

La Curva de Engel de los Servicios de Salud En Colombia. Una Aproximación Semiparamétrica¹

Por:

Jorge Barrientos Marín, Juan Miguel Gallego y Juan Pablo Saldarriaga²

–Introducción. –I. Antecedentes de la Literatura en Colombia. –II. Descripción de la Información. –III. Las Curvas de Engel y la Estrategia Empírica. –IV. Resultados empíricos. –Conclusiones. –Bibliografía

Resumen: en este trabajo estamos interesados en estimar curvas de Engel para diferentes rubros de gasto en salud, incluida la salud como un todo. Para llevar a cabo este objetivo, usamos información contenida en la Encuesta de Calidad de Vida de 1997. La técnica usada aquí está basada en modelos semiparamétricos, o Parcialmente Aditivos, combinando estructuras paramétricas y no paramétricas. Adicionalmente, corregimos los problemas de endogeneidad que pueden aparecer y detectamos si son suficientes modelos lineales o cuadráticos. Nuestros resultados indican que el transporte para ir al médico es un bien denominado inferior debido a la forma de la curva estimada. Así mismo la linealidad es fuertemente rechazada en muchos casos, pero el efecto cuadrático (imperfecto) no pudo ser rechazado en el caso del transporte.

Palabras clave: curva de Engel, gasto en salud, modelos semiparamétricos, endogeneidad, bootstrap, instrumentos.

JEL Clasificación: C14, D12, D1

Abstract: in this paper we are interested in estimating Engel curves for different budget share related to health services. To this end we use data drawn for ECV/1997. We based our estimation on additive partial linear models, which mix non-and parametric models. Endogeneity problems could arise in the estimation of Engel curves; in order to correct this problem, we use two stage generated regressors. We perform bootstrap testing procedures for detecting linear and quadratic effect of total expenditure on budget share. Because of shape of transportation for visiting doc, our results suggest that it is a sort of inferior good. Most of the times, linear effects are strongly rejected, but quadratic effects are not rejected for transportation.

Key Words: Engel curve, health expenditure, semiparametric models, endogeneity, bootstrap, instruments.

JEL Classification: C14, D12, D1

¹ Esta es una versión preliminar, por favor no citar.

² Jorge Barrientos Marín: investigador del Centro de Investigaciones y Consultorías-CI&C y Director del departamento de Economía, Universidad de Antioquia. Dirección electrónica: jbarr@economicas.udea.edu.co. Dirección postal: Ciudad Universitaria, bloque 13, A.A. 1226, Medellín, Colombia. Juan Miguel Gallego es estudiante del Doctorado en Economía, Universidad de Toulouse. Dirección electrónica: juan.jgallego@gmail.com. Juan P Saldarriaga es asistente de investigación del Centro de Investigaciones y Consultorías-CI&C. Dirección Electrónica: jpsaldarriaga@economicas.udea.edu.co. Los errores, opiniones e interpretaciones son responsabilidad exclusiva del autor. Los autores agradecen el apoyo financiero del Comité para el Desarrollo de la de Investigaciones-CODI, de la Universidad de Antioquia.

Introducción

El análisis empírico de las curvas de Engel ha sido tema de particular interés desde los estudios seminales de Working (1943), Leser (1963), Deaton y Muelbauer (1980a, 1980b), en los cuales desarrollaron estructuras paramétricas, hasta los más recientes estudios de Lys-souto (2002), Blundell (2003), Barrientos (2006, 2008), los cuales desarrollan análisis estadísticos basados en modelos noparamétricos o semiparamétricos. La mayoría de estos trabajos usan grandes encuestas de ingresos y gastos familiares. En esta investigación hacemos lo propio usando la Encuesta de Calidad de Vida de 1997 (ECV1997), pues a diferencia de la ECV2003, la primera contiene más información sobre rubros de gasto en salud.

En éste trabajo, estamos interesados en estimar la relación que existe entre el gasto total de los hogares en Colombia y la participación del gasto en salud como porcentaje del gasto total. Adicionalmente, estimamos la media condicional, controlando por las características del hogar y del jefe de hogar, como por ejemplo sexo, edad, escolaridad del jefe, estrato socioeconómico, número de personas por habitación y eventualmente consideramos el riesgo de enfermedad por hogar (REH).

La curva de Engel es una herramienta que relaciona, en principio, el ingreso de una persona o familia con el nivel de demanda de determinados bienes. Sin embargo, empíricamente resulta es de mayor interés la relación entre el gasto total y el gasto en determinados rubros, como alimentos, ropa, alquiler, etc. Nosotros extendemos el análisis empírico al gasto total en salud, en primera instancia, y en la medida de lo posible, a diversos rubros específicos de la salud como por ejemplo el gasto en medicinas, vacunas, aseguramiento, afiliación al sistema de seguridad social, hospitalización, odontología, consulta médica, entre otros.

El análisis de las curvas de Engel tiene otras interesantes aplicaciones, que requiere una teoría de la *integrabilidad* más completa (es decir, como desde las demandas individuales es posible recuperar las funciones de utilidad), información estadística más depurada y técnicas de estimación un poco más sofisticadas que las propuestas aquí. Las curvas de Engel permiten, por ejemplo, llevar a cabo análisis de bienestar social, pero también son im-

portantes para realizar estudios de pobreza y distribución de ingresos. Incluso, una curva de Engel eventualmente permite saber si un bien es de lujo, inferior o normal, entre otras categorías de bienes.

Un problema técnico que probablemente enfrentaremos es la endogeneidad de las variables explicativas de la demanda, el cual es muy frecuente en este tipo de estimaciones, en particular, porque el gasto total y el gasto en algún tipo de bien están conjuntamente determinados, lo que produce endogeneidad por simultaneidad. Vale la pena aclarar que la endogeneidad puede también tener origen en la errónea medición de las variables de interés. En cualquier caso, la solución está en las variables instrumentales.

