

UN MÉTODO BAYESIANO PARA LA DISTRIBUCIÓN DE LA DISPOSICIÓN A PAGAR

Ricardo Antonio Flores Huenchullanca*

RESUMEN

Esta investigación extiende la estimación de medidas de bienestar, sugeridas por Hanemann (1984), a medidas de bienestar basadas en la distribución de la disposición a pagar (DAP), para proporcionar información precisa a los decisores de política en relación a la provisión de bienes públicos. Para obtener la distribución de la DAP se usa un método bayesiano propuesto por Benoit y Van den Poel (2012). Finalmente, se realiza una simulación de Monte Carlo para probar el uso de esta metodología.

Palabras clave: disposición a pagar, regresión por cuantiles binaria, distribución asimétrica de Laplace, estimación Bayesiana.

JEL Código: C11, C53, E27

ABSTRACT

This investigation extends the estimation of the measurements of welfare suggested by Hanemann (1984), to welfare measurements based on the distribution of the 'Willingness to pay' (WTP); in this way providing precise information to the deciding parties of policies in relation to the supply of public goods. To obtain the distribution WTP, the Bayesian method proposed by Benoit and Van de Poel (2012) is applied. To test the use of this methodology, a Monte Carlo simulation is done to determine its viability.

Keywords: willingness to pay, binary quantile regression, asymmetric Laplace distribution, Bayesian estimation.

JEL Code: C11, C53, E27

* Magíster en Economía, Ingeniero Comercial y licenciado de la Universidad de Concepción. Analista de datos y programador. Correo electrónico: rflores@bcentral.cl
Fecha de recepción: 24 de abril de 2017
Fecha de aceptación: 16 de mayo de 2017

1. INTRODUCCIÓN

Esta investigación extiende la estimación de medidas de bienestar comúnmente utilizadas en valoración contingente (vc), introduciendo el uso de cuantiles de la disposición a pagar (DAP). Las medidas empleadas tradicionalmente son la media y la mediana, propuestas por Hanemann (1984). La estimación de la distribución de la DAP que se realiza en este estudio se basa en el método bayesiano propuesto por Benoit y Van den Poel (2012). Consideramos que este enfoque de estimación permite generar una información precisa para los decisores de política en relación con la provisión de bienes públicos.

El método de valoración contingente propuesto por Ciriacy y Wantrup (1947) surge ante la ausencia de mercado para los bienes públicos y la imposibilidad de conocer sus precios. Este método permite obtener el valor económico que los individuos le asignan a un bien a partir de la disposición a pagar que tienen por dicho bien. El procedimiento consiste en la aplicación de un cuestionario en el cual se simula un mercado hipotético para el bien en cuestión. El formato referéndum, propuesto por Bishop y Heberlein (1979), en que la respuesta de los individuos relacionada al pago del bien ambiental tiene un forma binaria (sí o no), es uno de los métodos que se usa con mayor frecuencia debido a que es incentivo compatible, es decir induce a revelar honestamente las preferencias de los entrevistados (Arrow *et al.*, 1993).

Con base en los trabajos de Bishop y Heberlein (1979), Hanemann (1984) discute cómo obtener compensaciones hicksianas en modelos con variable dependiente discreta, utilizando un proceso de maximización de utilidad aleatoria. Los resultados que obtiene este autor se basan en las medidas de tendencia central, media y mediana. Johansson *et al.* (1989) comentan el

trabajo de Hanemann (1984), exponiendo la posibilidad de obtener disposiciones a pagar negativas. A partir de esta discusión nace otra medida de bienestar, la media truncada. Esta nueva medida calcula la esperanza de la disposición a pagar considerando solo valores positivos en su integración, descartando los valores negativos.

