ingreso (bien inferior).

- Los precios de los demás bienes (sustitutos, complementarios): si una disminución en el precio de un bien provoca una reducción en la demanda de otro bien, entonces tales bienes son sustitutos. Asimismo, si un descenso en el precio de un bien incrementa la demanda de otro bien, entonces ambos bienes son complementarios.

- Los gustos del consumidor: si un bien se pone de moda entre los consumidores, es de esperar que la demanda de dicho bien se eleve.

De este modo, la función de demanda puede expresarse como:

$$Q_n^d = D(P_n, P_1, \dots, P_{n-1}, I, G)$$

Donde, Q_n^d es la cantidad demandada del bien n, P_n el precio del bien n, P_1 a P_{n-1} el precio de los demás bienes relacionados con el bien analizado, I el ingreso del consumidor y G representa sus gustos.

Metodología aplicada

En el presente trabajo se propone para la estimación de la función de demanda nacional de vino, un modelo uniecuacional y multivariado. En la estimación de la función se utiliza el Modelo de Heckman. También se realiza un ajuste de precios por calidad. Ambas cuestiones son explicadas en detalle a continuación.

Modelo de Heckman

Debido a que la variable considerada se encuentra censurada², es decir que posee muchos valores cero (alrededor del 70%) es posible, o bien utilizar solamente los casos positivos para estimar los parámetros del modelo, o bien incluir todos los casos.

Es importante diferenciar el consumo nulo de los hogares debido a que no se consumió vino en la semana de referencia, de los hogares en donde la causa del consumo nulo está en que habitualmente no se consume dicho bien. Es por ello, que en el presente trabajo se opta por incluir todos los casos.

Sin embargo, cuando se toman todos los casos, la distribución de la variable litros de vino consumidos genera un sesgo en la estimación, por lo que es necesario realizar un ajuste que corrija ese sesgo. A tal efecto, se utiliza el procedimiento propuesto por Heckman (1979). El mismo, consiste en estimar, en una primera etapa, un modelo *probit* cuya variable respuesta es uno si el hogar realiza un gasto positivo o cero caso contrario. En una segunda etapa se incorpora la variable generada en la primera etapa (λ), como explicativa del modelo que tiene como variable dependiente la cantidad demandada (Sosa Escudero, 1998).

² Se distinguen los datos censurados de los datos truncados en tanto que, “una muestra se considera truncada cuando ciertas observaciones son sistemáticamente excluidas de la muestra. En el caso de los datos censurados ninguna observación es excluida, pero cierta información en la misma es sistemáticamente alterada” (Sosa, 40).

1ª etapa:

Se estima por Máxima Verosimilitud, $\gamma = \beta / \sigma$ utilizando el modelo *probit*, usando todas las observaciones de la muestra. La variable binaria \hat{y}_i es la variable dependiente. Luego, se construye la variable $\lambda_i = \phi() / \Phi()$. El numerador de esta expresión representa la densidad probabilística y el denominador la probabilidad acumulada, ambas correspondientes a la distribución normal.

2ª etapa:

Se estima el modelo original, mediante MCO regresando la variable y_i en el vector de variables explicativas x_i adicionando la variable λ_i obtenida en el paso anterior. En esta etapa se utiliza la muestra truncada, es decir, solo las observaciones no censuradas.

Intuitivamente, la primera etapa del modelo de Heckman estima un modelo de la probabilidad de que los datos sean no censurados, con el propósito de obtener estimaciones de la variable λ_i . En la segunda etapa el modelo truncado es estimado, pero incorporando esta variable adicional, lo cual corrige el problema del sesgo encontrado.

Modelo a estimar

En el presente trabajo, el modelo teórico propuesto es:

$$Q_i^d = \beta_0 + \beta_1 P_i + \beta_2 P c_i + \beta_3 E j_i + \beta_4 I_i + \beta_5 E a_i + \beta_6 E m_i + \beta_7 C_i + \beta_8 A_i + \beta_9 \lambda_i + \mu_i$$

- El precio del vino. Este es un precio implícito, dado que no se dispone de información específica acerca del mismo, sino que se lo obtiene de la división del gasto total en vino por la cantidad de litros consumidos, todo ello para cada hogar. Asimismo, el precio se ajusto por el "efecto calidad". (P)
- El precio de la cerveza (bien sustituto). También se utiliza un precio implícito. (Pc)

- La edad del jefe de hogar. (*Ej*)
- El ingreso neto del hogar. (*I*)
- Nivel de educación del jefe del hogar. Se utilizan dos variables *dummy* para captar tres niveles de educación: baja, media y alta. (*Ea, Em*)
- La cantidad de miembros mayores de catorce años del hogar. Este dato fue calculado como la diferencia entre la cantidad total de miembros de la familia y la cantidad de menores de 14 años. (*C*)
- Influencia causada por valores consumo de vino elevado. En la muestra existen hogares con valores muy elevados de consumo de vino, respecto de los cuales no se dispone de la información necesaria para explicar ese consumo. Por lo tanto, se utiliza una variable *dummy*, considerando si el hogar pertenece al 5% de los hogares que más consumen vino o no. (*A*)
- Variable resultante de la aplicación de un modelo *probit*, según la primera etapa del modelo de Heckman utilizado. (λ)