Existen diversos métodos para aplicar variables instrumentales, desde el punto de vista paramétrico se suele utilizar la técnica de mínimos cuadrados en dos etapas, basados en el método generalizado de momentos. Newey y Powell (2003) desarrollaron un método no paramétrico en dos etapas, sin embargo es complicado porque requiere la elección de una base (vectorial) expandible en una primera etapa. Existe otra aproximación menos complicada, e incluso más rápida y eficiente debida a Sperlich (2005) y conocida como regresores generados o construidos en dos etapas. Barrientos (2008) aplica este método en un análisis de consumo de alimentos, ropa y transporte para Bogotá, con resultados satisfactorios.

La contribución de esta investigación puede resumirse como sigue: primero, investigamos el comportamiento, el patrón de consumo, de los agentes estimando un sistema de curvas de Engel en donde se incluyen en las regresiones algunas características del jefe de hogar. Segundo, ajustamos nuestras estimaciones por endogeneidad usando el método de Sperlich (2005), tomando el ingreso corriente de la familia y otras variables explicativas exógenas como instrumentos. Tercero, este trabajo sería el primer análisis noparamétrico de Curvas de Engel en el consumo de servicios de salud, en la literatura nacional e internacional.

La información obtenida de la Encuesta de Calidad de Vida de 1997, es la fuente con la información más completa y desagregada sobre el tema. Si bien, la Encuesta de Calidad de Vida de 2003 es más reciente, ésta no tiene información sobre los mismos rubros en salud

contenidos en la ECV97, la cual contiene 12 rubros de gasto en salud mientras que la ECV2003 incluye solo 4 rubros. Finalmente, se realizan contrastes de hipótesis de linealidad y efectos cuadráticos aplicando contrastes adaptativos. Las bandas de confianza para las regresiones no paramétricas son calculadas usando procedimientos basados en remuestreos (*bootstrap sampling*).

I. Antecedentes a la literatura en Colombia

El financiamiento de la salud en Colombia, como en todos los países del Mundo, es un problema de máxima importancia política y social, con enormes repercusiones sobre el bienestar de la población, el presupuesto del gobierno y los recursos de los hogares y las empresas. Para el caso colombiano, la financiación del sector proviene de fondos del sector público³ y del sector privado, algunos de estos, como las contribuciones al Sistema General de Seguridad Social (SGSSS) obligatorios, y otros voluntarios. La financiación privada a su vez puede provenir de las empresas, el principal rubro son los aportes al SGSSS pero hay empresas con programas adicionales, las ONG y los hogares. En lo que concierne a los hogares, el asunto de la financiación no es trivial, pues de acuerdo a Ramírez, Cortes y Gallego (2003) los hogares colombianos gastan el 9,62% de sus ingresos corrientes en salud, incluyendo sus cotizaciones al SGSSS. Este porcentaje es muy elevado y hace que sea un problema sustancial para la política pública el entender el patrón de este gasto y las variables que están asociadas con él.

El análisis descriptivo realizado en Ramírez, Cortes y Gallego (2003) muestra como el valor del gasto y su distribución por rubros varían según la región del país, el estrato socioeconómico de la familia y el régimen de afiliación al SGSSS. El componente de gastos en salud es un rubro muy importante dentro del consumo de los hogares colombianos, a saber, un hogar en 1997 con ingresos de \$500.000 mil pesos mensuales destinaba, en promedio, cerca de \$50.000 pesos para suplir sus necesidades en salud. Esta cifra entra a competir con otras necesidades del hogar como educación de sus miembros, alimentación, vivienda, recreación entre otras.

³ El análisis de la financiación pública se ha realizado en diversas oportunidades utilizando como fuentes de información los presupuestos y ejecuciones presupuestales de diversas agencias del Gobierno.

Ramírez, Cortes y Gallego (2003) muestran que en la mayoría de los rubros de gasto en salud la población se concentra cerca al gasto nulo, es decir, no gasta o gasta muy poco, y alrededor de cierto nivel de gasto. En consecuencia, las estimaciones de gasto no pueden ser explicadas a través de modelos paramétricos. Esto implica que en materia metodológica es importante desarrollar otros tipos de estudios que vallan más allá de estimaciones paramétricas. Por otro lado, para el caso del gasto en salud debe indagarse más por la distribución de los datos al interior de la ECV97; estos desarrollos deben ir encaminados principalmente a la implementación de pruebas de hipótesis y test de especificación a partir de las estimaciones no paramétricas.

El marco teórico de la investigación es la teoría microeconómica neoclásica,⁴ que es el marco de referencia, implícito o explícito de la mayor parte de las estimaciones de funciones de demanda, consumo o gasto. De acuerdo con esta teoría el consumidor se enfrenta a dos tipos de restricciones: la primera esta dada por las características personales, edad, sexo, educación, conocimiento de alternativas, actividad económica, clima en el que vive etc.; las cuales determinan lo que se ha denominado en la literatura ‘conjunto de posibilidades de consumo’; el otro tipo de restricciones es el dado por los recursos con los que cuenta el hogar (ingreso, riqueza) y por los precios que en el mercado tienen los bienes que va a consumir, ingreso y precios determinan el ‘conjunto de presupuesto’. De acuerdo con la teoría microeconómica, el consumidor elige la alternativa (lista exhaustiva de todas las cantidades de bienes y servicios) que prefiere dentro de las que satisfacen sus dos tipos de restricciones.

La solución de este problema de optimización bajo restricciones viene dada por la llamada función de demanda, en la cual la cantidad demandada del bien es función del ingreso, de los precios y de las características de los consumidores. Si se multiplica esta función por el precio del bien en cuestión se obtiene la llamada función de gasto. En este proyecto se propone estimar, independientemente, ecuaciones relacionadas con el gasto en salud.