En la práctica, la valoración que los individuos hacen de un bien ambiental difiere de un individuo a otro, y ello se ve reflejado en la heterogeneidad de la disposición a pagar. Por lo tanto, al usar la media como medida de bienestar, los beneficios de un proyecto que busca proveer un bien público no reflejarán la verdadera valoración del bien y tampoco será posible capturar dicha heterogeneidad, como en el ejemplo anteriormente expuesto. En este sentido, se justifica el uso de la mediana como medida de bienestar ya que, según Hanemann (1984, 1989), desde el punto de vista estadístico, la mediana es más robusta con respecto a los errores y datos atípicos de las respuestas de los encuestados y, a su vez, menos sensible al coeficiente de asimetría y curtosis de los datos Staving y Gibbons (1977). En este sentido, el criterio de Kaldor-Hicks, que usa la media para la agregación de resultados, no es adecuado desde el punto de vista teórico, cuando la disposición a pagar es heterogénea.

Adicionalmente, Hanemann (1989) plantea la necesidad de utilizar no solo el cuantil 0.5 (mediana) como medida de bienestar, sino también otros cuantiles de interés que puedan capturar información para una parte específica de la población. La estimación de la distribución de la disposición a pagar puede contribuir a responder necesidades de información específica. Por ejemplo, se puede estar interesado en conocer cuánto está dispuesto a pagar el 25% de la población por un proyecto que mejora la calidad del aire.

Para estimar los cuantiles de la disposición a pagar, utilizamos un método bayesiano de regresión binaria sugerido por Benoit y Van den Poel (2012). Este método presenta tres ventajas respecto del tradicional sugerido por Manski (1975, 1985), Horowitz (1992) y Kordas (2006). Primero, facilidad en la estimación e inferencia; segundo, convergencia de parámetros en muestras pequeñas; por último, las estimaciones funcionan correctamente cuando existen problemas de heterocedasticidad en la varianza del error.

2. MODELO TEÓRICO

2.1 Modelo de diferencia de la función indirecta de utilidad

En valoración contingente, los formatos de encuesta tipo *referendum* propuestos por Bishop y Heberlein (1987) tienen dos enfoques para modelar la elección discreta de los consumidores: el modelo de Hanemann (1984), denominado modelo de diferencia de la función indirecta de utilidad, y el de Cameron (1988), conocido como modelo de la función de variación. La diferencia entre ambos enfoques se centra en la forma de definir la variable latente y la interpretación de los parámetros. Esta investigación pretende extender el uso de las medidas sugeridas por Hanemann, por lo tanto empleamos el enfoque de diferencia de la función indirecta de utilidad. Sin embargo, en principio podemos ocupar ambos métodos, teniendo presente que la interpretación de los parámetros será diferente¹.

Lo que intentamos modelar es la respuesta de los individuos ante la pregunta: ¿estaría dispuesto usted a pagar una cantidad igual a \$A para realizar un proyecto de mejora en la calidad ambiental del recurso?

siendo las posibles repuestas “sí” o “no”. Entonces, sea $u_j = v_j(P, I; q_j)$, con $P > 0$ y $I > 0$, la función indirecta de utilidad de un individuo en particular, donde $j=1$ representa la mejora en la calidad del recurso y $j=0$ la situación contraria, P es un vector de precios que enfrenta el individuo por sus bienes, I es el ingreso familiar y q_j es un vector de calidad de los bienes que puede incorporar características socioeconómicas de los individuos. El investigador no puede observar la verdadera función indirecta de utilidad del individuo, solo una parte de ella, por lo tanto podemos asumir que cada función tiene un componente determinístico y otro estocástico, es decir

$$u_j = v_j(p, I; q_j) + \varepsilon_j, \quad j = 0, 1. \quad (1)$$

Donde ε_j es el componente estocástico no observado por el investigador. Siguiendo a McConnell y Kenneth (1990), un modelo de valoración contingente enfrenta al individuo a una elección en cuanto a una mejora en la calidad ambiental (de q_0 a q_1), por lo cual se debe pagar una cantidad A o no tener la mejora y no pagar. Sin embargo, la verdadera valoración del recurso (denotada por C) no es observable y lo único que es factible saber a partir de la respuesta de los individuos es si esta es mayor o menor que la cantidad ofrecida A_1 . Por lo tanto, el individuo aceptará pagar dicha cantidad siempre y cuando la utilidad de la mejora sea mayor que la utilidad sin la mejora, $u_1 > u_0$. Lo anterior se puede escribir en términos de probabilidad de la siguiente forma:

$$\begin{aligned} \Pr(\text{si}) &= \Pr[v_1(p, I - A; q_1) + \varepsilon_1 > v_0(p, I; q_0) + \varepsilon_0], \\ \Pr(\text{si}) &= \Pr[\Delta v > \varepsilon_0 - \varepsilon_1] = \Pr[\Delta v > \eta], \\ \Pr(\text{si}) &= F_\eta(\Delta v), \end{aligned} \quad (2)$$

donde F_η es la función de distribución acumulada de η y $\eta = \varepsilon_0 - \varepsilon_1$. Si se asume que η es normal, entonces se estima un modelo probit; por el contrario, si se asume que η se distribuye logísticamente, se estima el modelo

1. Para detalles de modelos, ver Vásquez *et al.* (2007), p. 160.

logit (Kotz *et al.*, 2001), el cual muestra que se puede construir una variable con distribución asimétrica de Laplace a partir de dos variables exponenciales estándar. Por lo tanto, esta investigación asume que ε_0 y ε_1 son i.i.d. con distribución exponencial estándar, luego η sigue una distribución asimétrica de Laplace (ALD) de la forma

$$\eta = \frac{\varepsilon_0}{\tau} - \frac{\varepsilon_1}{(1-\tau)} \sim ALD(0,1,\tau), \quad \tau \in (0,1) \quad (3)$$

La media y desviación estándar son iguales a 0 y 1, respectivamente. τ es el parámetro de asimetría.

Siguiendo a Hanemann (1984), se utiliza una función indirecta de utilidad lineal $u_j = v_j + \varepsilon_j$, donde $v_j = \alpha + \lambda I$ por lo tanto

$$\begin{aligned} v_0 &= \alpha_0 + \lambda I, \\ v_1 &= \alpha_1 + \lambda(I - A), \end{aligned}$$

indican la situación sin la mejora en la calidad ambiental y la situación con la mejora, respectivamente. En este caso se asume que las dos alternativas, aceptar o rechazar la mejora en la calidad ambiental, son sustitutos perfectos. Para encontrar Δv , perteneciente a la ecuación (2), diferenciamos v_1 y v_0 ,

$$\begin{aligned} \Delta v &= v_1 - v_0 \\ \Delta v &= \alpha_1 + \lambda(I - A) - (\alpha_0 + \lambda I), \\ \Delta v &= (\alpha_1 - \alpha_0) - \lambda A, \\ \Delta v &= \alpha - \lambda A. \end{aligned} \quad (4)$$

2.2 Medidas de bienestar clásica

Las medidas de bienestar hicksianas, como la variación compensada, representan la cantidad de dinero que un individuo está dispuesto a pagar por acceder a una

mejora en la calidad ambiental, y la denotamos por C . La cantidad de dinero C es usada como compensación para que el individuo mantenga su nivel de utilidad constante después de aceptar la mejora en la calidad ambiental, es decir se cumple

$$\begin{aligned} u_0 &= u_1, \\ \alpha_0 + \lambda I + \varepsilon_0 &= \alpha_1 + \lambda(I - C) + \varepsilon_1, \\ \alpha - \eta &= \lambda C, \end{aligned}$$

despejando con respecto a C

$$C = \frac{\alpha - \eta}{\lambda}. \quad (5)$$

Dado que η es una variable aleatoria, entonces C también es una variable aleatoria. Hanemann (1984) muestra que al asumir η con distribución normal o logística, luego las medidas de bienestar, media y mediana, serán igual a

