

Regresión Cuantilica para el Análisis del Gasto en Colombia: Una Aplicación con Datos de la ENPH

ELIÁN STEVEN BEJARANO GONZÁLEZ AND JUAN ANDRÉS GODOY REYES

Resumen

Este artículo utiliza un enfoque no paramétrico basado en regresión cuantílica para estimar la curva de Engel que describe la relación entre el gasto total del hogar y el gasto en mercado en Colombia. A diferencia de los modelos paramétricos centrados en la media, los métodos de cuantiles permiten que la curva de Engel condicional varíe de forma flexible a lo largo de distintos niveles de gasto, sin recurrir a supuestos de normalidad, homocedasticidad o estructura lineal de errores. Utilizando microdatos de la Encuesta Nacional de Presupuestos de los Hogares (ENPH), se evidencia una marcada heterogeneidad en las elasticidades del gasto entre cuantiles, revelando no linealidades y diferencias distributivas que permanecen ocultas bajo la estimación clásica por MCO. Los resultados destacan las ventajas de la inferencia no paramétrica para la estimación de curvas de Engel en datos económicos sesgados, con colas pesadas y alta heterogeneidad.

PALABRAS CLAVE: Curva de Engel; Regresión cuantílica; Métodos no paramétricos; Gasto de los hogares; Elasticidad; ENPH; Colombia.

Abstract (english)

This article applies a nonparametric quantile regression framework to estimate the Engel curve that characterizes the relationship between total household expenditure and market expenditure in Colombia. Unlike parametric mean-based models, quantile methods allow the conditional Engel curve to vary flexibly across expenditure levels, avoiding assumptions of normality, homoscedasticity, or linear error structure. Using microdata from the National Household Expenditure Survey (ENPH), we document substantial heterogeneity in expenditure elasticities across quantiles, revealing nonlinearities and distributional differences that remain concealed under classical OLS estimation. The results emphasize the advantages of nonparametric inference for Engel curve estimation in the presence of skewed, heavy-tailed, and heterogeneous economic data.

KEYWORDS AND PHRASES: Engel curve; Quantile regression; Nonparametric methods; Household expenditure; Elasticity; ENPH; Colombia.

1. INTRODUCCIÓN

El análisis del ingreso y del gasto de los hogares es un tema central en estadística aplicada y economía. Sin embargo, estas variables suelen presentar distribuciones altamente asimétricas y con valores atípicos que dificultan la aplicación de métodos que dependen de la media condicional. Dichas características hacen que los enfoques paramétricos tradicionales, como la regresión lineal estimada por mínimos cuadrados, puedan ocultar información relevante sobre la heterogeneidad entre segmentos poblacionales.

En este contexto, la regresión cuantilica constituye una herramienta poderosa, ya que permite analizar distintos puntos de la distribución condicional del gasto, ofreciendo un panorama más robusto y detallado de los patrones de consumo. Este trabajo aplica regresión cuantilica a datos de la ENPH para estudiar cómo varía la elasticidad del gasto en mercado respecto al gasto total en Colombia y en la ciudad de Bogotá.

2. PLANTEAMIENTO DEL PROBLEMA

El gasto de los hogares no responde de manera uniforme a los cambios en su ingreso o gasto total. Existen diferencias significativas entre hogares de bajos, medianos y altos niveles de consumo. Los modelos paramétricos tradicionales no distinguen adecuadamente esta heterogeneidad, lo que puede llevar a conclusiones sesgadas. Este artículo busca superar dichas limitaciones mediante el uso de cuantiles como mecanismo de modelación no paramétrica.

3. MARCO TEÓRICO

La regresión cuantilica, propuesta por Koenker y Bassett (1978), permite modelar la mediana y otros cuantiles de la distribución condicional, evitando supuestos de normalidad u homocedasticidad. Esta metodología es robusta frente a valores extremos y capta la variabilidad estructural en las relaciones económicas entre ingreso y gasto.

Los cuantiles condicionales reflejan la heterogeneidad que la media no puede capturar, y por ello son ideales para estudiar el comportamiento de consumo en hogares con diferentes niveles de recursos.