⁴ Para una presentación completa de esta teoría ver Mas Collé, Whinston y Green. (1995). Para aplicaciones al análisis de consumo de los hogares ver Deaton.(1997)

Adicionalmente, en Colombia en 1993 se implementó una fuerte reforma del sector salud lo cual determinó y afectó los patrones de gasto en salud. En este sentido, al estimar la variación de gasto en los distintos rubros en salud es necesario controlar por variables institucionales del SGSSS, como por ejemplo el tipo de afiliación del jefe de hogar, lo cual es importante para entender el nivel de gasto y su distribución. Adicionalmente, muchos de los gastos en salud, como por ejemplo gastos en medicamentos, transporte, y servicios de especialistas son gastos derivados de otros gastos en salud como gastos en consulta médica y aseguramiento. Por lo tanto, hacer análisis no paramétricos ayudaría a entender las estructuras de gasto y por ende poder tener elementos técnicos para analizar a futuro las diferentes propuestas de reforma a la ley 100 que el gobierno y el congreso pretendiesen implementar.

II. Descripción de la información

Tabla 1. Estadísticas Descriptivas

Variables	Media	Desv. Est	Min	Max
Hab. Hogar	3.2	1.5	1	14
Num. Pers. Hogar	4.2	2.1	1	19
Num. Pers/Hab.	1.6	1.2	0.08	13
REH	1.06	0.27	0.62	1.88
G. Afiliación S.S	33731.31	36314.51	1100	510390
G. Hospital.	30246.51	73377.86	1025	1333333
G. Medicinas	31599.97	47572.65	1100	715000
G. Transp. Vis. Dr.	11909.71	22394.27	1040	403000
G. Salud	349602.2	694185.7	10040	2.37e+07
G. Total	3422315	5079345	10460	1.10e+08
Ingreso Cte. Hogar	17722.28	10522.35	10058.33	52729.33
Escolaridad J.H.	5.5	4.5	0	20
Sexo. J.H	0.75	-----	0	1
Edad. J.H	46.8	15.3	13	98
Ocupación J.H_1	0.25	-----	0	1
Ocupación J.H_2	0.12	-----	0	1
Ocupación J.H_3	0.23	-----	0	1
Ocupación J.H_4	0.04	-----	0	1
Ocupación J.H_5	0.12	-----	0	1
Ocupación J.H_6	0.22	-----	0	1
Estrato 1	0.13	-----	0	1
Estrato 2	0.12	-----	0	1
Estrato 3	0.4	-----	0	1
Estrato 4	0.22	-----	0	1

Estrato 5	0.04	-----	0	1
Estrato 6	0.02	-----	0	1
Estrato 7	<0.01	-----	0	1
Tipo Afil. S.S 1	0.39	-----	0	1
Tipo Afil. S.S 2	0.12	-----	0	1
Tipo Afil. S.S 3	0.22	-----	0	1
Tipo Afil. S.S 4	0.22	-----	0	1
Tipo Afil. S.S 5	0.027	-----	0	1
Tipo Afil. S.S 6	0.003	-----	0	1
Atlántica	0.19	-----	0	1
Oriental	0.16	-----	0	1
Pacífica	0.18	-----	0	1
Central	0.15	-----	0	1
Antioquia	0.17	-----	0	1
Bogotá	0.08	-----	0	1
Orinoquia	0.04	-----	0	1

Fuente: Encuesta Calidad de Vida 1997. Cálculos propios.

La tabla 1 muestra las estadísticas de las variables relevantes usadas en esta investigación. De acuerdo con la descripción, los hogares están distribuidos en esta muestra del siguiente modo: 19% en la región atlántica, 18% en la pacífica, 17% en Antioquia, 16% en la oriental, 15% en la Central, 8% en Bogotá D.C., que es tomada como una región, y 4% en la Orinoquia. En las siete regiones, el número de personas y habitaciones por hogar es de 4.2 y 3.2 respectivamente, lo que da 1.6 personas por habitación. Por otro lado, cerca del 40% de los hogares se encuentran en estrato 3, 22% está en estrato 4, 13% en estrato 1, 12% en estrato 2, 2% en estrato 5 y menos del 1% en estrato 6.

Los hogares cuentan con un jefe de hogar, en los cuales el 75% son hombres con una edad que oscila entre los 46 y 47 años con un promedio de 5.5 años de educación, ciertamente muy bajo comparado con otros países de Latino América, aproximadamente 7.15 años de acuerdo a Cuesta, Ñopo y Pizzolito (2007). Alrededor del 25% de los jefes de hogar que reportó estar trabajando (en la última semana) como obrero o en el sector público, el 23% es profesional independiente, el 26% afirmó ser desempleado o pensionado o estar estudiando, el 12% dijo trabajar como jornalero en su propia finca o ser empleada del servicio, mientras que solo el 4% afirmó ser patrono o empleador.

Una variable que es importante en la determinación de la demanda por servicios de salud es el tipo de afiliación que tienen los jefes de hogar al SGSSS. La distribución de los hogares por tipo de afiliación indica que el 39% se encontraba en el régimen subsidiado, y solo el 12% estaba en el régimen contributivo, el 22% no estaba afiliado y otro 22% afirmó estar en el Instituto de Seguros Sociales (ISS), cerca de un 3% pertenece a regímenes especiales (Fuerzas Armadas (FF.AA), Magisterio, Ecopetrol, etc.) y menos del 1% estaba en otro tipo de entidades. Cabe señalar, que uno de los problemas que se pudo identificar, con esta muestra, es que para 1997 había un alto porcentaje de personas sin ningún tipo de protección en salud, lo cual ha sido una debilidad desde la reforma a la salud de 1993.