$$C = \frac{\alpha}{\lambda}. \quad (6)$$

Donde α y λ son los parámetros del modelo logit o probit. Sin embargo, la ecuación (6) es una estimación simplificada de la disposición a pagar, puesto que entrega solo un valor (promedio) de la DAP. No muestra su heterogeneidad y su inferencia resulta complicada, ya que se debe utilizar el método Delta para calcular su varianza. Además, C no tiene distribución teórica conocida, por lo tanto su inferencia es cuestionable. Para generar intervalos se usan técnicas de remuestreo, como bootstrap, pero estas siguen siendo válidas para la media de C . En este estudio extendemos el uso de la media y mediana a otros cuantiles de las medidas de bienestar. Para ello se utiliza la especificación expuesta en el siguiente punto.

3. DISTRIBUCIÓN DE LA DISPOSICIÓN A PAGAR (DAP)

Para obtener la distribución de la DAP, asumiendo un enfoque de regresión por cuantiles (C_τ^*), se retoma la idea anterior, donde $F[\Delta v] = \tau$ representa la ecuación asociada a la estimación por cuantiles. Al descomponer esta ecuación, podemos obtener C_τ^* de la siguiente forma:

$$\begin{aligned} F_\eta[\Delta v(C_\tau^*)] &= \tau, \quad \text{con } \tau \in (0,1), \\ F_\eta[\alpha_\tau - \lambda_\tau C_\tau^*] &= \tau, \\ F_\eta^{-1}[\tau] &= \alpha_\tau - \lambda_\tau C_\tau^*, \end{aligned}$$

despejamos con respecto a C_τ^* para obtener los cuantiles de la medida de bienestar,

$$C_\tau^* = \frac{\alpha_\tau - F_\eta^{-1}[\tau]}{\lambda_\tau}. \quad (7)$$

Donde F_η es la función de distribución asimétrica de Laplace acumulada, con parámetro de localización 0 y parámetro de escala igual a 1. α_τ y λ_τ son los parámetros que provienen de la estimación binaria por cuantiles. C_τ^* , es la probabilidad τ de que un individuo esté dispuesto a pagar una cantidad C menor o igual a c_0 , $\Pr(C \leq c_0) = \tau$

C_τ^* , es la probabilidad τ de que un individuo esté dispuesto a pagar una cantidad menor o igual a C . Si $\tau = 0.5$ entonces F_η^{-1} es igual a 0, por lo tanto, $C_\tau^* = \alpha_{0.5}/\lambda_{0.5}$, que es similar a la ecuación (6).

4. ESPECIFICACIÓN DEL MODELO ECONOMETRICO

Dado que la diferencia de la función indirecta de utilidad para cada individuo es una variable latente no observada, utilizamos un modelo de regresión binario,

$$\begin{aligned} y_i^* &= x_i \beta + \mu_i, \\ y_i &= 1, \text{ si } y_i^* > 0, \\ y_i &= 0, \text{ si } y_i^* \leq 0. \end{aligned} \quad (8)$$

Donde y_i es un indicador de la i -ésima respuesta del individuo ante la pregunta: ¿estaría dispuesto usted a pagar una cantidad igual a \$A, para realizar un proyecto de mejora en la calidad ambiental del recurso? y_i^* , es la variable latente correspondiente al diferencial de la función de utilidad indirecta (Δv_i), x_i es un vector de variables explicativas de orden $1 \times k$, β es un vector de parámetros de orden $k \times 1$ y μ_i es el error que se distribuye ALD, con parámetro de localización igual a 0, parámetro de escala igual a 1 y parámetro de asimetría igual a τ , $\mu_i \sim \text{ALD}(0,1,\tau)$.