La elección de las variables respondió a las siguientes consideraciones. En primer lugar, la inclusión de la variable precio del vino e ingreso del hogar como explicativas de la cantidad demandada de vino responden a la teoría microeconómica. En segundo lugar, se considera a la cerveza como un bien sustituto, no por sus características intrínsecas, sino porque en los últimos años se observa una tendencia al aumento del consumo de cerveza en detrimento del de vino³. Finalmente, se toma la edad del jefe del hogar y el nivel educativo alcanzado por el jefe del hogar como variables socioeconómicas relevantes para explicar la demanda de vino.

³ “En la última década no sólo aumentó el número de consumidores de bebidas alcohólicas sino también los hábitos ante una copa. Desde 1980, el consumo de vino cayó más del 30 por ciento. En el mismo período, el de cerveza aumentó un 500 por ciento” En: <http://www.alcoholinformate.org.mx/estadisticas.cfm?articulo=135>.

Ajuste de precios

Los precios del vino que se utilizan son implícitos y dado que, surgen de datos de corte transversal y que se considera al vino como un bien no homogéneo⁴, aparecen “efectos de calidad” que deben ser tenidos en cuenta. Por tal motivo y, siguiendo a Berges y Casellas (2002), se realizó una estimación del precio del vino por medio del método de Mínimos Cuadrados Ordinarios (MCO) en función de las siguientes variables:

- Cantidad de miembros mayores a catorce años que forman parte del hogar (Miem).
- Región geográfica a la que el hogar pertenece. Se toman cinco variables *dummy* (Metr, Pamp, Noro, Nore, Cuyo). Siendo Patagonia la categoría base.
- Nivel de educación del jefe del hogar. Se toma una variables *dummy* (Ealta), considerando educación alta si ha alcanzado al menos el nivel secundario completo, y educación baja si no.
- Quintil de ingreso al que pertenece el hogar. Se tomaron dos variables *dummy* (Ialto, Ibajo), considerando ingreso alto si pertenece al quintil 5, ingreso bajo si pertenece al quintil 1 e ingreso medio si pertenece el hogar a los quintiles 2, 3 o 4.
- Sexo del jefe del hogar. Se toma una variable *dummy* (sexo), que toma el valor uno para los varones y el valor cero para las mujeres.

De este modo, la función de precios (P_i) resultante sería:

$$P_i = \beta_0 + \beta_1 \text{Miem}_i + \beta_2 \text{Metr}_i + \beta_3 \text{Pamp}_i + \beta_4 \text{Noro}_i + \beta_5 \text{Nore}_i + \beta_6 \text{Cuyo}_i + \beta_7 \text{Ealta}_i + \beta_8 \text{Ialto}_i + \beta_9 \text{Ibajo}_i + \beta_{10} \text{Sexo}_i + \mu_i$$

⁴ Se entiende que el vino es un bien heterogéneo porque existen diferencias en cuanto al tipo de uva utilizado en su producción. Aún cuando éste coincida, podrían aparecer características distintas debido a las condiciones agroecológicas en las que la uva crece. Asimismo, existen diferencias en cuanto al manejo del proceso productivo.

Finalmente, el precio ajustado se calcula sumando el residuo de la regresión y el intercepto.

El precio de la cerveza no fue ajustado, dado que se considera que el bien es homogéneo, ya que no presenta importantes diferencias en cuanto a sus características y su calidad.

Análisis de la variable “Litros de vino consumidos por hogar por mes”

A partir de la Encuesta Nacional de Gastos Hogares realizada en 1996-1997, se dispone de datos correspondientes a 27260 hogares de todo el país, dividiendo el territorio en seis regiones: Metropolitana, Pampeana, Noroeste, Noreste, Cuyo y Patagonia.

A la hora de analizar la variable en cuestión se obtienen los siguientes resultados. El valor de la media es de 3,89 litros de vino por hogar por mes. Cabe aclarar que el relevamiento de la información se realizó sólo para una semana y luego es mensualizado. Es por esto que existen muchos casos de no consumo en la semana de referencia, en total 19149 hogares (70%). El análisis de las estadísticas descriptivas restringido sólo a los hogares que consumen vino se presenta más adelante.

La desviación estándar es de 8,692 litros⁵. En función de este dato se puede establecer un intervalo de confianza para la media poblacional estimada que iría desde 0 litros hasta 12,582 litros⁶.