Una variable que incluimos en el estudio, aunque puede ser discutible su efecto desde el análisis económico de la salud, es el riesgo de enfermedad por hogar (REH), en principio la motivación de su construcción fue capturar la estructura epidemiológica de del hogar, basados principalmente en la edad y el sexo de sus miembros, la REH está dada por la siguiente relación:

$$REH = \frac{\sum_i \alpha_i I_i}{\sum_i I_i} \quad (1)$$

Donde α_i es el factor epidemiológico del grupo u hogar i , e I es el número de individuos en el hogar que pertenecen al grupo epidemiológico. Dicho grupo está definido por el Ministerio de Salud basado en el Plan Obligatorio de Salud (POS). Esta variable es usada como instrumento principalmente; es probable que en algunos casos haya un problema de identificación si la incluimos en un modelo paramétrico con otras variables de estructura de los hogares.

La ECV1997 tiene varios componentes del gasto en salud, pero debido a que muchos encuestados no contestaron o respondieron gasto igual a cero, esto genero una pérdida significativa de información, de modo que solo pudimos tener en cuenta dentro de los rubros de gasto en salud el gasto en afiliación al sistema de seguridad social en salud, hospitalización,

medicinas, transporte para ir al médico, el gasto total en salud, la suma de los anteriores más otros gastos en salud y el gasto total del hogar.⁵

III. Las Curvas de Engel y la Estrategia Empírica

A. Las Curvas de Engel

La curva de Engel define la relación existente entre el gasto o consumo en un bien específico y el ingreso de los agentes. Empíricamente la relación que se establece es entre el gasto en un bien específico y el gasto total. La mayoría de las veces los investigadores suelen confiar más en la cifra de gasto que en la de ingreso que son reportadas por los encuestados. Desde los artículos de Working (1943), Leser (1963), Deaton y Muellbauer(1980a, 1980b) , Banks y otros (1997), Blundell y otros (1998, 2003), Bhalotra y otros (1998), Lysiotou y otros (2002) y Barrientos (2006, 2008) la relación empírica que se establece viene dada por:

$$w_{ij} = \frac{x_{ij}}{x_i} = g_j(\ln X_i, Z_i) + u_i \quad i = 1, \dots, N \quad (2)$$

Donde w_{ij} es la participación del gasto en el bien j en el gasto total del hogar i , $\ln X_i$ es el logaritmo del gasto total del hogar i , Z_i es un conjunto de variables socio-demográficas como género, edad y escolaridad del jefe de hogar, estrato del hogar, etc., $g(\cdot)$ es una función de gasto, la cual puede ser paramétrica, noparamétrica o una combinación de ambas, esta última relación da lugar a las relaciones semiparamétricas, la cual es la estructura de interés en esta investigación, y u_i es el termino de perturbación estocástica el cual se asume tal que $E(u|X, Z) = 0$. Debido a problemas con la escala de medición del gasto total y el gasto en bienes en este trabajo definimos la participación como:

$$W_{ij} = \frac{\ln x_{ij}}{\ln x_i} \quad (3)$$

⁵ Para el análisis de regresión los tamaños de muestra varían dependiendo del tipo de gasto que se tenga en cuenta.

La definición dada por la expresión (3) tiene la ventaja de reducir la escala, de modo que las rutinas del software (GAUSS) son más eficientes, pero no cambia el significado, pues son transformaciones monótonas y además permiten graficar la participación del gasto en los rubros contra el logaritmo del gasto total, cuando se lleva a cabo regresiones semiparamétricas o no paramétricas.

B. La Estrategia Empírica

1. Modelos Parcialmente Lineales

Nuestra estrategia empírica consiste recuperar la media condicional, $E(w|X, Z)$, basados en el siguiente modelo empírico:

$$w_t = \alpha + \beta Z_t + m(\ln X_t) + u_t \quad t = 1, \dots, N \quad (4)$$

Esta especificación tiene dos ventajas sobre los modelos paramétricos y no paramétricos puros, de un lado la parte lineal permite incluir variables que típicamente tienen un efecto lineal, como lo es el género del jefe de hogar u otras variables categóricas, y de otro lado, permite reducir el efecto perverso de un número de regresores crecientes que entran no paramétricamente y que es conocido como la “maldición” de la dimensionalidad.

Para obtener un estimador del parámetro α y β se usa la aproximación de Robinson (1988) del siguiente modo. La expresión (4) puede escribirse como:

$$\varpi = m(\ln X) + \xi \quad (5)$$

donde $\varpi = w - Z\beta$ y $\xi = u + Z(\beta - \hat{\beta})$ es un nuevo error compuesto para toda observación. El intercepto puede ser estimado como $\hat{\alpha} = \bar{\varpi} - \hat{\beta}\bar{Z}$ donde $\bar{\varpi}$ y \bar{Z} son las medias muestrales de las participaciones en el gasto y las variables socio-demográficas. De modo que la ecuación (5) puede estimarse consistentemente usando el procedimiento desarrollado por Nadaraya (1964) y Watson (1964).

Volvamos al problema mencionado en la introducción relacionado a la endogenidad por simultaneidad que se presenta en la estimación de las curvas de Engel. El procedimiento para la solución de dicho problema funciona del siguiente modo: sea X una variable endógena, con todo lo que ello implica, y sea W un conjunto de instrumentos (posiblemente contruidos con las variables exógenas Z e incluso con variables fuera de la base de datos original). La idea se basa en hacer una regresión noparamétrica (o incluso semiparamétrica) de $\ln(X)$ sobre W en una primera etapa, por ejemplo:

$$\widehat{\ln(X)} = f(W) \quad (6)$$

En un segundo paso, en lugar de usar $\ln(X)$ en la expresión (5), usamos nuestro regresor construido por medio de la regresión dada por (6). Este procedimiento es rápido y fácil, con el único costo de una regresión noparamétrica en un primer paso. Una explicación más detallada del procedimiento empleado aquí se encuentra en Sperlich (2005).