4. METODOLOGÍA

Los datos utilizados provienen de la Encuesta Nacional de Presupuestos de los Hogares (ENPH), específicamente de:

- **Gastos diarios urbanos — Mercados** (gasto en alimentos frescos, víveres y productos básicos).
- **Viviendas y Hogares** (gasto total del hogar y ciudad/dominio).

Se seleccionaron los hogares con información válida, se construyó el gasto mensual en mercado y se aplicó la transformación logarítmica para estabilizar la varianza.

Posteriormente, se estimaron:

- Regresiones cuantilicas para $q = 0.10, 0.50, 0.90$.
- Péndientes cuantilicas (elasticidades) a lo largo de $q \in [0.05, 0.95]$.
- Comparación entre Bogotá y Colombia agregada.

5. RESULTADOS

5.1. Regresiones cuantilicas para Colombia

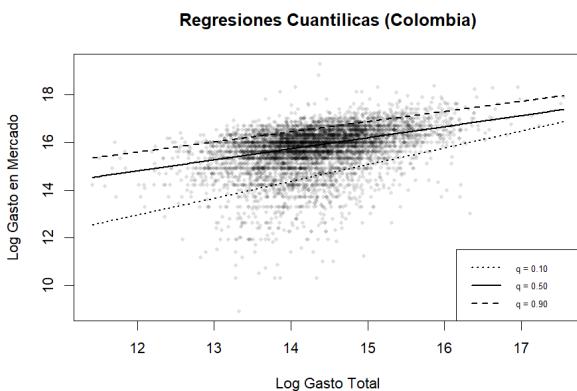


Figura 1: Regresiones cuantilicas ($q = 0.10, 0.50, 0.90$) del gasto en mercado respecto al gasto total para Colombia.

Los resultados muestran que:

- En el cuantil 0.10, la elasticidad es baja: hogares de bajo gasto responden débilmente ante aumentos en el gasto total.
- En la mediana (0.50), la respuesta es moderada.
- En el cuantil 0.90, la elasticidad es marcadamente mayor.

```

1 plot(data_colombia$log_total, data_colombia$log_mercado,
2       pch = 16, col = rgb(0,0,0,0.12), cex = 0.6,
3       xlab = "Log_Gasto_Total",
4       ylab = "Log_Gasto_en_Mercado",
5       main = "Regresiones_Cuantilicas_(Colombia)")
6
7 lines(xgrid, y_hat[,1], lty = 3, lwd = 2)
8 lines(xgrid, y_hat[,2], lty = 1, lwd = 2)
9 lines(xgrid, y_hat[,3], lty = 2, lwd = 2)
10
11 legend("bottomright",
12        legend = c("q=0.10", "q=0.50",
13                   "q=0.90"),
14        lty = c(3,1,2),
15        lwd = 2,
16        col = "black",
17        bty = "o",
18        cex = 0.6)

```

5.2. Elasticidades por cuantil

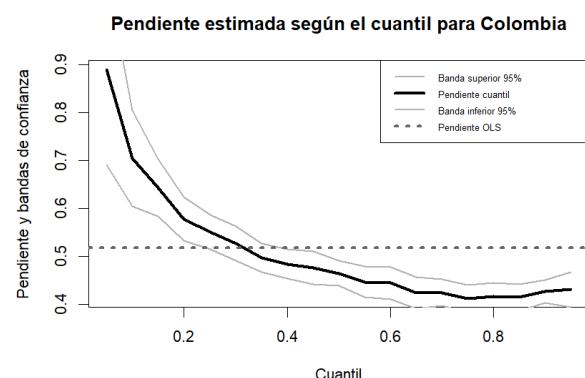


Figura 2: Elasticidad (pendiente cuantilica) del gasto en mercado para Colombia.

