Introduction to Consumer Theory

Mora Rodriguez, Jhon James

Universidad icesi

8 July 2013

Introduction to Consumer Theory

By

Jhon James Mora, PhD. ©
Icesi University

Abstract: This book shows the most recent topics in economic consumer theory. Also the book shows a review of econometric test of some topics in estimation of microeconometric models.

2013

INTRODUCCIÓN A LA TEORÍA DEL CONSUMIDOR

2013

Jhon James Mora

Tabla de contenido

Índice de Gráficas

Índice de tablas

Prólogo a la edición 2013

Desde la publicación de la primera versión en el 2002 del libro de Introducción a la Teoría del consumidor he recibido varios correos acerca de publicar la segunda versión de este libro.

Sin lugar a dudas, nuevos libros sobre Microeconometria como el de Cameron y Trivedi (2005) han cubierto muchos de los campos que habría abordado en una versión más reciente y que me llevaron a la decisión de no escribir una segunda versión.

En lugar de ello, he decidido publicar una versión revisada del mismo y tratar de incorporar algunos ado-file a fin de que sea útil bajar esta versión. Es así como, la presente versión revisada corresponde a la necesidad de ofrecer una versión on-line a las ya existentes en GOOGLE que son disponibles hoy día en diferentes páginas y ofrecer datos y archivos tipo ado a fin de replicar algunas de las técnicas más comunes usadas hoy día en econometría.

Agradezco mucho de los comentarios que me han enviado de la primera edición así como los comentarios en “SAN GOOGLE” del mismo. En particular deseo rescatar:

> *“Un excelente trabajo del Profesor Jhon James Mora del Departamento de Economía de la Universidad de ICESI en Cali, Colombia. 218 páginas en formato PDF. Trata toda la teoría del consumidor. Recomendable. http://www.microeconomia.org/moodle17/mod/resource/view.php?id=14”*

Algunos de los sitios donde se puede bajar la primera edición son:

http://www.eumed.net/libros-gratis/2005/jjm/libro-jjmora.pdf
http://books.google.es/books/about/Introducción_a_la_teor%C3%ADa_del_consumido.html?hl=es&id=-YoXBWRtHlgC
http://www.kilibro.com/book/preview/121267_introduccion-a-la-teoria-del-consumidor
http://www.microeconomia.org/documentos_new/libro-jjmora.pdf
http://isis.faces.ula.ve/computacion/emvi/libreria/2005/jjm/index.htm
http://www.slideshare.net/derecholaboralprivado/introduccion-a-lateoriadelconsumidordelapreferencia
http://www.librosgratis.me/75-libros-de-economia.html
http://www.carlospitta.com/Courses/Microeconomia/PDF/Madrid/libro-jjmora.pdf
http://www.buenastareas.com/ensayos/Introduccion-A-La-Teoria-Del-Consumidor-De-La-Preferencia/23668659.html
http://danieleconomist.blogspot.com/2010/02/introduccion-la-teoria-del-consumidor.html
http://librosfull.com/2008/09/12/introduccion-a-la-teoria-del-consumidor/
http://www.extpdf.com/jhon-pdf.html

Quedo en deuda con los otros portales en los cuales se puede bajar.

Prólogo a la edición 2002

El libro de Deaton, A y Muellbauer, J.(1980) *Economics and Consumer Behavior*, Cambridge, Cambridge U.Press, consiste en el punto de partida de lo que aquí se expondrá. La razón es muy simple, el libro de Deaton y Muellbauer es pionero en integrar la teoría del consumidor a los desarrollos en econometría, razón por la cual, siempre será de consulta fundamental. Desarrollos posteriores a la obra de Deaton y Muellbauer, como los modelos de autoselección de Heckman o las aplicaciones a la teoría medioambiental, han sido tratados en libros como el de Maddala (1983) o en los libros sobre medioambiente, como Azqueta (1994) *Valoración económica de la calidad ambiental*, Mc Graw-Hill (España) y, en pocas ocasiones, han sido tratados como el avance inevitable de la aplicación econométrica de la teoría del consumidor. Dos libros más merecen la atención del lector: el libro de Mas-Collel, A., Whinston, M.D y Green, J.R.(1995), *Microeconomic Theory*, Oxford University Press y el libro de Anderson, S.P., Palma, A y Thisse, J.F.(1995), *Discrete choice theory of product diferentiation*, MIT Press, Cambridge Mass. La impecable presentación de la teoría del consumidor realizada por Mascollel et-al hace de este libro una invaluable fuente de consulta. Por otro lado, el libro de Anderson et-al introduce al lector en los modelos de elección discreta presentando una gran ayuda en el análisis empírico.

Jhon James Mora R., PhD.
Profesor de Microeconometría
Universidad Icesi
Cali, Colombia

Introducción

Siendo las 6 de la mañana suena el reloj despertador. Usted se levanta, prepara su desayuno, y conduce hacia el trabajo. A las 8 A.M. se encuentra en la oficina y su mejor compañero de trabajo es Linda. A eso de las 11, toma un descanso y empieza a reflexionar sobre la mañana. Se da cuenta, de que al preparar el desayuno, ha invertido una gran cantidad de tiempo y que usted vive repitiendo en la oficina "mi tiempo es oro", pero al preparar el desayuno no valoró lo suficiente su tiempo o sino hubiese desayunado en una cafetería. De igual forma, observa que el tráfico estuvo bastante congestionado, y que conducir produce una desutilidad que hasta el momento nunca había considerado.

Al pensar en su compañera de trabajo, se acuerda que su madre nunca trabajó y que por el contrario hoy es muy común encontrar que una gran parte de nuestros compañeros de trabajo son mujeres y, entonces exclama: ¡Definitivamente han cambiado los tiempos! ¿Cuál podría ser la explicación de que Linda trabajara?

Linda, que se da cuenta de cuánto tiempo ha perdido usted en esta reflexión, como buena economista, recuerda sus primeras clases en la universidad y repite en voz alta "La economía es la ciencia que estudia el comportamiento humano como una relación entre los fines y los medios escasos que tienen usos alternativos" [†]. Usted algo desconcertado, por la abrupta exclamación de Linda, le exige que ahonde en sus explicaciones. Ve, de reojo, como Linda realiza unos "cálculos" algo complicados a primera vista, y entonces ella responde que si usted le paga $50.000 la hora, con mucho gusto le explica el significado de aquella frase. Más enojado que al principio, a usted le parece increíble que Linda "cobre" por una simple explicación. Entonces, decide suspender la conversación, continuar con el trabajo y, tal vez en la tarde, pasar por la biblioteca de la Universidad y buscar el significado de la frase de Linda. Dos días después de haber realizado una búsqueda amplia por la biblioteca y recordando aquella frase "mi tiempo es oro" piensa en que de pronto habría sido mejor haberle pagado a Linda por la explicación.

La situación anterior, es más común de lo que pensamos. La mayoría de nuestras decisiones están determinadas por la escasez de los recursos y servicios. La cuenta de Linda no es otra cosa que la valoración de su tiempo, que no es infinito, al contrario, Linda desearía tener más tiempo para realizar otras actividades.

En estas páginas, usted encontrará una descripción de cómo opera el proceso de asignación cuando agentes como usted y Linda enfrentan restricciones y, cómo se verifica si estas restricciones se cumplen.

[†] Esta frase la usaba el profesor de Microeconomía al entrar a clase en homenaje a Lionel Robbins [Classic Monograph, An essay on the nature and significance of economics science, Macmillan & Co.,Ltd, London, 1932, pág. 15]

Objetivo principal del libro

El objetivo principal de este libro, consiste en mostrar, como la microeconomía puede ser aplicada a situaciones reales, que los consumidores enfrentan cotidianamente. Para esto, se provee al lector de las herramientas comúnmente usadas por los economistas para explicar el comportamiento individual de los agentes. Por tal razón, se desarrolla el esquema formal de la microeconomía, de tal forma, que podamos aproximarnos a la conducta de los agentes, siempre y cuando actúen bajo los supuestos enunciados. Ya que este libro pretende ofrecer una introducción elemental, pero a la vez global, de la teoría del consumidor, no se demostrará rigurosamente algunos de los teoremas usados, lo cual no implica que no se desarrolle cuidadosamente como éstos pueden emplearse en las estimaciones sobre los diferentes aspectos que conciernen a la elección del consumidor. Finalmente, estas notas pretenden distinguirse de los libros tradicionales de microeconomía en dos aspectos: Primero, mostrando aquellos elementos de la teoría del consumidor que se mencionan brevemente en los libros tradicionales de microeconomía y segundo mostrando a partir de la definición del modelo, como se estimarían sus implicaciones más importantes.

Los modelos económicos

Los modelos económicos son representaciones abstractas de la realidad para estudiar algún fenómeno económico y social. Ya que no se pueden construir versiones del mercado laboral, del mercado del ocio, etc., se acude a la representación abstracta del fenómeno en cuestión[†]. Esta representación no es otra cosa que un modelo matemático, en donde, las ecuaciones desarrolladas representan características del comportamiento de los agentes. Las ecuaciones del modelo, buscan aproximarse a las interrelaciones en la economía y, a partir, del planteamiento de las ecuaciones llegar a su estimación. A lo largo del libro, el lector deberá identificar los siguientes elementos:

1. Un conjunto de supuestos, denotados como $A = \{ A_1 , ..., A_n \}$ que tiene que ver con el comportamiento de una construcción teórica, y que en últimas, están relacionados con el mundo real. Los supuestos son proposiciones universales de la forma: *todo x tiene la propiedad r*. Ejemplos de tales proposiciones serán "*todos los consumidores maximizarán su utilidad" o "todos los consumidores son tomadores de precios"* o "*todas las preferencias son separables"*.

[†] Sin embargo, el avance en la economía experimental ha permitido recrear el funcionamiento de los mercados mostrando la distancia entre la presentación formal, su estimación, y el verdadero funcionamiento de éstos. Aunque muchos de los resultados provenientes de la economía experimental usen pocos individuos, generalmente son estudiantes de últimos semestres, no dejan de ser interesantes las conclusiones en torno a lo que todavía ignoramos de la conducta de los agentes en y como actúan en el "medioambiente" social. Será de gran utilidad, después de conocer la teoría formal, revisar el libro de Hey,J.D(1991). Experimentos en economía (1996), Fondo de cultura económica (México), para una introducción al tema Montenegro, A.(1995). Introducción a la economía experimental, Ediciones Uniandes/Ecoe, (Colombia), para una buena presentación Douglas, D y Ch, Holt.(1993), Experimental economics, Princeton University Press.

2. Ya que los supuestos de comportamiento deberán estar relacionados con el mundo real, un segundo elemento consistirá en el conjunto de condiciones, bajo las cuales, los supuestos son comprobados. Este conjunto de condiciones se denotara como C = $\{ C_1 , ..., C_n \}$. Las condiciones, deberán incluir la forma de identificar los efectos sobre las variables. Por ejemplo, suponga que deseamos comprobar que un aumento en los costos de viaje a una zona recreativa disminuye la demanda por viajes a dicha zona en el año 2000. Esto requiere primero, observar las condiciones por las cuales la gente viaja a dicha zona, en otras palabras, sus preferencias por una serie de actividades como navegar en vela, montar a caballo, acampar, etc., los costos de viajar en el año 2000 incluyendo la depreciación del automóvil y los costos de oportunidad del salario en el año 2000. Este conjunto de condiciones específicas, sitúa a Juan demandando viajes a dicha zona en 2000.