2. Contrastes paramétricos contra noparamétricos

Nosotros estamos interesados en contrastar la hipótesis de linealidad (o de efectos cuadráticos), esto depende de la forma obtenida al estimar la función dada por (5). La hipótesis nula se define como $H_0: m(x) = m_\theta(x)$ para algún θ , con $m_\theta(x) = \theta x^q$ donde $q=1$ y $q=2$ denotan relaciones cuadráticas o lineales entre el gasto total y la participación en el gasto.

Estos contrastes pueden llevarse a cabo usando el método de Hardle y otros (2004). Con este objetivo empleamos tres tipos diferentes de estadísticos de contrastes. El primero está basado en la diferencia al cuadrado del modelo parcialmente lineal y el paramétrico. También se considera los estadísticos desarrollados por Gozalo y Linton (2001) y Rodríguez-Poo y otros (2005), dados por:

$$\hat{\epsilon}_1 = \frac{1}{n} \sum_t (\hat{m}(X_t) - \hat{m}_\theta(X_t))^2 p(X_t) \quad (7)$$

$$\hat{\epsilon}_2 = \frac{1}{n} \sum_t \hat{u}_t (\hat{m}(X_t) - \hat{m}_\theta(X_t)) p(X_t) \quad (8)$$

$$\hat{\epsilon}_3 = \frac{1}{n} \sum_t \left(\frac{1}{n k^d} \sum_j K_n(X_t - X_j) (w_j - \hat{m}_\theta(X_j)) \right)^2 p(X_t) \quad (9)$$

Donde $\hat{u}_t = w_t - \hat{m}_\theta(X_t)$, son los residuales para el modelo bajo la hipótesis nula, d es la dimensión de la matriz X y h es el parámetro de sobre-suavización para la estimación. Ahora enunciemos el procedimiento para calcular valores críticos basados en la idea de contrastes adaptivos de Horowitz y Spokoiny (2001). Basados en las estimaciones paramétricas y semiparamétrica construimos test estadísticos cuya distribución varía con el parámetro de suavidad del estadístico de contraste (diferente al parámetro de suavidad para la estimación) denotado por k ,

$$\hat{t}_{jk} = \frac{\hat{\epsilon}_{jk} - \hat{\beta}_j}{\hat{\sigma}_j} \quad j=1, 2, 3 \quad (10)$$

Donde $\hat{\beta}_j$ y $\hat{\sigma}_j$ son la media y la desviación estándar estimadas de \hat{t}_{jk} para $j=1, 2, 3$. Entonces nosotros calculamos el estadístico basado en los datos remuestrados (bootstrap) denotados por,

$$\hat{t}_{jk}^* = \frac{\hat{\epsilon}_{jk}^* - \hat{\beta}_j^*}{\hat{\sigma}_j^*} \quad j=1, 2, 3 \quad (11)$$

Este procedimiento crea una familia de estadísticos de contraste $\{\hat{t}_{jk}, k \in K_n\}$ donde la elección de k marca la diferencia entre la hipótesis nula y la alternativa. El estadístico de contraste que finalmente se usa esta dado por,

$$\hat{t}_{jk}^{**} = \max_{k \in K_n} \hat{t}_{jk}^* \quad (12)$$

Donde $K_n = \{k: k = c_l n^{-1/B}, l = 1, \dots, L\}$ define el conjunto admisible de parámetros de suavización, con $c_l = [l + (c_X(l-1)^{-1})] n^{-1/B}$, $c_X = C(\max X_t - \min X_t)$ y la contante $C \in (0,1)$. El procedimiento de contraste rechazará la hipótesis nula si al menos uno de los $k \in K_n$ (de los estadísticos originales) es mayor que su análogo obtenido mediante bootstrap. Un procedimiento similar aplicado a curvas de Engel se puede hallar en Barrientos (2007, 2008).

IV. Resultados empíricos

En este estudio consideramos cuatro rubros de gasto relacionados con el sector salud, el gasto total en salud, el gasto en seguridad social (pagos al sistema), el gasto en medicinas, el gasto en hospitalización y el gasto en transporte privado para ir al médico (no se encuentra el gasto privado para ir al trabajo o asuntos personales, ni el gasto en combustibles). Todos los rubros de gasto en ingresos están a precios constantes anuales de 1997.

En muchos rubros de gasto los encuestados reportaron gasto igual a cero, aplicando algunas restricciones a la información esto nos deja con una muestra variada dependiendo del tipo de gasto, por ejemplo, para analizar el gasto total en salud tenemos 5766 hogares y para el gasto en afiliación, medicinas, transporte y hospitalización tenemos 3577, 3770, 1669 y 2517 hogares respectivamente. En cada caso la muestra es suficientemente grande no solo para hacer estimaciones Semiparamétricas sino para los contrastes propuestos anteriormente. Todas las regresiones controlan por factores socioeconómicos del hogar (estrato, riesgo por enfermedad del hogar y región) y por las características del jefe de hogar (genero, edad, escolaridad, tipo de afiliación y de ocupación al sistema de seguridad social).

A. Curvas de Engel Semiparamétricas

Cada figura a continuación presenta la curva estimada junto con bandas al 95% de confianza. Errores normales son usados para generar las muestras bootstrap y calcular tanto las bandas como los valores críticos para los contrastes de hipótesis. Es valioso hacer notar que los parámetros de suavización (en adelante bandwidths y denotados por h_{CV}) para la estimación y contrastes son muy diferentes. El primero es obtenido por medio de selección automática *leave-one-out cross validation* (validación cruzada), mientras que el segundo es obtenido mediante métodos adaptivos.