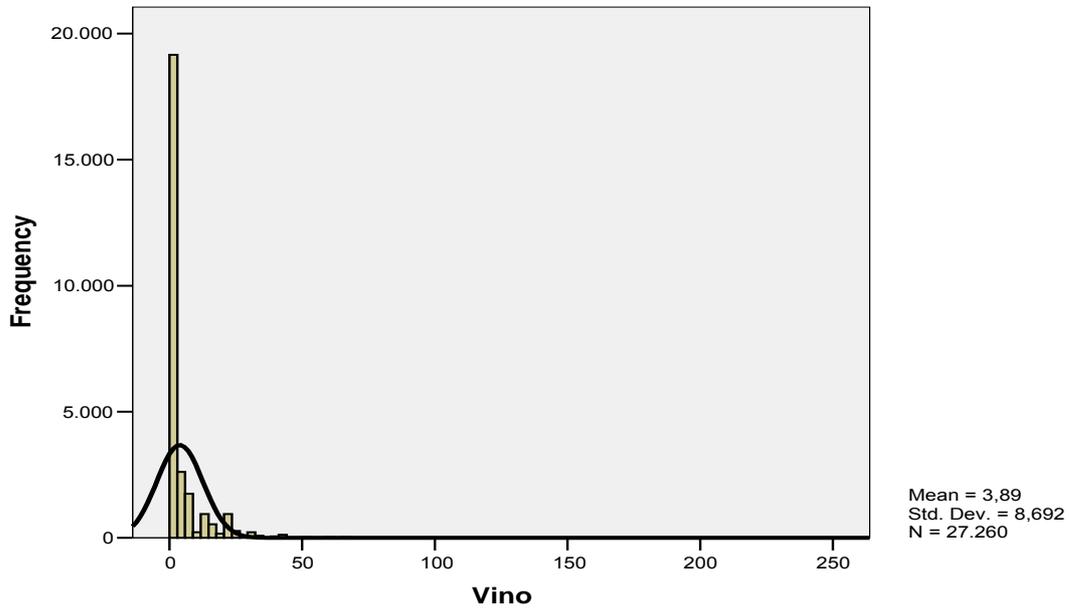
La variable presenta un comportamiento asimétrico sesgado hacia la derecha, lo que implica que la mediana es menor que la media, siendo el valor de asimetría de 4,37. Por otro lado, la curtosis es de 41,551, por lo tanto la distribución es leptocúrtica, es decir que hay una alta concentración de datos alrededor de la media. En función de lo anterior, podemos inferir que la variable no se ajusta a una distribución normal.

El valor mínimo de litros que se consumen es de 0, mientras que el máximo es de 217 litros. La cantidad de litros totales que se consumieron en la Argentina en el mes considerado es de 105989. Finalmente, el 25% de los hogares que más consumen, consumen como mínimo 4,33 litros.

⁵ De aquí en adelante, cuando nos refiramos a “litros”, se considerará litros de vino por hogar por mes.

⁶ El límite inferior del intervalo da un valor negativo que carece de relevancia.

Histogram



Statistics

Vino		
N	Valid	27260
	Missing	0
Mean		3,89
Std. Error of Mean		,053
Median		,00
Mode		0
Std. Deviation		8,692
Variance		75,544
Skewness		4,370
Std. Error of Skewness		,015
Kurtosis		41,551
Std. Error of Kurtosis		,030
Range		217
Minimum		0
Maximum		217
Sum		105989
Percentiles	10	,00
	25	,00
	50	,00
	75	4,33

Análisis por región

A continuación se realizó un análisis descriptivo de la variable por región. Los resultados son los siguientes:

Región	Media	Desvío estándar
Metropolitana	3.7	9.017
Pampeana	4.8	9.414
Noroeste	3.15	7.981
Noreste	3.03	7.614
Cuyo	4.45	9.228
Patagónica	3.84	7.675

Se observa que el mayor consumo de litros de vino es en la región Pampeana, mientras que el menor es en Noreste.

La variable presenta una fuerte asimetría en todas las regiones, en tanto que la región de Cuyo es una excepción, dado que el valor de asimetría es apenas 2,799 litros.

Análisis por quintil de ingreso

Luego se realizó una comparación de la variable en función de los quintiles de ingreso. La conclusión es la siguiente:

Quintil	Media	Desvío estándar
1	2.44	6.994
2	3.33	8.157
3	4.02	8.464
4	4.4	8.873
5	5.48	10.561

A partir de los valores mencionados se puede observar que el 20% de los hogares con mayor ingreso consumen aproximadamente el doble de litros de vino que el 20% de menor ingreso.