3. El último elemento consiste en los eventos E = $\{ E_1 , ..., E_n \}$ que son predecibles por la teoría. La teoría nos dice que el conjunto de supuestos A implica que si las condiciones C son válidas entonces el evento E podría ocurrir. Por ejemplo, si el comportamiento de Juan consiste en maximizar su función de utilidad sujeto a la restricción presupuestaria, lo cual se podría denotar como A, cuando las condiciones C se mantienen, entonces la disminución en la demanda de viajes a la zona en cuestión, el evento E, cuando aumentan los precios del viaje, será observado.

La estructura lógica se construye de tal forma, que el conjunto de supuestos A implica que si C es cierto, entonces E deberá ser cierto. Esto es, $(A \cdot C) \rightarrow E$, donde el símbolo $\rightarrow$ significa "implica". De esta forma, los supuestos A y las consecuencias C implican la observación de los eventos E.

Un elemento final, consiste en la forma funcional elegida. Suponga que usted desea considerar la ocurrencia de un evento, para describir este evento, definiremos una variable aleatoria Y. Asuma que la probabilidad del evento depende sobre un vector de variables independientes x* y un vector de parámetros desconocidos θ. El subíndice i denota el i- ésimo individuo. De esta forma, el modelo general se puede expresar como

$$p_i = p(Y_i) = G(x^*_i , \theta) ; i = 1,2,...., n.$$

Los Y_i son distribuidos independientemente. Suponga que Y_i es la probabilidad de que ocurra el evento, entonces x^*_i representará aquellas variables que permitan que Y ocurra. Dado que el "modelo" anterior es muy general, se deberá escoger alguna función $H(x^*_i , \theta)$, la cual se conoce que opera sobre un vector de parámetros θ , y de esta forma

$$p(Y_i) = F[H(x^*_i , \theta)]$$

Yo espero que usted pueda identificar, a lo largo del libro, la forma anterior. Y así, estas notas habrán contribuido a mejorar su percepción de la teoría del consumidor.

Contenido del Libro

El libro se ha estructurado en nueve capítulos, cubriendo desde los temas tradicionales que en cualquier libro de microeconomía se conoce como la teoría del consumidor, hasta aquellos temas que generalmente no se desarrollan como la función de producción de hogares, la diferenciación de productos y la utilidad aleatoria.

En los capítulos del 1 al 3 se desarrollará la teoría formal del consumidor. En el capítulo primero, se discute la forma de la restricción presupuestaría haciendo énfasis en las restricciones no lineales que un consumidor enfrentaría. En el segundo capítulo, se plantea formalmente la teoría del consumidor y se analizan aquellas formas funcionales que tienen una mayor tradición en economía.
En el capítulo tercero, se desarrolla el concepto de demanda del consumidor así como el concepto de excedente del consumidor y la recuperación de las preferencias del mismo a partir de su participación en el mercado.

En el capítulo cuarto, se discute la separabilidad de las preferencias y el presupuesto en dos etapas, es de especial atención en este capítulo la hipótesis de separabilidad y aditividad.

En el capítulo quinto, se presenta el modelo de función de producción de hogares. La importancia de este modelo radica en la formalización del uso del tiempo en diferentes actividades y, lo que esto significa en la elección de ocio y trabajo.

El capítulo sexto desarrolla la parte estadística que se usará en los capítulos posteriores. En este capítulo se presentan los modelos de Probabilidad Lineal, Logit, Logit Ordenado, Logit Condicionado, Probit, las diferentes modalidades del Tobit. También se incluye una sección sobre los contrastes de especificación y finalmente se incluye una sección sobre variables latentes.

El capítulo séptimo, presenta los modelos de utilidad discreta así como las diferentes versiones del mismo. En particular, la presentación de las funciones de densidad será de vital importancia para el desarrollo del capítulo noveno.

El capítulo octavo, presenta aplicaciones a la elección de ocio por parte del consumidor partiendo de las restricciones presupuestarias cuando los individuos deciden asignar su tiempo entre trabajo, ocio y otro tipo de actividades.

Finalmente, el capítulo noveno, desarrolla aplicaciones de la teoría del consumidor al medioambiente. El énfasis de estos modelos radica en el cálculo del excedente del consumidor cuando los bienes medioambientales entran en la función de utilidad.

1 Barreras a la elección

Las oportunidades de elegir un conjunto de bienes son directamente observables, por lo cual, las variaciones en las oportunidades influyen directamente sobre la elección de los individuos, existiendo una relación entre éstos y las variaciones en el comportamiento como efecto de la variación en el conjunto de oportunidades.

A menudo, cuando usted va a comprar algún bien, no sólo encuentra el bien que desea sino que además encuentra nuevos productos que le hacen reflexionar sobre los bienes que llevará. Esta situación, tan sólo muestra que la información que usted poseía sobre los bienes ha variado y, por lo tanto, que el conjunto de oportunidades ha cambiado.

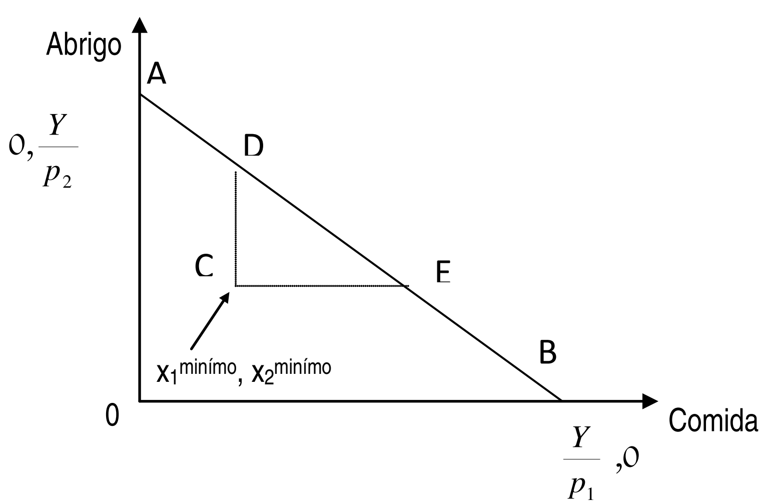
1.1 El conjunto de oportunidades

El conjunto de oportunidades más típico, se puede describir cuando los hogares tienen un ingreso exógeno Y, el cual se gasta durante un período determinado en todos los m bienes o en algunos. Dado que los bienes, o la cantidad de ellos, son positivos a precios positivos la restricción puede ser escribirse como

(1.1) $Y \geq \sum_{i=1}^{m} p_i x_i$ cuando m = 2 tendremos $Y \geq p_1x_1 + p_2x_2$

Donde Y es el ingreso, p los precios y x_i las cantidades. Supongamos, que existen dos bienes, Comida(x_1) y Abrigo(x_2) a unos precios p_1 y p_2, entonces, la gráfica que ilustra dicha restricción será

Gráfica 1.1 Restricción de supervivencia



1.1.1 Restricciones básicas

Suponga que existen cantidades mínimas de los dos bienes para sobrevivir $x_1^{\text{Mínimo}}$ y $x_2^{\text{Mínimo}}$. La elección, estará determinada por el triángulo CDE (gráfica1.1). De la anterior gráfica, un ingreso más bajo que $Y= p_1x_1^{\text{Mínimo}}+ p_2x_2^{\text{Mínimo}}$ no estaría en el conjunto de elección del individuo. Las restricciones pueden tomar diferentes formas: muy pocos abrigos y más alimentos pueden ser más necesarios que una gran cantidad de abrigos. Alternativamente x_1 y x_2 pueden ser tomados, como una composición de bienes en dos diferentes períodos.

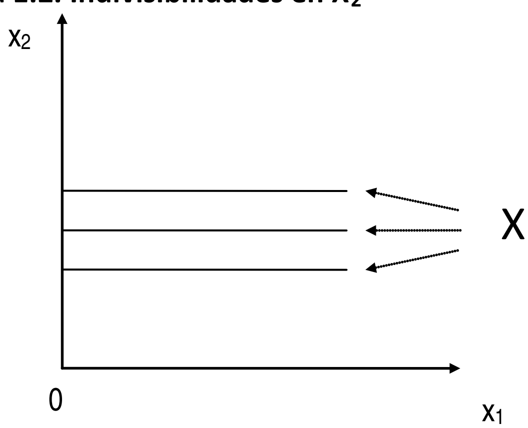
Si el consumidor comienza el período 1 sin dinero y si puede ahorrar o pedir prestado a una tasa de interés de cero, si el ingreso se distribuye en los períodos Y_1 y Y_2, y todo se gasta, entonces la restricción presupuestaria será

(1.2) $Y_1+Y_2 \geq p_1x_1+ p_2x_2$ con $Y_1+Y_2=Y$

En la anterior restricción, existe como supuesto implícito un mercado eficiente y cero costos de transacción. No siempre es posible derivar directamente *el conjunto de oportunidades*, supongamos los siguientes casos

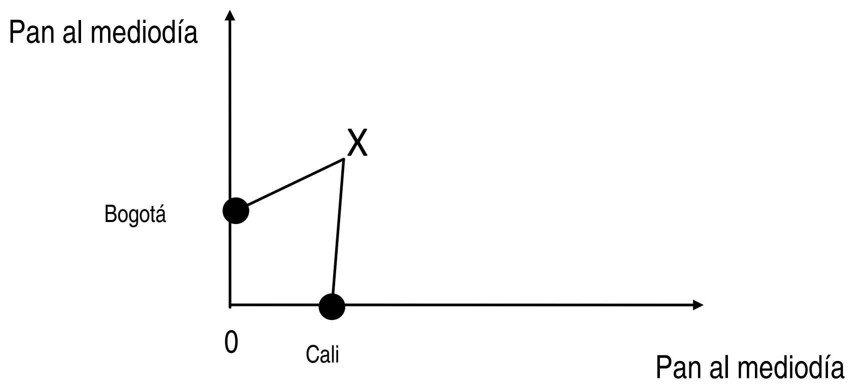
A- El primer bien es perfectamente divisible, pero el segundo es disponible en cantidades discretas

Gráfica 1.2. Indivisibilidades en X_2



B- El bien pan puede ser consumido al medio día por un individuo, ya sea en Santa fe de Bogotá o en Santiago de Cali, pero no al mismo tiempo en ambas ciudades:

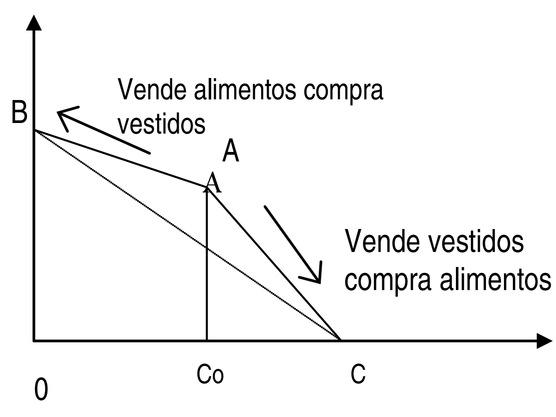
Gráfica 1.3. Imposibilidad geográfica de consumir un bien



1.2 Restricciones no lineales

Consideremos una economía de trueque y sea A la dotación inicial de alimentos y vestidos. Ahora, suponga la existencia de dos grupos: el grupo de los glotones y el grupo de los bien vestidos; el grupo de los glotones tiene comida y desea vestidos y el grupo de los bien vestidos tiene vestidos y desea comida. Los dos grupos viven en una isla y cada grupo vive aislado del otro, dado que no existe un medio único de intercambio, tampoco existirá una razón única de intercambio.