La curva es decreciente: hogares de bajo gasto presentan elasticidades más altas que los hogares con gastos elevados.

```

1 plot(taus, slopes_col,
2       type = "l", lwd = 3, col = "black",
3       xlab = "Cuantil",
4       ylab = "Pendiente y bandas de
5           confianza",
6       main = "Pendiente estimada segun el
7           cuantil para Colombia")
8
9
10 lines(taus, upper, col = "gray70", lwd =
11        2)
12 lines(taus, lower, col = "gray70", lwd =
13        2)
14
15 abline(h = coef(ols_col)[2],
16         lty = 3, lwd = 3, col = "gray40")
17
18 legend("topright",
19        legend = c("Banda superior 95 %",
20                  "Pendiente cuantil",
21                  "Banda inferior 95 %",
22                  "Pendiente OLS"),
23        col = c("gray70", "black", "gray70",
24               "gray40"),
25        lwd = c(2,3,2,3),
26        lty = c(1,1,1,3),
27        bty = "o",
28        bg = "white",
29        cex = 0.6)

```

```

1 plot(data_bogota$log_total, data_bogota$  

2      log_mercado,  

3      pch = 16, col = rgb(0,0,0,0.12), cex  

4      = 0.6,  

5      xlab = "Log Gasto Total",  

6      ylab = "Log Gasto en Mercado",  

7      main = "Regresiones Cuantilicas (Bogota)")  

8  

9 lines(xgrid_bog, y_hat_bog[,1], lty = 3,  

10    lwd = 2) # q = 0.10  

11 lines(xgrid_bog, y_hat_bog[,2], lty = 1,  

12    lwd = 2) # q = 0.50  

13 lines(xgrid_bog, y_hat_bog[,3], lty = 2,  

14    lwd = 2) # q = 0.90  

15  

16 legend("bottomright",  

17        legend = c("q=0.10", "q=0.50",  

18                  "q=0.90"),  

19        lty = c(3,1,2), lwd = 2, col = "black",  

20        bty = "o",  

21        cex = 0.6)

```

5.3. Regresiones cuantilicas para Bogotá

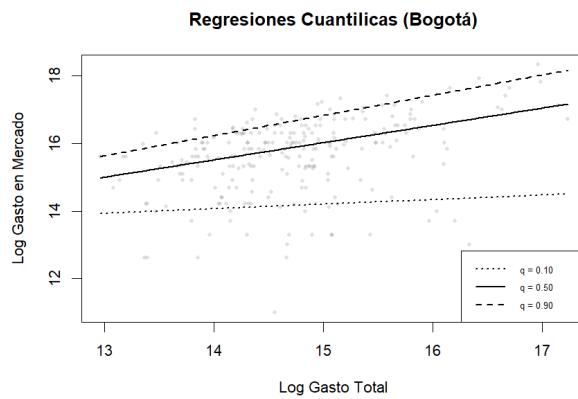


Figura 3: Regresiones cuantilicas ($q = 0.10, 0.50, 0.90$) del gasto en mercado respecto al gasto total para Bogotá.

A mayores cuantiles, la pendiente es más alta, lo que indica que los hogares con mayores niveles de gasto total presentan un gasto en mercado más sensible (mayor elasticidad). En cambio, el cuantil 0.10 casi no crece, mostrando que los hogares de bajo gasto mantienen un gasto en mercado más estable y menos dependiente del gasto total.

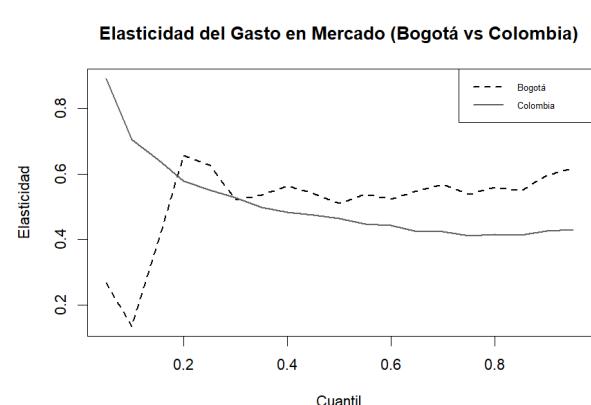


Figura 4: Comparación de elasticidad entre Bogotá y Colombia.