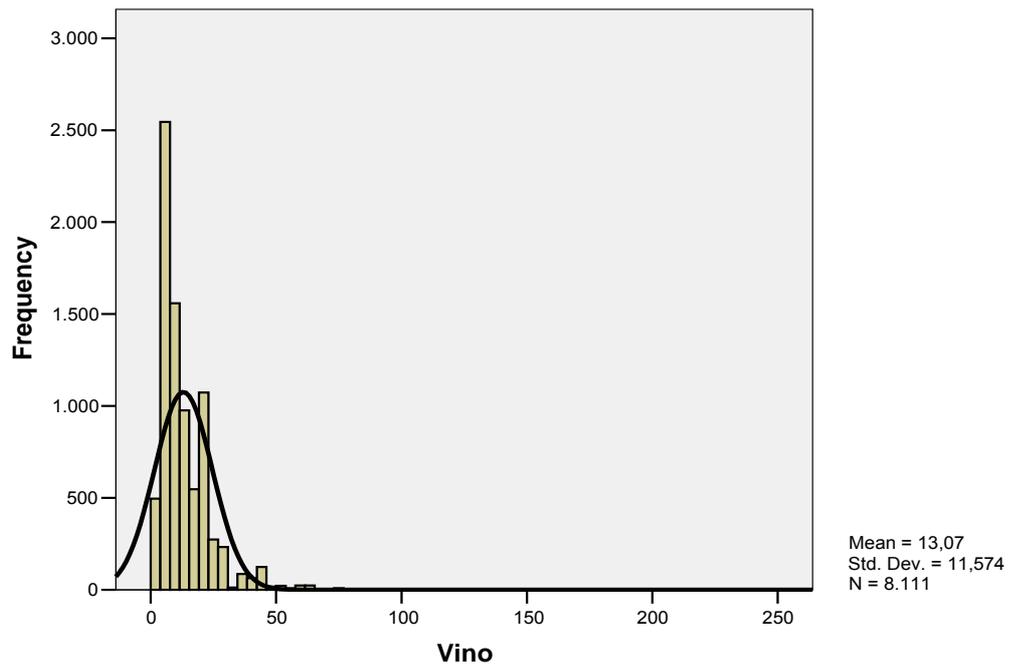
Análisis sin considerar los casos nulos

Teniendo en cuenta lo explicado anteriormente respecto de que se registraron 19149 casos con valor cero, entonces se procedió a realizar un análisis descriptivo considerando solamente los casos en los que hay valores mayores a cero.

Para este caso, la media es de 13,07 litros de vino, y la desviación estándar es 11,574. El intervalo de confianza va desde 1,496 litros hasta 26,644 litros por mes. Por su parte, la mediana es de 8,66, con lo cual se sabe que el 50% de los hogares que menos consumen, consumen como máximo esa cantidad de litros, mientras que el 50% de los que más consumen, consumen más de 8,66 litros. Finalmente, la cantidad de litros que es consumida con mayor frecuencia es 4 litros. Dado que la media, el modo y la mediana no coinciden, entonces podemos decir que la distribución no es normal.

La variable presenta una distribución asimétrica sesgada a la derecha y además es leptocúrtica. Sin embargo, se asemeja más a una normal, que considerando todos los casos.

Histogram



Statistics

Vino		
N	Valid	8111
	Missing	0
Mean		13,07
Std. Error of Mean		,129
Median		8,66
Mode		4
Std. Deviation		11,574
Variance		133,953
Skewness		3,433
Std. Error of Skewness		,027
Kurtosis		29,550
Std. Error of Kurtosis		,054
Range		215
Minimum		1
Maximum		217
Sum		105989
Percentiles	10	4,33
	25	4,33
	50	8,66
	75	17,32

Estimación del modelo

Luego de realizada la estimación por el método de Mínimos Cuadrados Ordinarios, la función de demanda resultante es:

$$\hat{Q}_i^d = 16,889 - 1,707P_i + 0,456P_{c_i} + 0,017E_{j_i} + 0,001I_i - 1,837E_{a_i} - 1,657E_{m_i} + 0,316C_i + 98,424A_i - 2,929\lambda_i$$

Significatividad del modelo

Se realizó un Test F para establecer la significatividad global del modelo. La hipótesis nula es que el modelo no es globalmente significativo, y resulta rechazada con el 95% de confianza, dado que el valor p es 0,000. Es decir, que el modelo es globalmente significativo.

ANOVA^b

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	220580.7	9	24508,964	229,329	,000 ^a
	Residual	865775.2	8101	106,873		
	Total	1086356	8110			

a. Predictors: (Constant), valores de vino alto, lambda_e, preciosimp cerv, Precioajustado, edumedio, Edad del jefe del hogar, cantmiemmay, ingreso neto total del hogar, edualto

b. Dependent Variable: Vino

Bondad del ajuste

Para analizar la bondad del ajuste del modelo, es decir encontrar en qué medida ajusta la línea de regresión muestral a los datos, se utilizó el coeficiente R² ajustado. El valor del mismo es de 20,3 % que parece ser un valor bajo pero que, en realidad es aceptable para un modelo de corte transversal, debido a que en estos modelos no se dispone de la

información necesaria para lograr explicar una mayor proporción de la variabilidad la cantidad demandada de litros de vino.