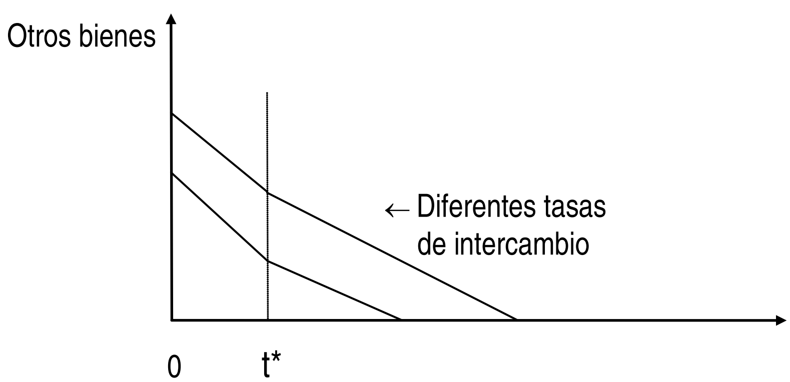
Gráfica 1.4. Diferentes tasas de intercambio



Como podrá observar, sin un medio general de intercambio, la información y los costos de transacción evitan al grupo que desea intercambiar ropa por alimentos a través de AC "iniciar un intercambio" con aquellos que desean cambiar alimentos por ropa a través de AB. De igual forma, sin un patrón general monetario, la tasa de intercambio difiere en las dos direcciones, por lo tanto, los grupos tienen diferentes tasas de intercambio. Una economía totalmente monetizada, donde se utilice como patrón de intercambio el dinero, eliminará las divergencias entre dichas tasas de intercambio, lo que se puede observar a través de la línea discontinua BC. Sin embargo, en una economía monetizada, con un mercado desarrollado, los problemas informacionales podrían causar divergencias entre los precios de compra y venta, generando no linealidades.

La no linealidad es más común de lo que se piensa; por ejemplo, la existencia de pagos diferenciales en las tarifas de agua: Al consumir x cantidades de m^3 de agua a un precio tendremos unos precios relativos entre el agua y los otros bienes, y al consumir más agua y pagar más por este consumo los precios relativos cambiarán. Supongamos que existe un punto óptimo de la tarifa t^* (gráfica 1.5): a la izquierda se paga una mayor tasa, lo que induce a consumir menos agua y más de los otros bienes, pero no tanto

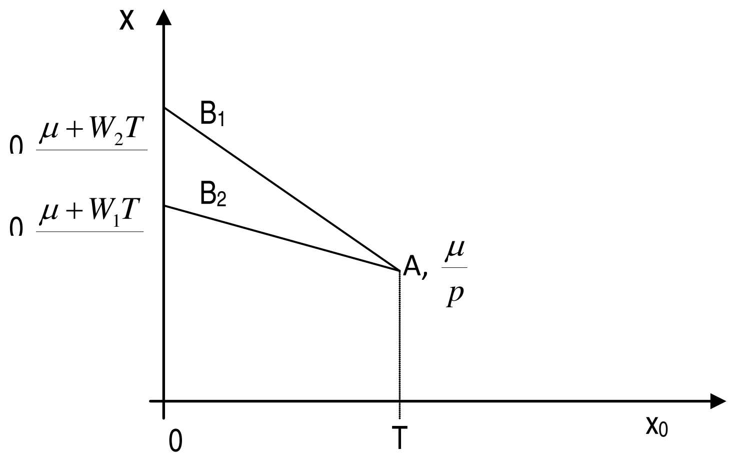
Gráfica 1.5. No linealidad en tarifas



En las elecciones de trabajo, es frecuente también, que existan no linealidades sobretodo en la elección de ocio o en el comportamiento intertemporal. Supongamos un individuo que elige el número de horas que puede trabajar. Y, cada hora se paga a una tasa fija de salario w. Adicionalmente, el individuo tiene algún tipo de transferencia en el ingreso μ como herencias, premios de loterías, etc. Si T es el número de horas disponibles y x_0 es el ocio, la restricción presupuestaria vendrá dada por:

(1.3) $$\mu + w(T - x_0) \geq \sum_{i=1}^{n} p_i x_i$$

Gráfica 1.6. No linealidad en la fuerza de trabajo



Observe, que si x es la cantidad de bienes que el individuo puede comprar trabajando un período de tiempo determinado, la restricción (1.3) se puede escribir como

(1.4) $\mu + wT - wx_0 = px$

Y, cuando x sea igual a 0, y μ también sea igual a 0[1] entonces $wT= wx_0$ ó $T = x_0$, por lo tanto todo el tiempo disponible se usa en ocio. Si no existe ocio entonces $\mu + wT = px$

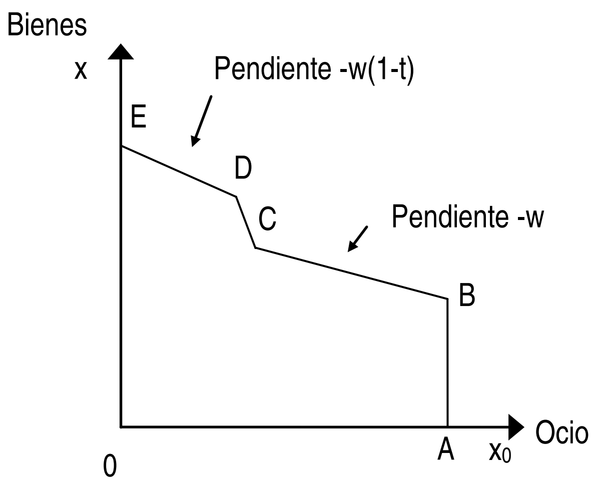
(1.5) $x = \frac{\mu + wT}{p}$

La elección del consumidor entre AB_1 ó AB_2, en la gráfica (1.6), es la elección de cuánto el individuo decide trabajar y, por lo tanto, el desplazamiento a través de AB_1 ó AB_2 depende de los gustos ya que AB_1 y AB_2 implican diferentes salarios percibidos, diferentes cantidades de los otros bienes y diferentes elecciones de ocio.

Suponga ahora, que existen impuestos al ingreso como retención en la fuente e incentivos por productividad. Después de un cierto número de horas de trabajo, éstas se pagarán a un mayor salario, como se puede ver en la línea E-D de la gráfica (1.7). Sin embargo, la existencia de impuestos hace que el incremento no sea igual al incremento en el salario sino menor, lo cual se traduce en una pendiente menor (– w) en la línea C-B como se puede observar

[1] Si el individuo no tiene transferencias deberá en algún momento decidir trabajar. De esta forma, a la izquierda de A existe una tasa de salario que lo incita a trabajar.

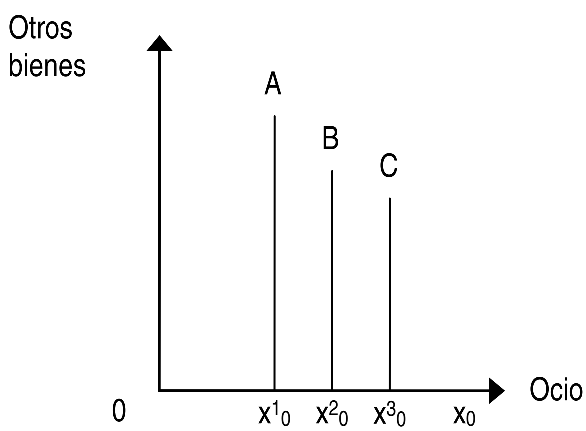
Gráfica 1.7. Efecto de un impuesto en la decisión de trabajar



Aunque B también es posible, a partir de C las horas se pagan a un mayor salario. Por otro lado, cuando se trabaja más allá de D, se deberá pagar impuestos.

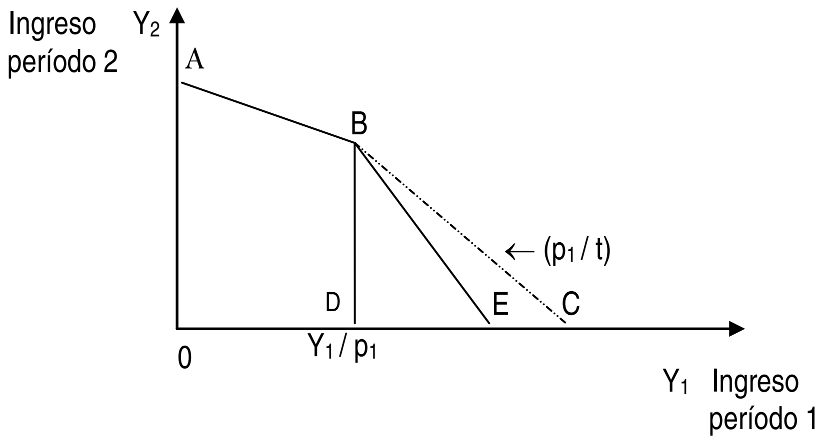
Cuando las elecciones realizadas incluyen tres tipos de trabajos, cada uno con diferente hora de trabajo y salario, podemos observar

Gráfica 1.8. Elecciones de trabajo y ocio a diferentes salarios



Otra forma de no linealidades, es introducida cuando en el conjunto de oportunidades, la relación ocio- ingreso, es diferente de acuerdo al período en el cual estos son consumidos, veamos:

Gráfica 1.9. Elección intertemporal con mercados de capital imperfectos



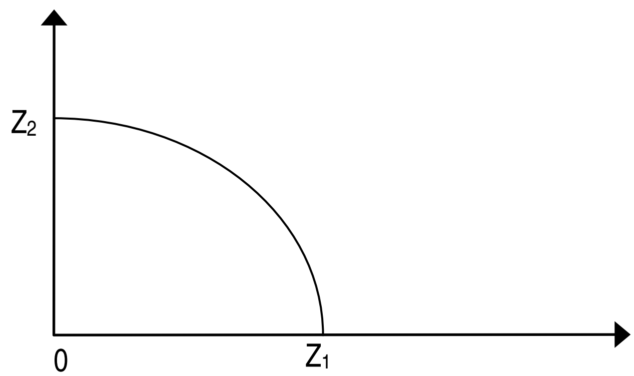
Si el consumidor, desea gastar más ingreso en el período 1, deberá prestar a una tasa de interés y pagar en el período 2. Cuando los consumidores no pueden conseguir dinero prestado, las elecciones se realizan en un mercado imperfecto de capitales, esto se puede observar si la restricción presupuestaria no es ABC sino ABD. Los consumidores en B gastan todo su ingreso y esta situación permanecerá sin cambiar, en tanto, tienen una propensión marginal unitaria. Si la situación no es tan extrema y asumimos la existencia de la tasa de interés, esto es, el consumidor paga una mayor tasa por pedir prestado, la restricción será menos severa y será descrita por la línea ABE. La pendiente está determinada por el hecho de que la tasa de interés de pedir prestado será diferente de la tasa de interés para prestar.

Uno de los más importantes desarrollos, desde la postguerra, en la teoría del consumidor, consiste en aquellos modelos donde el hogar se ve como una función de producción, esto es, cuando la combinación de los bienes con el ocio se hace a través de una función de producción de hogares. La función de producción de hogares, nos muestra la producción de un número limitado de bienes básicos considerados como el objeto real de la elección del consumidor.

Si existen dos bienes básicos Z_1 y Z_2, su producto es limitado por el tiempo disponible en el hogar, por la tasa de salario y, por los precios del bien en el mercado.