En todos los casos las regresiones no paramétricas son hechas usando *quartic kernel* definido como $\frac{15}{16}(1 - x^2)^2 I(|u| \leq 1)$ donde $I(\cdot)$ es una función indicadora. El mismo conjunto

de instrumentos, W , fueron usados para todas las regresiones: el ingreso familiar, la edad y escolaridad del jefe de hogar y el riesgo por enfermedad del hogar (REH).

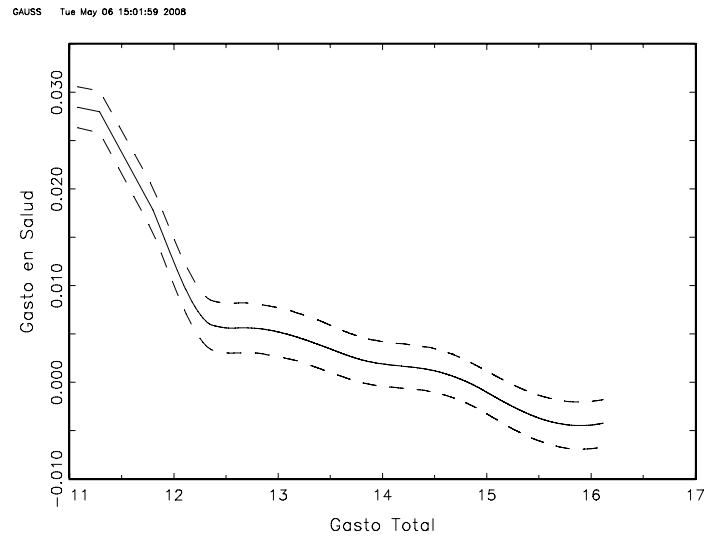


Figura 1: *Gasto total en salud v.s. gasto total*, $h_{CV} = 1.3$

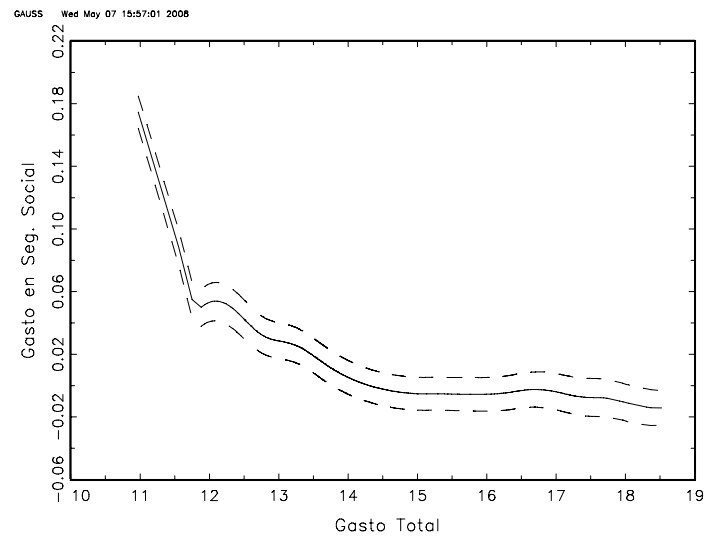


Figura 2: *Gasto en afiliación a la S.S v.s. gasto total*, $h_{CV} = 0.9$

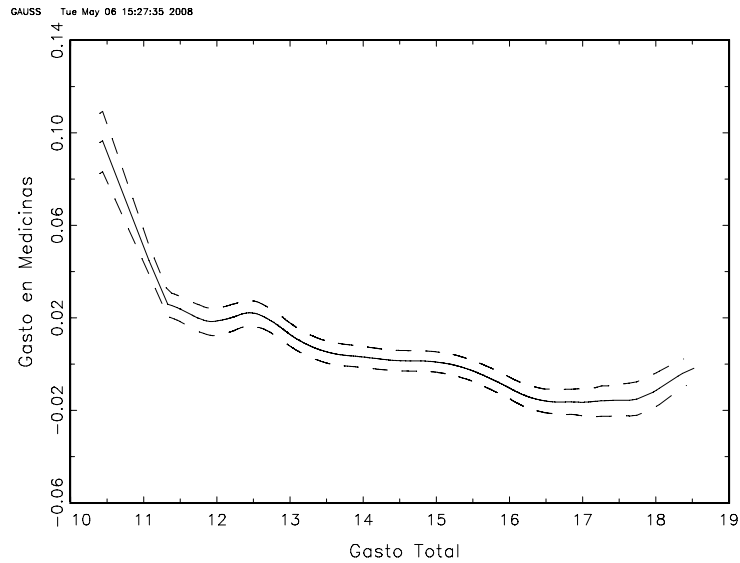


Figura 3: *Gasto en medicinas v.s. gasto total*, $h_{CV}=0.9$

La Figura 1 muestra la estimación semiparamétrica de la curva de Engel correspondiente al gasto (total) en salud. El eje vertical se refiere a la participación del gasto total en salud ajustado por los factores socioeconómicos y del jefe de hogar. Esta estimación, así como las correspondientes a las curvas para el gasto en seguridad social (Figura 2) y gasto en medicinas (Figura 3), son decrecientes, con una mayor pendiente para los hogares que menos gastan y una mayor pendiente para los de ingreso más alto. En este sentido, los hogares más pobres podrían estar sacrificando servicios de salud para dedicar prioridades en alimentación y transporte.

Probablemente, lo anterior indique que las familias no dedican demasiado de su ingreso al gasto en afiliación al sistema de seguridad social en salud, de hecho la información revela que cerca del 22% de los jefes de hogar no está afiliado y el 39% está en el régimen contributivo. Es interesante notar que tanto el gasto en afiliación como en medicinas sigue un patrón similar al gasto total. De hecho, en esta muestra, cerca del 40% del gasto total en salud corresponde a gasto en afiliación y medicinas.