Model Summary^b

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	,451 ^a	,203	,202	10,338

a. Predictors: (Constant), valores de vino alto, lambda_e, preciosimp cerv, Precioajustado, edumedio, Edad del jefe del hogar, cantmiemmay, ingreso neto total del hogar, edualto

b. Dependent Variable: Vino

Significatividad e interpretación de los coeficientes estimados

- $\hat{\beta}_0 = 16,889$ → Resulta significativo con el 95% de confianza.
Indica que se espera que en promedio se demanden 16,889 litros de vino, cuando las demás variables toman valor cero.
- $\hat{\beta}_1 = -1,707$ → Resulta significativo con el 95% de confianza.
Indica que se espera que en promedio la cantidad demandada de vino disminuya en 1,707 litros cuando el precio del vino aumenta en \$1, manteniendo todo lo demás constante. El signo de este coeficiente coincide con el predicho por la teoría microeconómica.
- $\hat{\beta}_2 = 0,456$ → Resulta significativo con el 95% de confianza.
Indica que se espera que en promedio la cantidad demandada de vino aumente en 0,456 litros cuando el precio de la cerveza aumenta en \$1, manteniendo todo lo demás constante. El signo de este coeficiente coincide con el predicho por la teoría microeconómica.
- $\hat{\beta}_3 = 0,017$ → Resulta significativo al 94% (valor p = 0,056).

Indica que se espera que en promedio la cantidad demandada de vino aumente en 0,017 litros cuando el jefe del hogar tenga un año más, manteniendo todo lo demás constante. El signo de este coeficiente coincide con el predicho por la teoría microeconómica.

- $\hat{\beta}_4 = 0,001$ → Resulta significativo con el 95% de confianza.

Indica que se espera que en promedio la cantidad demandada de vino aumente en 1 litro cuando el ingreso neto del hogar aumenta en \$100, manteniendo todo lo demás constante. El signo de este coeficiente coincide con el predicho por la teoría macroeconómica para los bienes normales.

- $\hat{\beta}_5 = -1,837$ → Resulta significativo con el 95% de confianza.

Indica que se espera que en promedio la cantidad demandada de vino sea inferior para aquellos individuos con nivel de educación alto que para aquellos con niveles de educación bajo, en aproximadamente 1,837 litros, manteniendo todo lo demás constante.

- $\hat{\beta}_6 = -1,657$ → Resulta significativo con el 95% de confianza.

Indica que se espera que en promedio la cantidad demandada de vino sea inferior para aquellos individuos con nivel de educación medio que para aquellos con niveles de educación bajo, en aproximadamente 1,657 litros, manteniendo todo lo demás constante.

- $\hat{\beta}_7 = 0,316$ → Resulta significativo con el 95% de confianza.

Indica que se espera que en promedio la cantidad demandada de vino aumente en 0,316 litros por cada miembro adicional del

hogar mayor de catorce años, manteniendo todo lo demás constante.

- $\hat{\beta}_8 = 98,424$ → Resulta significativo con el 95% de confianza.

Indica que se espera que en promedio la cantidad demandada de vino aumente en 98,424 litros si el hogar pertenece al 5% de los hogares que mayor cantidad de vino consumen, manteniendo todo lo demás constante.

Coefficients^a

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	16,889	1,206		14,005	,000
	Edad del jefe del hogar	,017	,009	,020	1,910	,056
	ingreso neto total del hogar	,001	,000	,070	5,689	,000
	preciosimpcerv	,456	,200	,023	2,274	,023
	cantmiemmay	,316	,095	,038	3,334	,001
	lambda_e	-2,929	,678	-,054	-4,319	,000
	edumedio	-1,657	,320	-,067	-5,170	,000
	edualto	-1,837	,462	-,055	-3,976	,000
	Precioajustado	-1,707	,108	-,169	-15,837	,000
	valores de vino alto	98,424	2,444	,400	40,278	,000

a. Dependent Variable: Vino

Evaluación de los supuestos del modelo MCO

Análisis de la independencia de los residuos

Dado que existe aproximadamente un 80% de la variabilidad de la cantidad demandada de vino que no es explicada por el modelo, es de esperarse que exista una estructura en los residuos.

Se realizó una Prueba de rachas, bajo la hipótesis nula de que los residuos se distribuyen aleatoriamente. El valor p obtenido es de 0,026, con lo cual se rechaza la hipótesis nula al 5% de significatividad. Esto quiere decir, que existe en los residuos una estructura no captada por el modelo. Sin embargo, como no se cuenta con la información necesaria para corregir el problema, la estructura encontrada no puede ser evitada.