Sin embargo, si se está interesado en un modelo de regresión binaria por cuantiles, siguiendo a Ji *et al.* (2012), el estimador del τ -ésimo cuantil de una regresión binaria está dado por

$$\beta_\tau = \arg \min_{\beta} \sum_{i=1}^n \rho_\tau(y_i - 1(x_i' \beta > 0)), \quad \tau \in (0,1), \quad (9)$$

donde $1(\cdot)$ es una función indicadora y $\rho_\tau(\cdot)$ es la función de pérdida. Definida por

$$\rho_\tau = \mu\{\tau - 1(\mu > 0)\}. \quad (10)$$

Koenker y Machado (1999) fueron los primeros en mostrar la relación entre la función de pérdida anterior y la función de distribución asimétrica de Laplace (ALD), es decir, se puede mostrar que es equivalente a minimizar el problema asociado a la ecuación (9) y maximizar la función de verosimilitud de una variable aleatoria ALD. Esta dualidad se da por la igualdad entre sus funciones de pérdida asimétricas. Algo similar ocurre entre el problema de mínimos cuadrados ordinarios (MCO) y el error aleatorio con distribución normal. En este último caso, ambos poseen una función de pérdida cuadrática, siguiendo a Benoit y Van den Poel (2012), quienes usan un enfoque bayesiano basados en una distribución asimétrica de Laplace (ALD) para estimar la regresión por cuantiles binarias.

5. SIMULACIÓN DE MONTE CARLO

Con el fin de ejemplificar el uso de la metodología propuesta para estimar la distribución de la disposición a pagar, se simula la función de utilidad aleatoria para luego derivar la DAP dado un cuantil específico de su distribución. Primero,

$$\Delta v = 5 - 0.5A + \varepsilon.$$

Es la función de utilidad aleatoria, donde $\varepsilon \sim N(0,1)$ y $A \sim unif(1, 20)$. A partir de la ecuación (6) se obtiene la media de la DAP que es igual a $C = \frac{\alpha}{\lambda} = \frac{5}{0.5} = 10$. Asumiendo que si $y = 1$ si $\Delta v > 0$, y $y = 0$ si $\Delta v < 0$, se reescribe la ecuación anterior como:

$$y = 5 - 0.5A + \varepsilon.$$

Luego, se estima el método de regresión por cuantiles binarias. Con un tamaño de muestra de 2000 se realizan 1000 repeticiones del experimento de Monte Carlo. El resultado de los parámetros y la DAP se muestran en la siguiente tabla:

Tabla 1: cinco cuantiles de la distribución de la disposición a pagar

| cuantil | α | λ | $C_{\tau}^*(dap)$ |
|---------|----------|-----------|-------------------|
| 0.05 | 14.07 | -2.76 | 3.12 |
| 0.25 | 6.81 | -0.90 | 7.08 |
| 0.50 | 7.27 | -0.72 | 10.07 |
| 0.75 | 11.21 | -0.88 | 13.18 |
| 0.95 | 37.16 | -2.32 | 18.00 |

La interpretación de la disposición a pagar para un cuantil específico, en el caso de (C_{τ}^*) según la ecuación (7), es la siguiente: dado el cuantil 0.05 se observa una disposición a pagar (DAP) igual a 3.12, es decir el 5% de la población está dispuesto a pagar una cantidad menor o igual a 3.12, por otro lado, el 50 % de la población estaría dispuesto a pagar 10.07, que es similar a la cantidad teórica previamente estimada ($C = \frac{\alpha}{\lambda} = \frac{5}{0.5} = 10$).

Si solo se consideran las medidas de bienestar media y mediana, no es posible capturar la evidente heterogeneidad de la DAP. Por otro lado, la estimación por cuantiles sugiere que el cálculo de los beneficios individuales y agregados de un proyecto que intente proveer un bien ambiental será sensible a los cuantiles de la DAP.