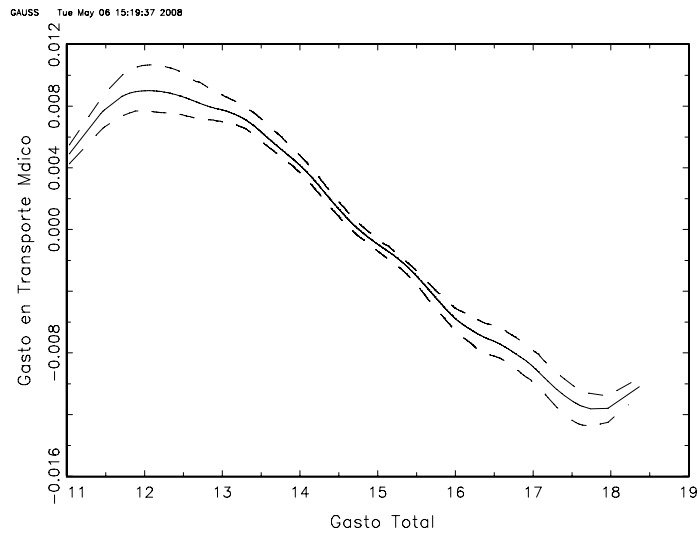


Figura 4: *Gasto en transporte médico v.s. gasto total, $h_{CV}=1.5$*

La Figura 4 muestra la curva de Engel para el gasto en transporte exclusivo para ir al médico. De acuerdo al tipo de bien, transporte público, cuando el ingreso se incrementa las personas dejan de usarlo. Esta forma para la curva de Engel puede indicar que el transporte es en general un bien inferior, en el sentido de que cuando el ingreso de los hogares mejora las personas deciden emplear transporte privado (vehículo propio).

Más interesante es destacar que para el caso de la comunidad de Madrid (España) y para Bogotá D.C, la curva de Engel para el transporte (público) presenta la misma estructura semi-cuadrática. Es decir, la naturaleza inferior del transporte público es manifiesta con independencia del grado de desarrollo, ver Barrientos (2007, 2008).

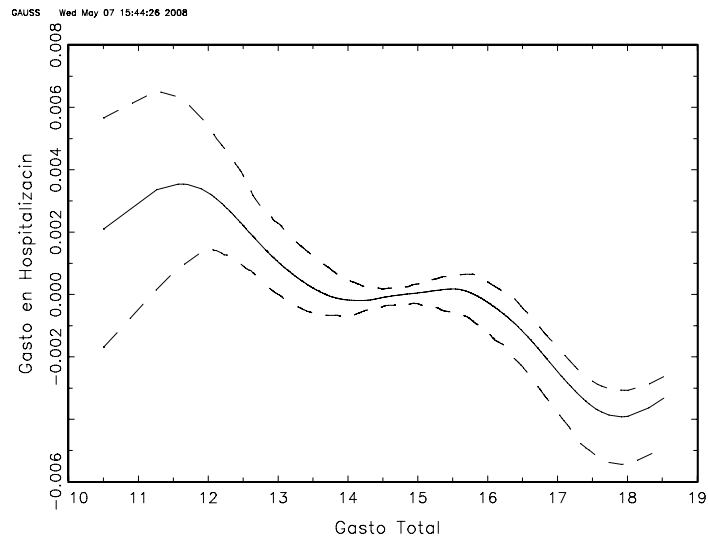


Figura 5: *Gasto en Hospitalización v.s gasto total*, $h_{CV} = 2.8$

La Figura 5 muestra la forma de la curva de Engel para el gasto en hospitalización. A lo largo de la muestra está tiene una tendencia decreciente e indica que la hospitalización es una carga que no es uniformemente distribuida, de hecho para algunos hogares de ingresos medios altos es constante, pero en general para los de mayores ingresos es decreciente.

B. Efectos lineal o cuadrático

Si los modelos se han elegido correctamente los resultados cuantificarán el grado en el cual las variables afectan el comportamiento del consumidor. Es importante resaltar si las especificaciones lineales o cuadráticas en un modelo paramétrico puro son suficientes para recuperar la media condicional subyacente. En esta sección llevamos a cabo estos contrastes de hipótesis. De acuerdo a lo expuesto en la Sección III.B.2 usamos el estadístico de contraste (12) definiendo un conjunto de parámetros de suavización K_n de cardinalidad $L=10$, las bandas están en un rango de 0.24 a 0.87.

Como mencionamos anteriormente, los bandwidths óptimos para estimar la función de regresión son obtenidos por validación cruzada. Es también importante destacar que el bootstrap para contrastes de hipótesis es generado de manera diferente a aquella para generar

contrastes de hipótesis. En la práctica tenemos que sobre-suavizar para los contrastes y sub-suavizar para bandas de confianza (Barrientos 2006).

Usamos 1000 muestras bootstrap para los contrastes de hipótesis, de acuerdo al algoritmo en la sección III.B.2. Puesto que los parámetros de sobre-suavizado son cruciales, los contrastes se elaboran con varios parámetros con resultados muy similares, de modo que se reportan los p -valores para uno de estos parámetros.

De acuerdo a los procedimientos desarrollados, en el caso de gasto total en salud, gasto en afiliación y gasto en medicinas, los tres estadísticos de contrastes rechazan la hipótesis de linealidad al 7% de significancia. De otro lado, el efecto lineal en el caso del transporte para ir al médico es fuertemente rechazado con todos los estadísticos. Sin embargo, ni el estadístico de contraste (8) y (9) son incapaces de rechazar el efecto cuadrático (20% y 90% respectivamente). Estos resultados están acordes con la forma de la Figura 4. Respecto al gasto en hospitalización los resultados indican que la linealidad es rechazada (al 9% en promedio) con todos los estadísticos y el efecto cuadrático no es rechazado con los estadísticos (8) y (9) (al 25% en promedio). Esto nos da una idea de lo confiable y robusto de los resultados.