Runs Test

	Unstandardized Residual
Test Value ^a	,0000000
Cases < Test Value	4989
Cases >= Test Value	3122
Total Cases	8111
Number of Runs	3747
Z	-2,219
Asymp. Sig. (2-tailed)	,026

a. Mean

Análisis de multicolinealidad

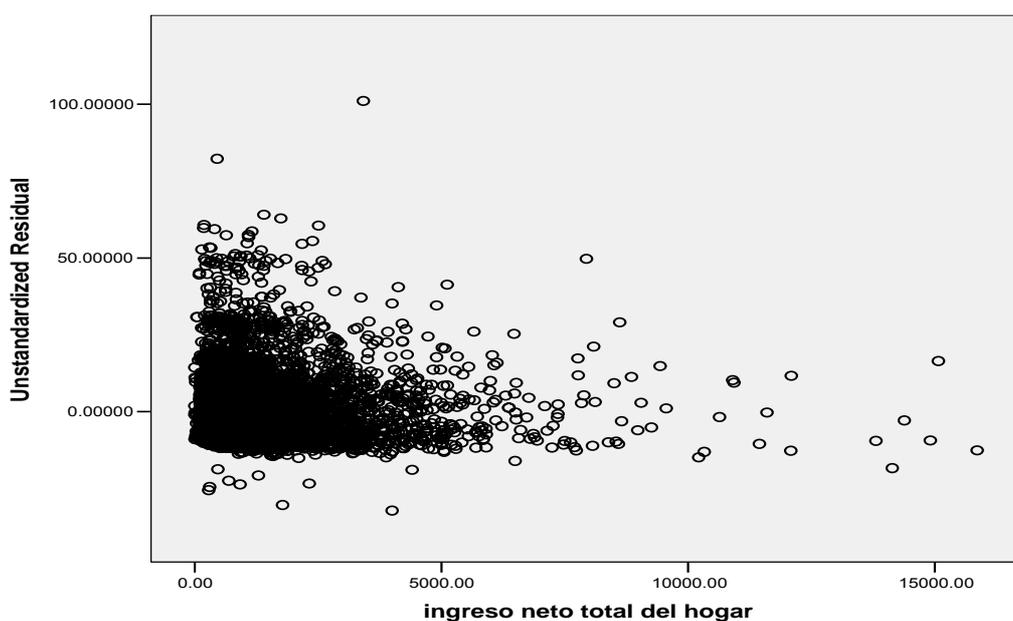
Para diagnosticar la existencia de multicolinealidad se calculó el Índice de Condición (IC), sin incluir la variable λ . El máximo valor resultó ser de 14,369, lo cual indica la existencia de multicolinealidad moderada. Se considera multicolinealidad moderada cuando el IC resulta menor a 30.