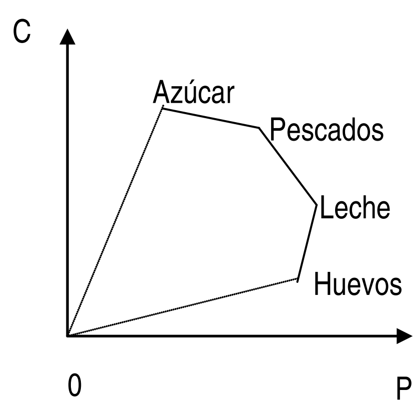
Si la función de producción asigna una igual sustitución entre tiempo y bienes de mercado como insumos, una restricción típica podrá ser

Gráfica 1.10. Restricción de presupuesto para una función de producción de hogares con tecnología de coeficientes fijos



Otra forma de no linealidad, se encuentra en el problema de la dieta, el cual muestra un hogar que requiere de dos productos proteínas (P) y calorías(C). Dado que los siguientes alimentos azúcar, pescados, leche y huevos tienen proteínas y calorías, la restricción podría venir especificada de la siguiente forma

Gráfica 1.11. El problema de la dieta



Si el gasto total se realiza en Azúcar, Pescado, Leche y Huevos los puntos muestran los límites a la elección. Los segmentos mostrarán canastas mixtas compradas, pero la elección no se realiza sobre los ejes {P, C} pues ellos mostrarán un 100% de proteína o un 100% de calorías.

Otro ejemplo de no linealidades, proviene de los modelos que asignan una cantidad determinada de tiempo sobre un sitio en modelos de demanda por recreación. Suponga que un consumidor elige x, los viajes a un lugar y Z una canasta Hicksiana. En cada viaje se consume t, siendo t el tiempo sobre el lugar.

(1.6) $Y = xc_x + xtc_t + c_Z Z$

Donde Y es el ingreso monetario, c_x es el costo por viaje, c_t es el costo en el lugar por unidad de tiempo (t en horas) y c_Z es el precio de una canasta Hicksiana . Asumiendo que exista la siguiente restricción de tiempo.

(1.7) $T^* = \gamma_x x + xt + \theta Z$

Donde T* es el tiempo disponible por consumir una serie de actividades, θ es el tiempo gastado en consumir Z y γ_x es el tiempo de viaje por cada paseo. Si T*, γ_x, t y θ son medidas en las mismas unidades(horas, días, años, etc.) y T es el tiempo disponible para trabajar o consumir

(1.8) $T^* = T - h$

Donde h es el tiempo gastado en trabajar. El agente elige la cantidad de tiempo para trabajar si h es endógeno. Cuando el agente elige una determinada cantidad de tiempo para trabajar los ingresos serán

(1.9) $wh = w(T - T^*) \Rightarrow \quad w(T - \gamma_x x - xt - \theta Z)$

(1.10) $y_0 + wh = xc_x + xtc_t + c_Z Z$

Siendo y_0 un ingreso exógeno, por ejemplo transferencias, herencias, loterías, etc. Reacomodando términos

(1.11) $y_0 + w(T - \gamma_x x - xt - \theta Z) = xc_x + xtc_t + c_Z Z$
(1.12) $y_0 + wT = xc_x + \gamma_x xw + xtc_t + wxt + c_Z Z + w\theta Z$
(1.13) $y_0 + wT = x(c_x + \gamma_x w) + xt(c_t + w) + Z(c_Z + w\theta)$

Si hacemos $p_x = c_x + \gamma_x w$, $p_t = c_t + w$ y $p_z = c_Z + w\theta$

(1.14) $y_0 + wT = x(p_x + p_t t) + p_z Z$

Los ingresos totales se pueden asumir como un dato y deduciendo a los ingresos totales, el consumo de la canasta hicksiana $y_0 + wT - p_z Z$, lo que queda será lo que se gaste en consumir x unidades de viaje. Haciendo $y_0 + wT - p_z Z = c_1$ tendremos

(1.15) $c_1 = xp_x + xp_t t$

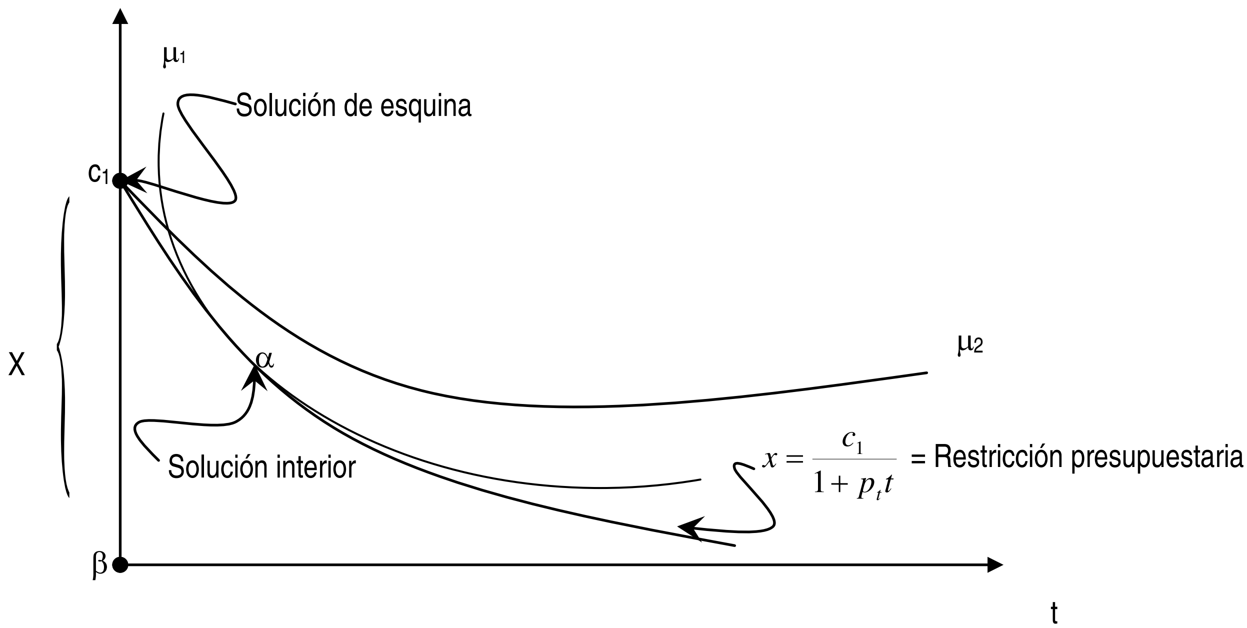
(1.16) $\frac{c_1}{p_x} = x + x\frac{p_t}{p_x}t$

(1.17) $\frac{c_1}{p_x} = x(1 + \frac{p_t}{p_x}t)$

Por lo tanto, la restricción del presupuesto no será lineal en el precio de x, p_x. Haciendo $p_x = 1$ se obtiene

(1.18) $x = \dfrac{c_1}{1 + p_t t}$

Gráfica 1.12. No linealidad en la demanda por recreación



En la gráfica (1.12) la solución interior, α, resulta de la intersección de la restricción presupuestaria con la curva de indiferencia μ_1. Si no existe ninguna restricción sobre la función de utilidad, la solución será c_1 con la curva de indiferencia μ_2, lo que se conoce como solución de esquina. Cuando existe débil complementariedad, $t = 0$ y $\partial u/\partial x = 0$, en el punto c_1, el individuo gastará xp_x, entonces podría moverse a β y ahorrar dinero sin haber reducido su utilidad, de esta forma, c_1 nunca sería elegido. Así, cuando existe débil complementariedad soluciones como α y β serán relevantes.

En resumen, la existencia de no linealidades en la restricción presupuestaria es muy común y estas no linealidades producirán variaciones de diferente proporción en el conjunto de oportunidades de elección de los agentes, afectando así, las elecciones realizadas por los agentes económicos.

1.4. Múltiples restricciones

En algunas situaciones, el consumidor no se enfrenta a una sola restricción, sino a varias restricciones, por lo cual podría estar racionado en un conjunto de bienes. Suponga, un individuo que deberá realizar una serie de elecciones entre una serie de bienes como deportes, ocio, educación, etc., y que denominaremos x_i. De igual forma, el consumir una unidad (i) requiere una cantidad de tiempo i=1,2,…….,n. Por lo cual, las restricciones para el consumidor vienen dadas como

$$\begin{aligned} &p_{deportes} deportes + p_{ocio} ocio + p_{educacion} educacion + ... + p_i X_i \leq Y \\ &t_{deportes} deportes + t_{ocio} ocio + t_{educacion} educacion + ... + t_i X_i \leq T \end{aligned} \tag{1.19}$$

Siendo Y el ingreso del individuo, t_i el tiempo dedicado a la actividad (i) y T el tiempo total disponible.

2 Preferencias individuales

Un elemento fundamental en la teoría microeconómica, consiste en cómo los individuos realizan sus decisiones y, cómo seleccionan alternativas de un conjunto disponibles de las mismas. La teoría postula, que cada individuo ordena las alternativas de acuerdo a su preferencia relativa. De esta forma, cuando el individuo realiza una elección, éste selecciona la alternativa con aquello que más tiene de todo lo posible. En este capítulo, se desarrollará el marco teórico asociado con el concepto de preferencia.

2.1 Preferencias individuales

Asuma la existencia de n alternativas, éstas pueden contener n diferentes objetos que usted puede poseer, n posibles candidatos por los cuales votar, n diferentes empleos a optar, etc. En general, cuando hay n alternativas en algún orden que desea, usted podrá expresar un orden de preferencias por las mismas. Cuando algunas alternativas tienen el mismo nivel en su lista usted tendrá indiferencia entre las mismas. Existen dos propiedades importantes en su lista: *Primero*, es posible comparar dos alternativas diciendo cual de las dos es mayor, de esta forma, una es más preferida que la otra, o cuando ella tiene el mismo nivel. *Segundo*, dada la naturaleza de las preferencias esta no es cíclica[2], esto es, si la primera alternativa es mayor que la segunda, y también mayor que la tercera, entonces la primera alternativa es mayor que la tercera. Ahora, usted puede establecer un orden, y si solamente algunas de las alternativas son posibles, entonces podrá seleccionar aquella alternativa que más prefiera. Un orden más general, se establece para un infinito número de elementos y aún cuando la lista con dicho orden sea complicada, se puede decir que el ordenamiento se mantiene.

2.1.1 Definición Formal

Sea X el conjunto de posibles alternativas consideradas por un individuo, el conjunto X simplemente puede ser un conjunto finito de alternativas o en el caso de un consumidor éste puede representar el conjunto de canastas de consumo disponibles. Una relación binaria sobre X, es un subconjunto R de $X \times X$ con el conjunto de pares ordenados (x , q) donde $x \in X$ y $q \in X$. Los pares en el subconjunto de R se dicen que satisfacen esta relación. Una relación de preferencia es un caso especial y se escribe $x \succsim q$ si $(x, q) \in X \times X$ satisface esta relación. Si $x \succsim q$ entonces se dice que x es preferido a q. Esta relación puede entenderse en el sentido débil como "al menos es tan bueno como" más que en el sentido es "mejor que". De igual forma, una relación estricta de preferencias $\succ$ se define como $x \succ q \Leftrightarrow x \succsim q$ pero no $q \succsim x$, y se lee x es preferido a q. La relación $\sim$ se conoce

---

[2] Esto significa que deberá cumplir el teorema de aciclicidad: Si para un entero finito n, $x_1>x_2$, $x_2>x_3$, $x_3>x_4$,….., $x_{n-1}>x_n$ entonces $x_n \neq x_1$ (Kreps 1995).

como indiferencia y se define por x $\sim$ q $\Leftrightarrow$ x $\succsim$ q y q $\succsim$ x y se lee x es indiferente a q. En orden a cualificar la relación de preferencias, la relación $\succsim$ deberá satisfacer las siguientes propiedades fundamentales

2.1.1.1 Reflexibidad

Para todo x $\in$ X, x $\succsim$ x. Este supuesto nos dice que la canasta x, en el sentido débil, es preferida a sí misma, es decir que al menos es tan buena como ella misma.