CONCLUSIÓN

La estimación de la distribución de la disposición a pagar permite observar su heterogeneidad y genera información precisa que podría ser usada por los decisores de política, en relación a la provisión de bienes públicos. Al considerar medidas de bienestar basadas solo en la media y en la mediana, el análisis costo-beneficio puede presentar errores en el cálculo de los beneficios asociados a un proyecto de mejora en la calidad ambiental. Con respecto a la medida de bienestar más adecuada, se sugiere lo siguiente: si la DAP entre los cuantiles es homogénea, no hay problema con la elección de la medida de bienestar puesto que no habrá cambios significativos en los cuantiles de la DAP. Por el contrario, si la DAP es heterogénea, la elección de una medida de bienestar sobre un cuantil específico afecta el cálculo de los beneficios individuales y agregados de un proyecto. No obstante, como menciona Hanemann (1989), esta decisión implica un tema de carácter ético y, por lo tanto, se podría decir que ninguna medida de bienestar, en principio, es mejor que otra.

BIBLIOGRAFÍA

- Arrow, K. *et al.* (1993). Report of the NOAA Panel on Contingent Valuation.
- Benoit, D.F. y Van den Poel, D. (2012). *Binary quantile regression: a Bayesian approach based on the asymmetric Laplace distribution*. *Journal of Applied Econometrics* (27), pp. 1174-1188.
- Bishop, R.C., Heberlein, T.A. (1979). *Measuring values of extra-market goods: Are indirect measures biased?* *American Journal of Agricultural Economics* (61), pp. 926-930.
- Cameron, T.A. (1989). *A New Paradigm for Valuing Non-Market Goods Using Referendum Data: Maximum Likelihood Estimation by Censored Logistic Regression*. *Journal of environmental economics and management* (55), pp. 355-379.
- Ciriacy-Wantrup, S.V. (1947). *Capital returns from soil-conservation practices*. *Journal Farm Economics* (29), pp. 1181-1196.
- Hanemann, W.M. (1984). *Welfare Evaluations in Contingent Valuation Experiments with Discrete Responses*. *American Journal of Agricultural Economics* (66), pp. 332-341.
- Hanemann, W.M. (1989). *Welfare evaluations in contingent valuation experiments with discrete response data: Reply*. *American Journal of Agricultural Economics* (71), pp. 1057-1061.
- Horowitz, J.L. (1992). *A smoothed maximum score estimator for the binary response model*. *Econometrica* (60), pp. 505-531.

Ji, Y., Lin, N., Zhang, B. (2012). *Model selection in binary and tobit quantile regression using the Gibbs sampler*. Computational Statistics & Data Analysis (56), pp. 827-839.

Económicos y Aplicaciones. Thomson Learning. Buenos Aires.

Johansson, P.-O., Kristom, B. y Maler, K.G. (1989). *Welfare evaluations in contingent valuation experiments with discrete response data: comment*. American Journal of Agricultural Economics (71), pp. 1055-1056.

Koenker, R. y Machado, J.A.F. (1999). *Goodness of fit and related inference processes for quantile regression*. Journal of the American Statistical Association (94), pp. 1296-1310.

Kordas, G. (2006). *Smoothed binary regression quantiles*. Journal of Applied Econometrics (21), pp. 387-407.

Kotz, S., Kozubowski, T.J. y Podgórski, K. (2001). *The Laplace Distribution and Generalizations: A Revisit with Applications to Communications, Economics, Engineering and Finance*. Birkhauser: Boston, MA.

McConnell, K.E. (1990). *Models for Referendum Data: the Structure of Discrete Choice Models for Contingent Valuation*. Journal of Environmental Economics and Management (18), pp. 19-34.

Manski, C.F. (1975). *Maximum score estimation of the stochastic utility model of choice*. Journal of Econometrics (3), pp. 205-228.

Manski, C.F. (1985). *Semiparametric analysis of discrete response: asymptotic properties of the maximum score estimator*. Journal of Econometrics (27), pp. 313-333.

Vásquez, L.F., Cerda, U.A., Orrego, S.S. (2007). *Valoración Económica del Ambiente*. Fundamentos Económicos,