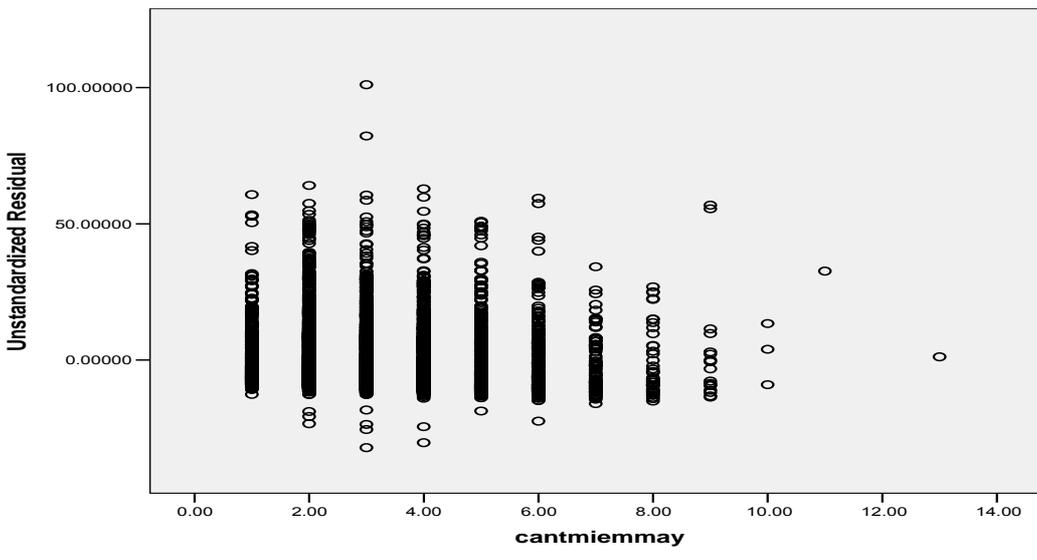
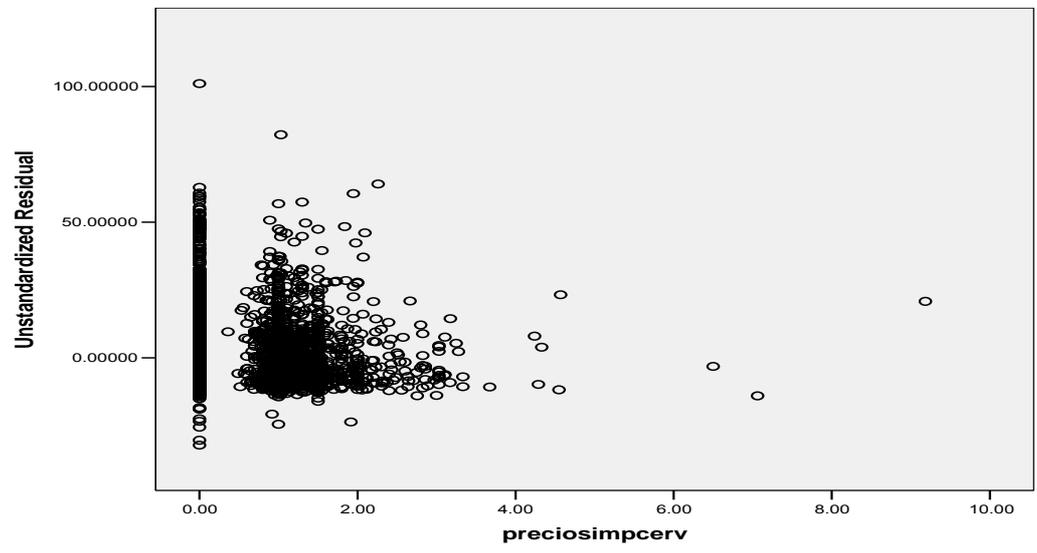
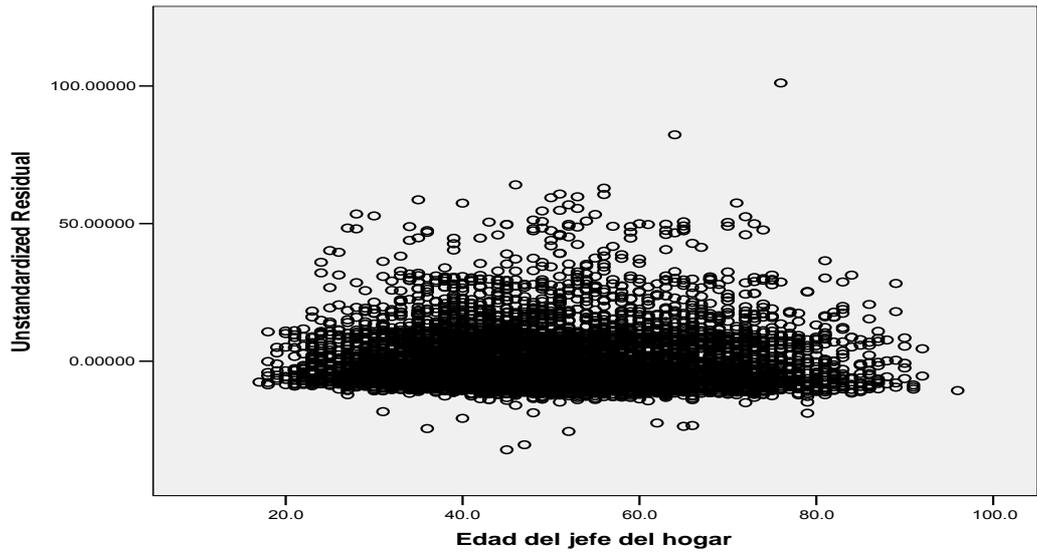
Model	Dimension	Eigenvalue	Condition Index
1	1	5.117	1.000
	2	1.103	2.154
	3	.994	2.269
	4	.766	2.585
	5	.380	3.671
	6	.316	4.026
	7	.188	5.219
	8	.112	6.749
	9	.025	14.369