2.1.1.2 Completitud

Para todos los elementos x , q en X se cumple que x $\succsim$ q o q $\succsim$ x o ambos. Este supuesto simplemente nos dice que dos canastas pueden ser comparadas.

2.1.1.3 Transitividad

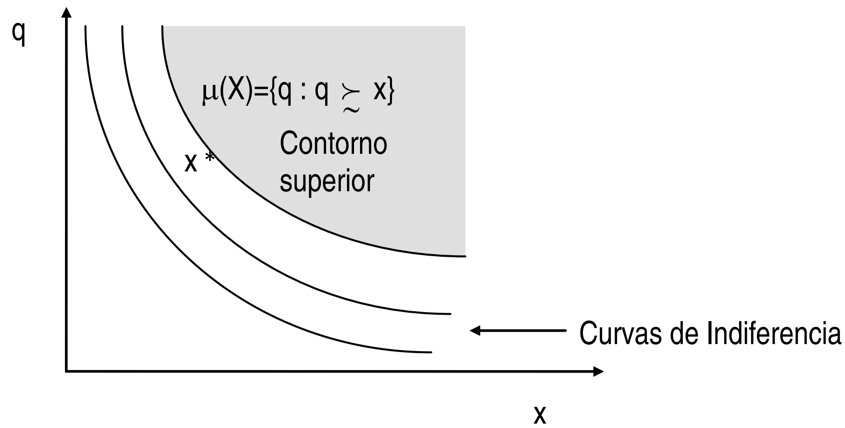
Para todo x , q y z en X, si x $\succsim$ q y q $\succsim$ z entonces x $\succsim$ z. La propiedad de transitividad previene la circularidad en las relaciones de preferencias.

Los axiomas 2.1.1.1 al 2.1.1.3 definen un conjunto de elección preordenado.

2.2 Preferencias sobre $\Re_+^m$

Las relaciones de preferencias se usan para caracterizar los deseos de los consumidores, por varias combinaciones de bienes. Los bienes son indexados de 1 hasta m. Una canasta de bienes, es una colección de varias cantidades de esos m bienes y, la cantidad de cada bien en una canasta es un número real positivo. También podemos ver una canasta, como la representación de un vector m-dimensional de números no negativos, comúnmente se asume que los bienes son divisibles. Tomemos X = $\Re_+^m$ como el ortante no negativo de $\Re^m$, en este conjunto una relación de preferencia en el caso de dos dimensiones puede verse de la siguiente forma:

Gráfica 2.1. Preferencias en dos dimensiones



2.2.1.1 Continuidad

Una relación de preferencias $\succsim$ sobre X= $\Re^m_+$ se dice que es continua si para cada x $\in$ X los conjuntos { y $\in$ X | y $\succsim$ x } y { y $\in$ X | x $\succsim$ y } son cerrados. Esto es, para alguna canasta x defina A(x) como el conjunto donde x es “al menos tan bueno como y ” y B(x) “no existe un mejor conjunto que x ” de tal forma que

(2.1) A(x)= { y $\in$ X | y $\succsim$ x } y B(x) = { y $\in$ X | x $\succsim$ y }

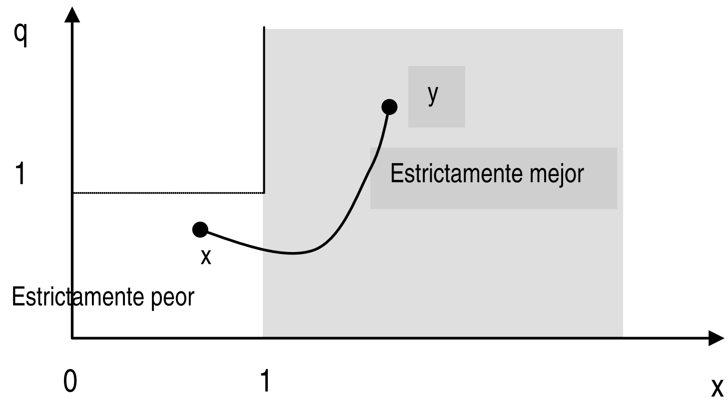
Observe que A(x) y B(x) serán cerrados, pues contienen sus propios límites para x en el conjunto de elección. A A(x) se le denomina el conjunto superior y a B(x) el conjunto inferior. De lo anterior, se deduce

(2.2) A(x) ={ y $\in$ X | y $\succ$ x } y B(x) = { y $\in$ X | x $\succ$ y }

Serán abiertos. Se podrá observar entonces que (2.1) y (2.2) se utilizan para evitar conductas discontinuas.

Observe que para un orden de preferencias $\succsim$ la intersección entre los conjuntos superior e inferior, para algún x, definirá el conjunto de indiferencia I(x) = { y $\in$ X | y $\sim$ x } el cual es cerrado. Con referencia a la gráfica 2.1 la continuidad asegura que los puntos sobre la frontera de μ(X) = { y $\in$ X | y $\succsim$ x } sean equivalentes al elemento x. Observe sin embargo, que la continuidad no asegura la posibilidad de que la superficie de la curva de indiferencia pueda contener conjuntos cerrados. Por ejemplo, en el conjunto de preferencias lexicográficas

Gráfica 2.2. Preferencias Lexicográficas: Se salta de estrictamente peor en x a estrictamente mejor en y



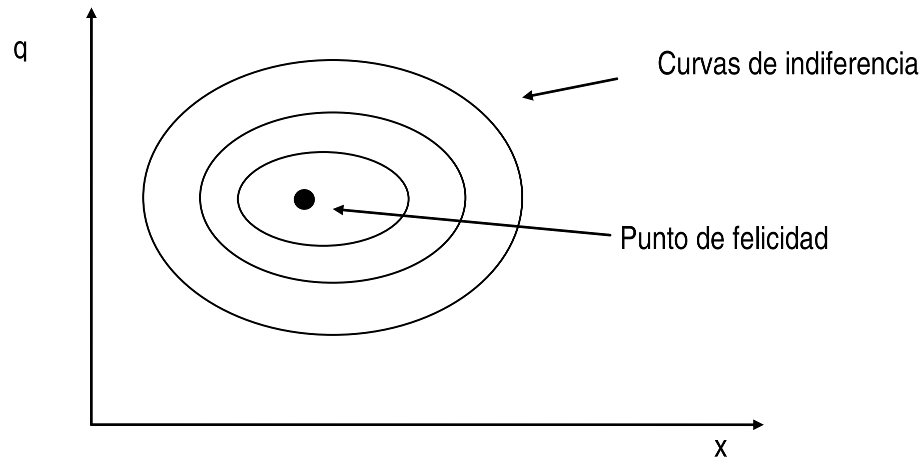
El orden de preferencias lexicográficas no es continuo. Como se puede observar en el caso de dos dimensiones, considérese un conjunto cuyo contorno superior corresponda al elemento x = (1,1), esto es, el conjunto de elementos q $\succsim$ x cuya gráfica es (2.2), claramente no es cerrado debido a que la frontera del conjunto por abajo (1,1) no esta contenida en el conjunto. Por ejemplo (1,½) < (1,1) y de igual forma (1,1½) $\geq$ (1,1).

2.2.1.2 Insaciabilidad

Una relación de preferencia $\succsim$ sobre X se dice que no es saciada si para todo x $\in$ X existe un q $\in$ X siendo q $\succ$ x.

Lo contrario a dicha afirmación es la existencia de un elemento x_0 en X que sea preferido a otro elemento. Tal elemento se denominará "punto de felicidad" y se puede observar en la siguiente gráfica

Gráfica 2.3. Puntos de felicidad



Existen otras propiedades relacionadas con la insaciabilidad, a saber

2.2.1.2.1 Insaciabilidad local

Una relación de preferencia $\succsim$ sobre X = $\Re^m_+$ satisface la insaciabilidad local si para algún x en X y algún ε > 0 existe algún q en X con $\|x-q\| < \varepsilon$ tal que q $\succ$ x [3]. Esto significa, que para alguna canasta x, existen canastas cercanas que son estrictamente preferidas a x, lo cual podría de igual forma entenderse como que es posible siempre mejorar aunque sólo se introduzcan pequeñas variaciones en la cesta de bienes.

2.2.1.2.2 Fuerte monotonicidad

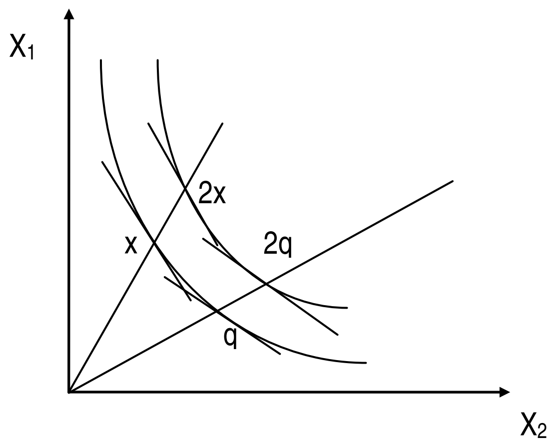
Una relación de preferencia $\succsim$ sobre X=$\Re^m_+$ es fuertemente monótona si x $\in$ X, q $\in$ X y q $\succsim$ x, q $\neq$ x implica que q $\succ$ x. La fuerte monotonicidad implica insaciabilidad local. Si una cesta de bienes contiene como mínimo la misma cantidad de bienes que otra y más de alguno de ellos, esta cesta es estrictamente mejor que la otra, lo cual significa suponer que los bienes son buenos, pues si uno de ellos fuese un mal, como la contaminación o la basura no se cumpliría este supuesto. En términos generales esta propiedad se conoce también como que "los bienes son buenos".

[3] Donde $\|x-q\| < \varepsilon$ es la distancia euclideana entre los puntos x y q: $\|x-q\| = (\sum_{i=1}^{n}(x_i - q_i)^2)^{\frac{1}{2}}$

2.2.1.2.3 Homotecia

Una relación de preferencias monótonas $\succsim$ sobre X= $\Re_{+}^{m}$ es homotética si todos los conjuntos de indiferencia están relacionados por un rayo de expansión proporcional, esto es, sí x $\sim$ q entonces αx $\sim$ αq para algún $\alpha \geq 0$, veamos

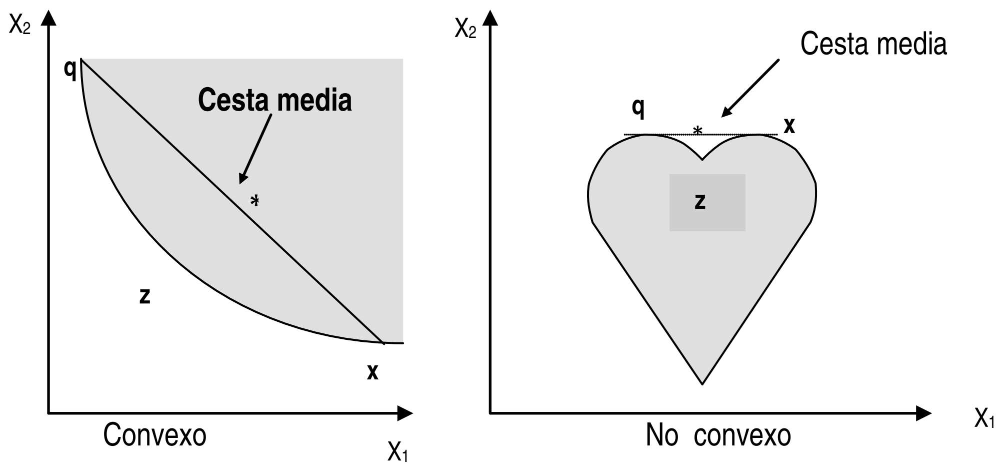
Gráfica 2.4. Preferencias homotéticas



2.2.1.3 Convexidad

Una relación de preferencia $\succsim$ sobre X= $\Re_{+}^{m}$ es convexa si dado algún x, q y z en X tal que x $\neq$ q, x $\succsim$ z , q $\succsim$ z entonces para todo α, 0< α<1, se cumple que αx + (1- α)q $\succsim$ z y se dice que es estrictamente convexa si x $\succ$z , q $\succ$ z y para todo α, 0< α<1, se cumple que αx + (1- α)q $\succ$ z. De igual forma, esta propiedad podría leerse como que la canasta media, es preferida a los extremos, veamos

Gráfica 2.5. Conjuntos convexos y no convexos



2.3 La función de utilidad

Si la ordenación de preferencias es completa, transitiva, reflexiva y continua, entonces las preferencias se pueden representar a través de una función de utilidad continua. La función de utilidad, es una función con valores reales, definida sobre el conjunto X, de tal forma, que el orden de las preferencias sobre X se preserva por la magnitud de u. De esta forma, una función de utilidad tiene la propiedad de que dados dos elementos x y q en X se cumple que $u(q) \geq u(x)$ sí y solo sí $q \succsim x$.