Bogotá muestra elasticidades más bajas en la parte inferior de la distribución, y mayor estabilidad en los cuantiles altos.

```

1 plot(taus, slopes_bog,
2       type = "l", lwd = 2, col = "black",
3             lty = 2,
4       ylim = range(c(slopes_bog, slopes_
5                     col)),
6       xlab = "Cuantil",
7       ylab = "Elasticidad",
8       main = "Elasticidad del Gasto en el
9               Mercado (Bogotá vs Colombia)")
10
11 lines(taus, slopes_col,
12        lwd = 2, col = "gray40", lty = 1)
13
14 legend("topright",
15        legend = c("Bogotá", "Colombia"),
16        col = c("black", "gray40"),
17        lty = c(2,1), lwd = c(2,2),
18        bty = "o",
19        cex = 0.6)

```

6. CONCLUSIONES

- La elasticidad del gasto en mercado no es constante y varía significativamente a lo largo de los cuantiles.
- Bogotá presenta un comportamiento más homogéneo y menor elasticidad en los hogares de bajo gasto.
- La regresión cuantílica proporciona evidencia más rica que OLS, revelando heterogeneidad estructural.

7. RECOMENDACIONES

Los resultados obtenidos mediante regresión cuantílica y el enfoque no paramétrico para la estimación de la curva de Engel permiten derivar varias recomendaciones metodológicas y analíticas relevantes:

- **Incorporar métodos no paramétricos en estudios de consumo.** Dado que la relación entre el gasto total del hogar y el gasto en mercado muestra comportamientos heterogéneos a lo largo de la distribución, se recomienda que futuros estudios utilicen regresión cuantílica u otros métodos no paramétricos, especialmente cuando los datos presentan asimetría, colas pesadas o heterocedasticidad.
- **Extender el análisis a otros dominios geográficos.** La comparación entre Bogotá y el agregado nacional

evidenció diferencias relevantes en las elasticidades. Por ello, sería pertinente realizar estimaciones adicionales para otras ciudades principales y regiones del país, con el fin de identificar patrones de consumo diferenciados.

- **Incluir covariables socioeconómicas.** Factores como tamaño y composición del hogar, nivel educativo, ocupación y ciclo de vida pueden aportar explicaciones importantes sobre la heterogeneidad encontrada. Incorporar estas covariables en modelos cuantílicos permitiría obtener estimaciones más completas de las curvas de Engel.
- **Evitar conclusiones basadas únicamente en OLS.** El análisis mostró que los modelos paramétricos basados en la media pueden ocultar variaciones clave entre hogares. Por lo tanto, se recomienda no depender exclusivamente de modelos OLS cuando se analizan patrones de consumo o elasticidades del gasto.
- **Replicar el análisis para diferentes categorías de gasto.** El enfoque utilizado puede aplicarse a otros rubros relevantes de la canasta familiar, como transporte, educación, salud o vivienda. Esto permitiría caracterizar de forma más integral la estructura del gasto en los hogares colombianos.
- **Aprovechar la riqueza de los microdatos de la ENPH.** Dado que la encuesta contiene información detallada sobre múltiples dimensiones del gasto, se recomienda realizar análisis más desagregados, ya sea por estrato socioeconómico, tipo de hogar o región, para fortalecer la interpretación de las curvas de Engel no paramétricas.

REFERENCIAS

- [1] Cameron, A. Colin, and Pravin K. Trivedi (2005). *Microeconomics: Methods and Applications*. Cambridge University Press.
- [2] Koenker, R., & Bassett, G. (1978). Regression Quantiles. *Econometrica*, 46(1), 33–50.
- [3] Giraldo Henao, R. (2025). *Estadística No Paramétrica: Alternativas a Pruebas Clásicas, Estimación de la Densidad y Análisis de Regresión*. Apuntes de Clase (Documento en Revisión). Departamento de Estadística, Universidad Nacional de Colombia, Sede Bogotá.

Elián Steven Bejarano González. Departamento de Estadística, Universidad Nacional de Colombia, Bogotá, Colombia. E-mail address: elbejaranog@unal.edu.co

Juan Andrés Godoy Reyes. Departamento de Estadística, Universidad Nacional de Colombia, Bogotá, Colombia. E-mail address: jgodoy@unal.edu.co