Conclusiones

Los resultados generales indican que la modelación del efecto del gasto total sobre el rubro de gastos específicos es un poco más complicada de lo que supone en los análisis paramétricos, e incluso en los análisis semiparamétricos en una sola dimensión. Como se han preguntado diversos autores ¿podría ser la no linealidad detectada con las pruebas de hipótesis un problema de especificación del modelo? En el futuro, esto debería ser incluido en un análisis más sofisticado.

Es necesario incluir más variables socio-demográficas para capturar la variedad en el consumo del agente promedio. Las preferencias y los gustos de los consumidores cambian con el tiempo, este es un punto esencial del análisis de curvas de Engel, sin embargo, existe una

fuerte limitante en la información existente debido a la ausencia de información estadística para formar datos de panel. Quizá, con una información más completa se pueda hacer inferencia sobre el tipo de bien que es la salud en un país desarrollado.

Finalmente, las elasticidades juegan un papel primordial para conocer las características de los bienes (normales, inferiores, Giffen, de lujo, etc). No obstante, esto no es fácil pues requiere no solo estimar una función desconocida sino su primera derivada, pero sería muy útil, pues por definición ésta debe cambiar entre hogares.

Bibliografía

BHALOTRA, Sonya y Attfield, Cliff (1998) "Intrahousehold Resources Allocation in Rural Pakistan: A Semiparametric Analysis". *Journal of Applied Econometrics*, 13, No 5, 463-480.

BANKS, J. Blundell, R., y Lewbel, A. (1997). "Quadratic Engel Curves and Consumer Demand". *The Review of Economics and Statistics*, 79, No 4, 527-539.

BARRIENTOS, Jorge (2008) *On the Consumer Behaviour in Urban Colombia: the Case of Bogotá. (Submitted)*.

_____ (2007) *Estimation and Testing Additive Partially Linear Model in a System of Engel Curves*. IVIE Working Paper, WP-AD 2006-23.

BARRIENTOS, M., Jorge y Sperlich Stefan (2006). "The Size Problem of Kernel Based Bootstrap Tests when the Null is Nonparametric" (Submitted).

BARRIENTOS, M., Jorge (2005). "A Note On Bandwidth Choice when the Null Hypothesis is Semiparametric". *Revista de Economía del Rosario*, 8, No 2, 114-129, 2005.

Blundell, R., Duncan, A., y Pendakur, K. (1998). "*Seminparametric Estimation and Consumer Demand*". *Journal of Applied Econometrics*, 13, No 5, 435-461.

CUESTA, JOSÉ; ÑOPO, HUGO Y GEORGINA, PIZZOLITO. (2007) "Using Pseudo-Panels to Measure Income Mobility in Latin America. Working Paper, IDB and World Bank.

BLUNDELL, Richard; Browning, Martin y Ian A. Crawford (2003). "Nonparametric Engel Curve and Revealed Preference". *Econometrica*, 71, No 1, 205-240.

DEATON, A y J. Muellbauer (1980a). "An Almost Ideal Demand System". *American Economic Review*, **70**, 321-326.

_____ (1980b). *Economic and Consumer Behavior*. Cambridge University Press, Cambridge.

DEATON, Angus. *The Analysis of Household Survey. A Microeconometric Approach to Development Policy*. Johns Hopkins University Press, 1997.

LESER, C. E. V (1963). "Form of Engel Functions". *American Economic Review*, 31, No 4, 694-703.

GOZALO, P. L. and O. B. Linton (2001) "Testing Additivity in Generalized Nonparametric regression models with Estimated Parameters". *Journal of Econometrics*, 104: 1-48.

HARDLE, W., Huet, S., Mammen, E., and Sperlich, S (2004) "Semiparametric Additive Indices for Binary Response and Generalized Additive Models". *Econometric Theory*, 20, 265-300.

HOROWITZ, J, L and V. Spokoiny (2002) "An Adaptive, Rate-optimal Test of Parametric Mean-Regression Model Against A Nonparametric Alternative". *Econometrica*, 69, No. 3, 599-631.

LESER, C. E. V (1963). "Form of Engel Functions" *Econometrica* ,31, No 4, 694-703.

LYSSIOTOU, P; Pashardes, P y Stengos, Thanasis (2001). "Age Effects on Consumer Demand: An Additive Partially Linear Regression Model". *The Canadian Journal of Economics*, 35, No 1, 153-165.

MASCOLLEL, Whinston y Green. *Microeconomic Theory*. Oxford University Press, 1995.

NADARAYA, E. A. (1964). "On Estimating Regression". *Theory Probability Applied*, 10.

NEWAY, W. K y J. Powell (2003) "Instrumental Variables Estimation of Nonparametric Models". *Econometrica*, 71, 1565-1578.

RAMIREZ, Manuel; Cortes, Darwin y Gallego, Juan (2002). "El gasto en Salud de los Hogares Colombianos: un Análisis Descriptivo". *Lecturas de Economía*. Julio-Dic, No 57.

ROBINSON, P (1988) "Root N -Consistent Semiparametric Regression". *Econometrica*, 56, 931-54.

RODRIGUEZ-POO, J. M, S. Sperlich and P. Vieu (2005) "And Adaptive Specification Test For Semiparametric Models". Working Paper, University of Carlos III de Madrid.

SPERLICH, S. (2005). "A Note on Nonparametric Estimation with Constructed Variables and Generated Regressors". Working Paper, University of Carlos III de Madrid.

WATSON, G. S (1964) "Smooth Regression Analysis". *Sankhyā Ser. A* 26.

WORKING, H. (1943) "Statistical Laws of Family Expenditure". *Journal of the American Statistical Association*, 38, 4-56.