No todas las relaciones de preferencias pueden ser representadas por funciones de utilidad, pero si la relación de preferencias es continua sobre $\Re_+^m$ entonces ésta puede ser representada por una función de utilidad. Por lo tanto, sí las preferencias son reflexivas, transitivas, completas y continuas, u(x) representa dichas preferencias y deberá cumplirse: Primero, u(x) es estrictamente creciente sí y solo sí las preferencias son monótonas. Segundo, u(x) es cuasiconcava sí y solo sí las preferencias son convexas y tercero u(x) es estrictamente cuasiconcava sí y solo sí las preferencias son estrictamente convexas [4].

2.3.1 Invarianza de la función de utilidad

Sea $\succsim$ un orden de preferencias continuo tal que u(x) es una función de utilidad que represente estas. Si $f(\cdot)$ es una función estrictamente creciente de una variable

[4] Una función $f: D \to R$ es cuasiconcava si $\forall$ x_1 y x_2 en D : $f(x^t) \geq \min [f(x_1), f(x_2)]$ $\forall$ $t \in [0,1]$

singular y, f(u(x)) es la función compuesta y está es una transformación monótona positiva de u(x), entonces está también representa una función de utilidad. De lo anterior se deduce

1) Que $u(x_1,x_2)$ represente $\succsim$ significa que $u(x_1,x_2) > u(q_1,q_2) \Leftrightarrow (x_1,x_2) \succ (q_1,q_2)$

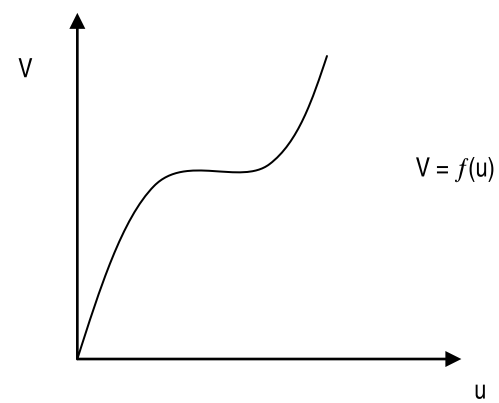
2) $f(\cdot)$ es una transformación monótona de $u(x_1,x_2) > u(q_1,q_2) \Leftrightarrow f(u(x_1,x_2)) > f(u(q_1,q_2))$

3) De (2) se observa que $f(u(x_1,x_2)) > f(u(q_1,q_2)) \Leftrightarrow (x_1,x_2) \succ (q_1,q_2)$

Proposición 2.1

Si una relación de preferencia es representada por una función de utilidad sobre $\Re_+^m$, entonces una función de la forma v(x) = f(u(x)), donde f es estrictamente creciente en el rango de v sobre u, será también una función que represente la misma relación de preferencia. Si f y u son continuas entonces v será también continua. Esto se deduce de (1),(2) y (3).

Gráfica 2.6. Transformación monótona de u



La invarianza en la función de utilidad deberá incluir como requisitos adicionales: mantener la invarianza en la descripción, la invarianza en el procedimiento y la invarianza en el contexto[5].

[5] Ver también la descripción de Kreps(1995) sobre el encuadramiento(framing).

2.3.1.1 Invarianza en la descripción

Este requisito requiere que las preferencias entre las opciones no dependan de la forma en la cual ellas son presentadas. De esta forma, dos descripciones del mismo problema deberán llevar a la misma elección [Arrow (1982), Tversky y Kahneman(1986)]. Tversky y Kahneman (1986) proveen el siguiente ejemplo, que viola esta propiedad.

Problema 3 (126 individuos participaron en el experimento): Asuma que usted se enriquece en $300 más que hoy. Usted tendrá una elección entre

A) Una ganancia segura de $100 [72% de los individuos eligieron esta opción].

B) 50% de oportunidad de ganar $200 y 50% de oportunidad de no ganar nada [28% de los Individuos eligieron esta opción].

Problema 4 (128 individuos participaron en el experimento): Asuma que usted mismo se enriquece en $500 más que hoy. Usted tendrá una elección entre

A) Una pérdida segura de $100 [36% de los individuos eligieron esta opción].

B) 50% de oportunidad de no perder nada y 50% de oportunidad de perder $200 [64% de los Individuos eligieron esta opción].

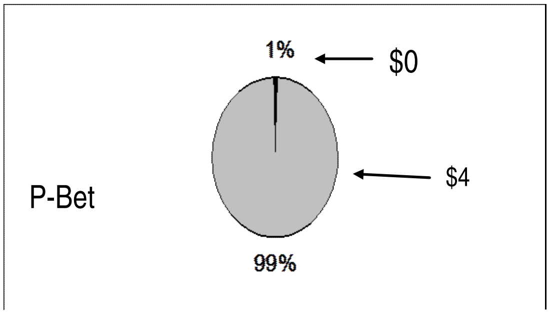
Dado que los dos problemas son idénticos, la variación en la descripción tiene un gran efecto en las preferencias.

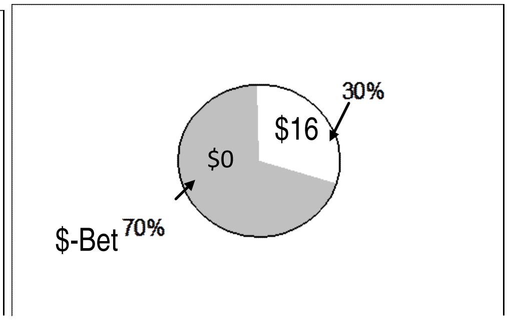
Mora (2008) realizo el anterior experimento con estudiantes de la Universidad Icesi controlando por el tipo de carrera y semestre y no encontraron que los efectos como el sexo o el estrato social fuese determinante a la hora de explicar la no consistencia en la preferencias.

2.3.1.2 Invarianza en el procedimiento

Esta propiedad requiere que los métodos de “extraer” las preferencias mantengan el mismo orden en ellas. Por lo que, dos procedimientos deberán mantener el mismo orden en las preferencias. Este fenómeno, está asociado directamente a la existencia de inversión en las preferencias descrito inicialmente por Sarah Lichtenstein y Paul Slovicv y ampliado por Tversky y Kahneman. Considere un individuo que tiene la oportunidad de jugar dos loterías representadas en la gráfica siguiente [6]

[6] Esta versión de la inversión en las preferencias es tomada de Plott,Ch.R(1996)





Lotería A Lotería B

La lotería A da un pago de $4 con una gran certeza y un pago de $0 con una pequeña probabilidad. La lotería B da un pago de $16 con una probabilidad de un 30 % y un pago de $0 con una probabilidad de 70%. La lotería A es llamada P-Bet debido a que la probabilidad de ganar es muy grande y la lotería B es llamada $-Bet debido a que la cantidad a ganar es muy grande. Cuando los individuos son preguntados por su elección la mayoría elige A. Sin embargo, cuando se les pregunta cuánto pagarían por el derecho a jugar las loterías, el mismo individuo desearía pagar más por el derecho a jugar la lotería B. De esta forma, la inconsistencia en el comportamiento es evidente, mostrando asimetría en el cambio. [Ver también Tversky (1996) y McFadden (1999)].

2.3.1.3 Invarianza en el contexto

El último requisito consiste en la invarianza en el contexto, definido por el conjunto de opciones bajo consideración. De acuerdo con Tversky(1996) uno de los supuestos básicos en una elección racional consiste en que cada alternativa tiene una utilidad que depende solamente sobre esa alternativa. Esto significa que una opción no preferida, no puede preferirse si se adicionan nuevas alternativas al conjunto de elección. Lo contrario mostraría que no existe invarianza en el contexto. Esta hipótesis implica que si no existe invarianza, la "´parte del mercado´ de x puede ser incrementada por adicionar a {x,y} una tercera alternativa z que es claramente inferior a x pero no a y" (Tversky:194). Un ejemplo sobre la violación de este supuesto, es provisto por los autores anteriores: A un grupo de 106 encuestados se les ofreció elegir entre $6 y un bolígrafo Cross, el porcentaje que seleccionó el bolígrafo fue del 36% y el resto prefirió el dinero. A un segundo grupo de 115 encuestados se les ofreció elegir tres opciones: $6, el bolígrafo Cross y un bolígrafo menos atractivo, el 2% eligió el bolígrafo menos atractivo, mientras el porcentaje que eligió el Cross aumento del 36% al 46%.

2.4 El problema básico del consumidor

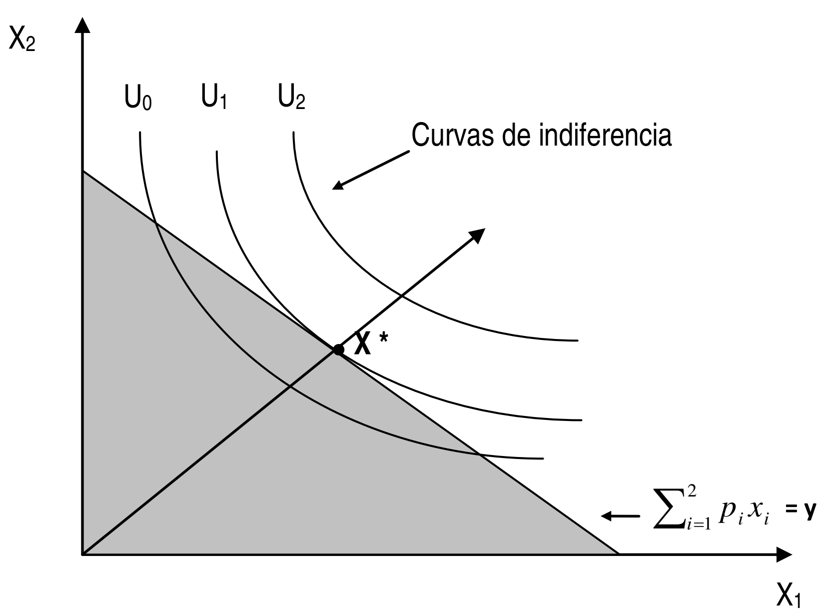
Deberemos ahora introducir los precios en nuestro modelo básico. Cualquier consumidor ha experimentado que sus deseos de elegir m bienes se ven frustrados cuando decide ir al mercado, a un centro comercial, etc. Dicha frustración, no es más que la confirmación de que aun cuando se tienen preferencias por los bienes, éstas, por sí solas no bastan, esto es, existen restricciones como la cantidad de dinero que poseemos en nuestros bolsillos para comprar dichos bienes. De una manera más formal, asumamos que existen m bienes, las cuales son infinitamente divisibles.