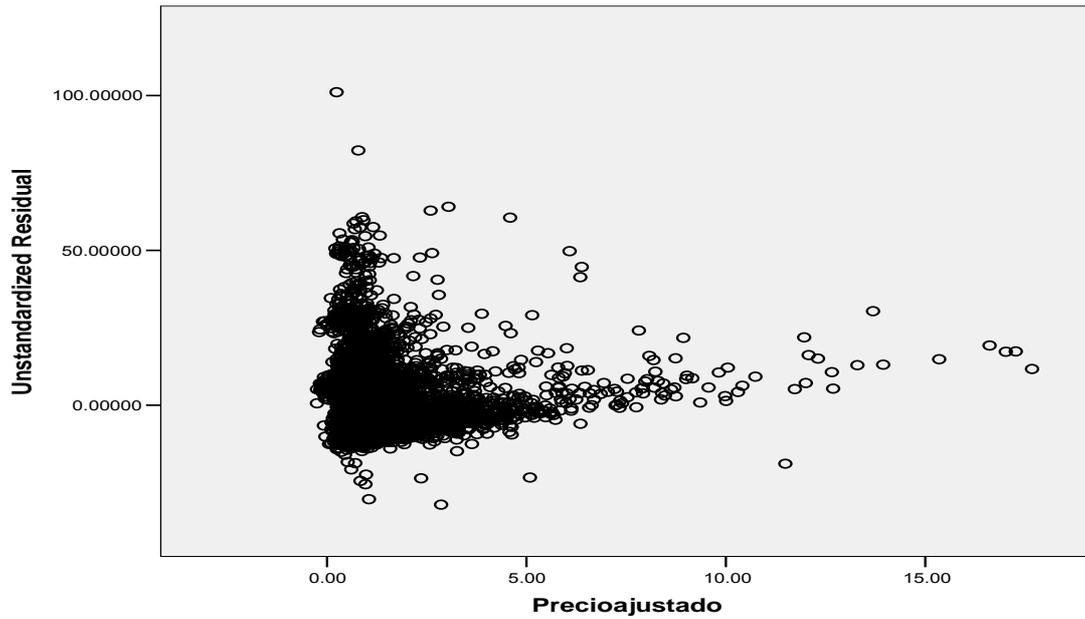
a. Dependent Variable: Vino

Análisis de heteroscedasticidad

Para diagnosticar la existencia de heteroscedasticidad se utilizó una prueba no formal. Se graficó el valor de los residuos contra cada una de las variables que podrían hacer que la varianza no sea homoscedástica. De los gráficos se extrae la conclusión de que no existe un grave problema de heteroscedasticidad, ya que no se observa ningún patrón de comportamiento demasiado fuerte en los residuos.







Análisis de normalidad de residuos

Se realizó un Test Kolmogorov-Smirnov para analizar la normalidad de los residuos, bajo la hipótesis nula de que su distribución es normal. Con el 95% de confianza se rechaza la hipótesis nula, por lo tanto los residuos no se distribuyen normalmente.

One-Sample Kolmogorov-Smirnov Test

		Standardized Residual
N		3122
Normal Parameters ^{a,b}	Mean	.9686779
	Std. Deviation	.96177410
Most Extreme Differences	Absolute	.193
	Positive	.193
	Negative	-.157
Kolmogorov-Smirnov Z		10.787
Asymp. Sig. (2-tailed)		.000

a. Test distribution is Normal.

b. Calculated from data.

Análisis de influencia

Se analizó la posible existencia de datos influyentes (SAS Institute Inc., 1999) utilizando como medidas de influencia la Distancia de Cook, el Leverage, y el Diffit, siendo los resultados:

- Distancia de Cook: Los valores obtenidos se comparan con el valor informal de referencia $F_{(0,5; 11; 8100)} = 0,94$ → Dado que no existen valores superiores a 0,94, se puede decir que no existen observaciones que al ser removidas tengan gran efecto sobre los valores ajustados.
- *Leverage*: Los valores obtenidos se comparan con una medida de referencia⁷ = 0,002 → Existe un conjunto de datos observados cuyo *leverage* es superior al punto de corte.
- *Diffit*: El valor de referencia⁸ considerado es 0,007 → Existe un conjunto de observaciones cuyo valor de *Diffit* es superior al punto de corte, y que son influyentes en la estimación del modelo.

Las observaciones influyentes coinciden.

⁷ Medida de referencia: $2p/n$, siendo p el número de parámetros estimados y n el número de observaciones.

⁸ Medida de referencia: $2\sqrt{p/n}$.

Conclusiones

Luego de realizado el trabajo, se puede concluir que la cantidad demandada de litros de vino de los hogares depende, según el modelo estimado de las siguientes variables. En primer lugar, al aumentar el precio del vino, su cantidad demandada disminuye, ceteris paribus. En segundo lugar, la cantidad demandada aumenta al aumentar el precio de la cerveza, ceteris paribus. En tercer lugar, al aumentar el ingreso, aumenta también la cantidad, ceteris paribus.

En cuanto a las variables socioeconómicas consideradas, se puede decir que al aumentar el número de miembros del hogar mayores de catorce años, aumenta la cantidad demandada de vino. Asimismo, al aumentar la edad del jefe del hogar aumenta la cantidad demandada. Por otro lado, los hogares con educación media y/o alta consumen menos del bien, que los hogares con educación baja.

Las variables que producen mayor efecto sobre la cantidad demandada son el precio ajustado del vino, y la variable binaria que capta el alto consumo de algunos hogares. Cabe destacar, que esta última es la variable que posee un coeficiente beta estandarizado mayor.

Los signos de los coeficientes estimados se corresponden con los predichos por la teoría económica. Si bien el coeficiente R cuadrado es bajo, esto es lo generalmente sucede en análisis de series de corte transversal. Esto se debe a que no se dispone de la información necesaria para mejorar la estimación.

Bibliografía

Berges, M. y Casellas, K. (2002) “Estimación de un sistema de demanda de alimentos. Un análisis aplicado a hogares pobres y no pobres”.

Gould y Lazear (2000) *Teoría macroeconómica*, 3º ed. en español, México: Fondo de Cultura Económica.

Gujarati, D. (1992) *Econometría*, 2º ed. En español, México: Mc Graw Hill.

Lipsey (1977) *Introducción a la economía positiva*, 10º ed., Barcelona: Vicens-vives.

SAS Institute Inc. (1999), *SAS OnlineDoc* [en cd-rom], version 8. Cary, NC: SAS Institute Inc.

Sosa Escudero, W. (1998) *Tópicos de Econometría Aplicada*, Buenos Aires, Universidad de San Andrés.