Un consumidor selecciona una canasta que contiene dichos bienes descritas por el m vector en $X = (x_1,x_2,\ldots,x_m)$ donde x_i , i=1,..,m, representa la cantidad de la bien i. Las preferencias del consumidor, para varias posibles canastas, se representa por la relación de preferencias $\succsim$ sobre $\Re^m_+$.

Asociado a cada bien existe un precio, medido en alguna unidad monetaria $p_i \geq 0$, de tal forma, que el costo de elegir x_i será $p_i\ x_i$. El costo total de elegir la canasta X = $(x_1,x_2,\ldots,x_m)$ será $\sum_{i=1}^{m} p_i x_i$ = p X. Asumiremos por simplicidad, que el consumidor tiene un presupuesto de y (unidades monetarias).

Así, la elección del consumidor es restringida a la restricción de presupuesto

Gráfica 2.7. Elección del consumidor



Por lo tanto, el consumidor elegirá una canasta determinada de acuerdo a la restricción

$$(2.3)\quad \sum_{i=1}^{m} p_i x_i \leq y$$

Esta situación, se representa en la gráfica (2.6) donde la restricción de presupuesto se define como el área sombreada debajo de la recta $p \cdot x = y$. Como se podrá observar, la diagonal de esta región será perpendicular al vector de precios p. Si el presupuesto cambia, digamos aumenta, dicha región también aumentará desplazándose a la derecha y, de lo contrario a la izquierda. Un consumidor que actúe racionalmente, de acuerdo a la información que posee, preferirá una serie de canastas que estén en el interior del área sombreada, pues éstas son asequibles a su ingreso, de ahí que la restricción (2.3) nos indique que tanto U_0 como U_1 son asequibles al consumidor, mientras U_2 de acuerdo a dicha restricción no lo es.
¿Cómo se deberá elegir entre U_0 y U_1? Para resolver dicho interrogante se usará el supuesto 2.1.1.5.1 por lo cual la desigualdad se convertirá en igualdad y, entonces la canasta óptima será X^* con la curva de indiferencia U_1. Mas-Collel, Whinston y Green (1995) definen esta solución de la siguiente forma: la demanda Walrasiana x(p,y) satisface la ley de Walras, si para cada p >> 0 e y > 0 se cumple que $p \cdot x = y \ \forall \ x \in X(p,y)$, esto es, el consumidor gasta totalmente su riqueza (1995:23). De manera formal, diremos que el problema básico del consumidor dados unos precios p y un presupuesto y, será solucionar

$$(2.4)\quad \begin{array}{l} \max\limits_{x} u(x) \\ \text{Sujeto a } p.x \leq y \ ; \ x \in X \end{array}$$

Dado que el problema está bien definido, entonces una solución única existirá. Si p es estrictamente positivo (>>) y $u(\cdot)$ es continua el problema de maximización de la utilidad tiene una solución.

2.4.1 Restricciones múltiples

Suponga una función de utilidad continua y cuasi-cóncava. Cada $g_i(x)$ es convexa y X es un conjunto convexo. Suponga también que existe algún $x \in X$ con $g_i(x) < 0 \ \forall \ i=1,2,..,s$. Si x^* es una solución existe una serie de $\lambda_1, \lambda_2, \ldots, \lambda_s$ positivos, pero no todos iguales a cero, tal que x^* es una solución al problema

$$(2.4.1)\quad \begin{array}{l} Max\ \mu(x) \\ Sujeto\ a \sum\limits_{i=1}^{s} \lambda_i g_i(x) \leq 0 \quad ; \quad x \in X \end{array}$$

Así, el problema (1.12) del capítulo 1 podría escribirse como

$$\begin{aligned} &Max\ \mu(x) \\ (2.4.2)\quad &Sujeto\ a: Deportes(\lambda_{ingreso} p_{deportes} + \lambda_{tiempo} t_{deportes}) + Ocio(\lambda_{ingreso} p_{ocio} + \lambda_{tiempo} t_{ocio}) + .. \\ &..+x_m(\lambda_{ingreso} p_m + \lambda_{tiempo} t_m) \leq \lambda_{ingreso} Y + \lambda_{tiempo} T \end{aligned}$$

Como podrá observarse λ_{tiempo} es una clara interpretación del valor del tiempo, este será el valor por medio del cual, una unidad de tiempo, por ejemplo una hora, puede ser convertida en dinero.

2.5 Dualidad

Uno de los aspectos importantes en la teoría del consumidor, consiste en la dualidad. La dualidad es una de las "herramientas" más usadas en la estimación de modelos. Básicamente la dualidad expresa la relación entre los bienes por un lado y los precios por el otro. De esta forma, el consumidor podrá elegir entre maximizar la función de utilidad sujeto a la restricción de presupuesto o, minimizar su gasto en una serie de bienes siempre y cuando, la función de utilidad permanezca constante. El problema se plantea como

$$(2.5)\quad \begin{aligned} &\max_{x}\ \mu(x) \\ &\text{Sujeto a } p.x \leq y\ ;\ x \in X \end{aligned}$$

ó

$$(2.6)\quad \begin{aligned} &\min_{x}\ p.x \\ &\text{Sujeto a } \mu(x) = \mu^{o};x \in X \end{aligned}$$

De la solución al problema (2.5) resultarán las demandas marshallianas, mientras de la solución a (2.6) resultarán las demandas Hicksianas o demandas compensadas. Las funciones Hicksianas satisfacen la cantidad de bienes x a los precios p cuando la utilidad permanece constante, de ahí su nombre

(2.6.1) $X_i = g_i\ (y,p) = h_i\ (u,p)$

(2.6.2) $\mu = v(x_1,\ldots,x_m) = v[g_1(y,p),g_2(y,p),..,g_m(y,p)] = \varphi\ (y,p)$

(2.6.3) $x = \sum_{i=1}^{m} p_i h_i(u,p) = C(u,p)$

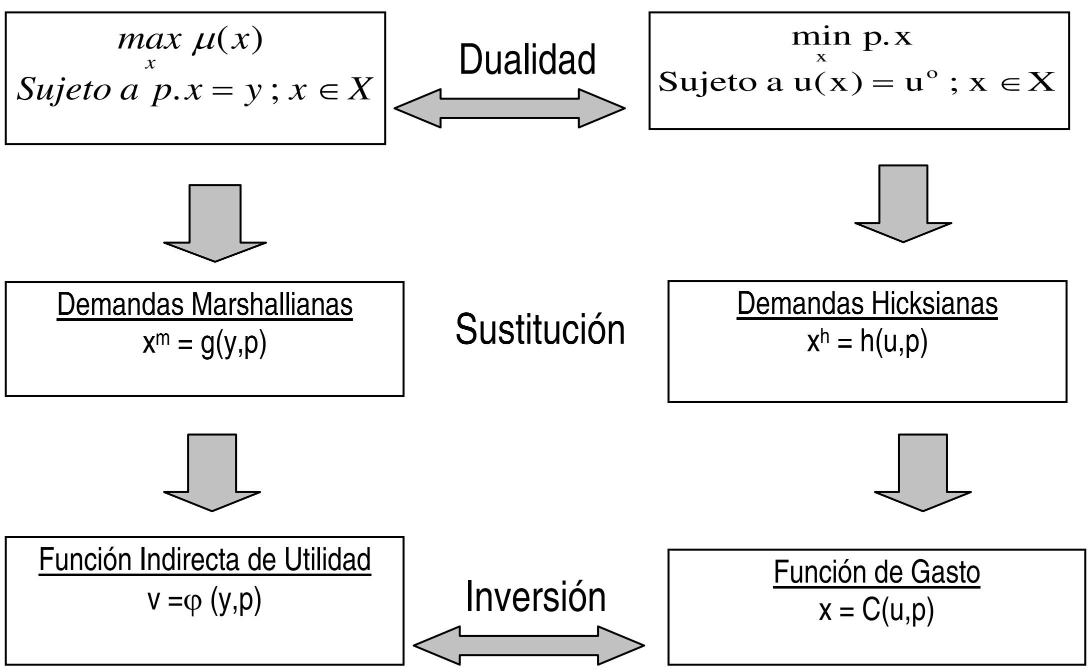
Donde φ (y,p) es la función indirecta de utilidad o el máximo sostenible de utilidad dados los precios y el ingreso

(2.6.4) $\varphi(y,p) = \max_{x} [v(x);\ p.x = y]$

Y, la función C(u,p) es el mínimo gasto que mantiene la utilidad constante, dado los precios p, y claramente es una solución al problema dual

(2.6.5) $C(u,p) = \min_{x} [\ p.x = ;\ u(x) = v]$

La función de gasto y la función indirecta de utilidad están íntimamente relacionadas, pues a partir de invertir C(u,p) = x se encuentra u en función de x y p, por lo cual obtenemos v=φ (y,p). Similarmente la inversión de u=φ (y,p) nos lleva directamente a x = C(u,p). Esto se puede observar mejor en el siguiente esquema:



2.5.1 Propiedades de la función indirecta de utilidad

Entre las propiedades usuales de la función indirecta de utilidad, tenemos

1) Es homogénea de grado cero en (p,y), esto es, v(tp, ty)= v(p,y) ∀ t >0.
2) No es creciente en p y es estrictamente creciente en y.

3) Es cuasi convexa con respecto a p, esto es el conjunto {p: $v(p,y) \leq c$} es convexo para cada $y > 0$ y algún c.
4) La derivada de la función indirecta de utilidad con respecto a los precios e ingreso se conoce también como la **Identidad de Roy** y es una forma conveniente de recuperar la demanda Marshalliana

(2.7) $x_i^m = g_i\ (y,p) = v = -\dfrac{\partial\varphi / \partial p_i}{\partial\varphi / \partial y}$

5) Es continúa en p e y.

2.5.2 Propiedades de la función de gasto

Entre las propiedades usuales de la función de gasto, están

1) La función de gasto en homogénea de grado uno en precios, formalmente para algún escalar
$\theta > 0 : C(u, \theta p) = \theta\, C(u,p)$. Esto es, si los precios se doblan se deberá desembolsar dos veces más cantidad de dinero para estar en la misma curva de indiferencia.
2) La función de gasto es creciente en μ no decreciente en p y creciente en al menos un precio. Esto se deriva del axioma de insaciabilidad ya que dados unos precios, el consumidor tiene que gastar más para estar mejor, ya que un incremento en precios requiere más cantidad de dinero para permanecer mejor.
3) La función de gasto es cóncava en precios. Cuando el precio de un bien cambia, mientras los otros precios y la utilidad permanecen constantes, la concavidad implica que el costo aumenta no más que linealmente, esto es esencial, porque el consumidor minimiza sus gastos reacomodando sus compras en orden a tomar las ventajas de la estructura de precios. Formalmente una función escalar f(z) definida sobre un vector z, se dice que es cóncava si para $0 \leq \theta \leq 1$ se mantiene

(2.8) $f[\theta z^1+(1-\theta)z^2] \geq \theta f(z^1) + (1-\theta)f(z^2)$

La estricta concavidad se mantiene, si en (2.8) se reemplaza $\geq$ por >. Supongamos los siguientes bienes y precios

	x_i	p_i	$\sum x_i\, p_i$
1	2	4	8
2	4	5	20
3	6	4.5	27

Y definamos p_3 como $\theta p_1+ (1-\theta)p_2$ y si θ es igual a 0.5 entonces p_3=4.5. Entonces si

$C(u,p_1)= \sum x_1p_1= 8 \leq \sum x_3p_1= 24$
$C(u,p_2)= \sum x_2p_2= 20 \leq \sum x_3p_2= 30$

Se deberá demostrar que $C(u, \theta p_1+ (1-\theta)p_2)\geq \theta[C(u, p_1)]+ (1-\theta)[C(u, p_2)]$ luego si $C(u, \theta p_1+ (1-\theta)p_2)= C(u,p_3)= \sum x_3p_3= 27$ y dado que $\theta[C(u, p_1)]=0.5*8$ y $(1-\theta)[C(u, p_2)]=10$, entonces

$C(u, \theta p_1+ (1-\theta)p_2)\geq \theta[C(u, p_1)]+ (1-\theta)[C(u, p_2)]$ ya que 27 >14.

4) La función de gasto es continua en p y la primera y la segunda derivada con respecto a los precios existe.
5) Cuando ellas existan, las derivadas parciales de las funciones de gasto con respecto a los precios serán las funciones de demandas Hicksianas

(2.9) $$\frac{\partial C(u,p)}{\partial p_i} = h_i(u,p) = x_i^{h}$$

La anterior propiedad se conoce también como el **Lema de Sheppard.**

2.5.3 Propiedades de las funciones de demandas Marshallianas y Hicksianas

Como se ha encontrado anteriormente, de la solución a (2.5) se obtiene la demanda Marshalliana mientras de la solución a (2.6) se obtiene la demanda Hicksiana. Ahora, derivaremos algunas de las propiedades para dichas demandas

2.5.3.1 Adición

El valor total de las demandas Hicksianas y Marshallianas serán los gastos totales

(2.10) $$\sum_{i=1}^{m} p_i h_i(u,p) = \sum_{i=1}^{m} p_i g_i(y,p) = y$$

2.5.3.2 Homogeneidad

Las demandas Hicksianas son homogéneas de grado cero en precios, las demandas Marshallianas en el gasto total y en los precios también lo son, esto es, para algún escalar θ >0, se cumple

(2.11) $h_i (u, \theta p) = h_i (u, p)= g_i (\theta y, \theta p) = g_i (y, p)$

2.5.3.3 Simetría

Las derivadas transversales de los precios, en las demandas Hicksianas son simétricas

(2.12) $$\frac{\partial h_i(u,p)}{\partial p_i} = \frac{\partial h_j(u,p)}{\partial p_j}$$

Para todo i $\neq$ j. Esta propiedad además se deriva del teorema de Young.

2.5.3.4 Negatividad

La matriz de $n \times n$ formada de los elementos de $\frac{\partial h_i}{\partial p_j}$ es semidefinida negativa

(2.13) $$\sum_i \sum_j \lambda_i \lambda_j \frac{\partial h_i}{\partial p_j} \leq 0$$

Si λ es proporcional a p, la desigualdad llegara a ser una igualdad y, la forma cuadrática (2.13) será cero. El resultado anterior se mantiene en tanto $\frac{\partial h_i}{\partial p_j}$ es la matriz de segundas derivadas de la función de gasto, que es una función cóncava y semidefinida negativa. Por conveniencia $\frac{\partial h_i}{\partial p_j}$ se reemplaza por s_{ij} y se conoce como la matriz de sustitución o la matriz de Slutzky, donde los elementos diagonales de dicha matriz serán no positivos para todo i, esto es

(2.14) $s_{ii} \leq 0$

De esta forma, un incremento en precios con la utilidad constante deberá generar una mayor demanda para aquel bien o aquellos que permanecen sin cambiar. La expresión (2.14) podría considerarse también como la famosa **"ley de la demanda"**. Por supuesto (2.14) nos muestra una función de demanda compensada, pero a diferencia de la visión tradicional (2.14) proviene de la función de gasto, la cual es cóncava aunque las preferencias sean o no convexas. La propiedad de negatividad, no nos dice nada acerca de las curvas de indiferencia y, por supuesto si éstas fueran cóncavas al origen dichas demandas mostrarían soluciones de esquina, las cuales no son explicadas por (2.14).

De esta forma, las propiedades básicas de las curvas de demanda son la adición, la homogeneidad de grado cero en precios e ingreso. Por lo tanto, serán simétricas las respuestas en precios y formarán una matriz semidefinida negativa. Si la simetría y negatividad son comprobables entonces es posible observar la matriz de sustitución de

Slutzky. Al derivar la Hicksiana con respecto a p_j y usando el lema de Sheppard encontraremos

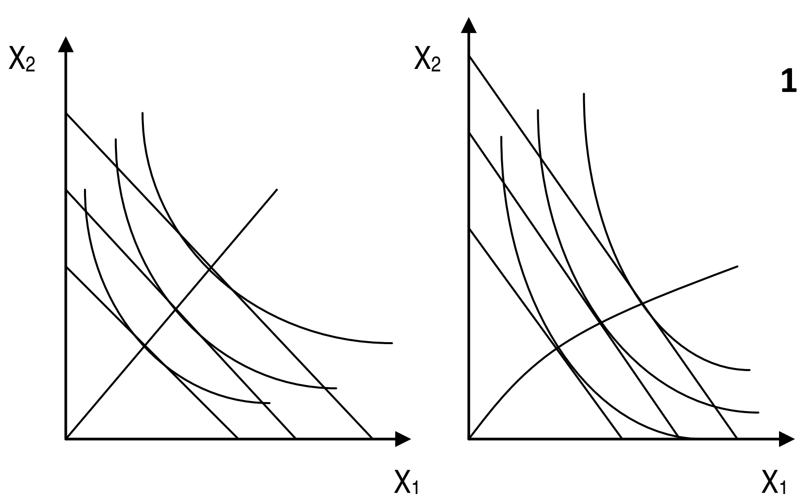
$$(2.15)\quad s_{ij} = \frac{\partial h_i}{\partial p_j} = \frac{\partial g_i}{\partial y} x_j + \frac{\partial g_i}{\partial p_j}$$

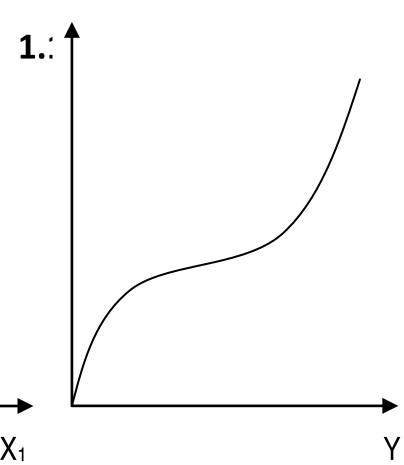
Donde el último término es la derivada de precios no compensados con respecto a p_j. El término compensado se refiere a la cantidad de veces x_j (la derivada del mínimo costo con respecto a p_j) que $\frac{\partial g_i}{\partial x}$ (el gasto total derivado de x_i) deberá ser adicionado. Cada una de las magnitudes de (2.15) en principio puede observarse directamente variando a x y p. La ecuación (2.15) se descompone en un efecto sustitución del cambio en precios y un efecto ingreso $-x_j \frac{\partial g_i}{\partial y}$. Un s_{ij} positivo solo puede ocurrir en el caso de un bien inferior, por ejemplo bienes giffen, ya que todos los bienes giffen son inferiores pero lo contrario no es cierto. Los bienes giffen son muy raros y sólo se citan casos excepcionales como la papa ya que si aumenta su precio su demanda no disminuirá. La matriz de sustitución tiene una función importante en clasificar los bienes: Si los bienes i e j son complementarios s_{ij} es negativo y si los bienes i e j son sustitutos s_{ij} es positivo.

2.6 Trayectorias de expansión

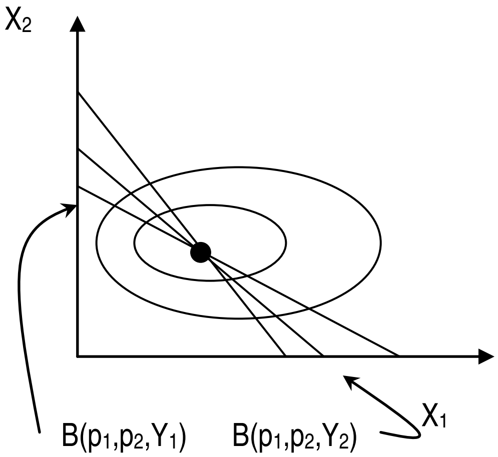
Usualmente, la función de demanda cambia ante un cambio en precios o ingreso, éste cambio puede observarse en términos de la estática comparativa: Suponga que los precios están fijos, pero el ingreso del consumidor lentamente se incrementa, entonces a partir de la colección de puntos resultantes se podría trazar una trayectoria en el ortante no negativo que se denomina trayectoria de expansión del ingreso. Esta trayectoria puede ser proyectada, en un plano definido por dos bienes, mostrando dicha trayectoria la expansión del ingreso relativo a estos dos bienes de la siguiente forma

Gráfica 2.8 a. y b. Trayectorias de expansión **Gráfica 2.8. c. Curva de Engel**

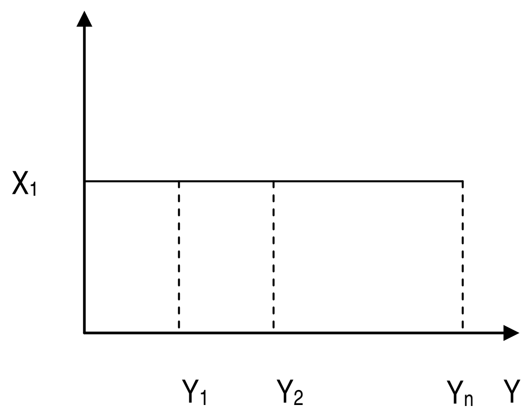




Gráfica 2.8. d. Puntos de felicidad y restricciones presupuestarias B_i



Gráfica 2.8. e. Curva de Engel para la gráfica 2.8. d



Como podrá observarse, la gráfica (2.8.a) muestra un conjunto de bienes normales: el consumo de los bienes x_1 y x_2 se incrementa cuando el ingreso se incrementa. En la gráfica (2.8.b) el bien x_2 es un bien inferior, esto es, el consumo cae en tanto el ingreso aumenta. Cuando la demanda de un bien se grafica como una función del ingreso, el resultado se conoce como la *curva de Engel* para dicho bien. Deberá además observarse, que en el caso de que las preferencias sean homotéticas, se cumple que $u(tx)= tu(x)\ \forall t > 0$, entonces, la trayectoria de expansión y la curva de Engel será una línea de trazo continuo como en la gráfica (2.8.c).

2.7 La tasa marginal de sustitución

Considere la curva de indiferencia $\mu(x_1,x_2) = \mu_1$ cuya representación viene dada por la gráfica (2.6) para algún valor fijo de μ. Esta curva puede ser pensada como la función $x_2(x_1)$. De dichas curvas se define la tasa marginal de sustitución del bien 2 por el bien 1 como

(2.16) $TMS_{21} = -\dfrac{\partial x_2}{\partial x_1}$

La tasa marginal de sustitución es la pendiente de la curva de indiferencia y, su sentido económico, no es otro, que la cantidad que se está dispuesto a renunciar del consumo del bien 1 por consumir unidades adicionales del bien 2, por esta razón la tasa marginal de sustitución definida de la anterior forma decrece cuando x_1 crece. Si nosotros diferenciamos $u(x_1,x_2) = U_1$ con respecto a x_1, se encuentra

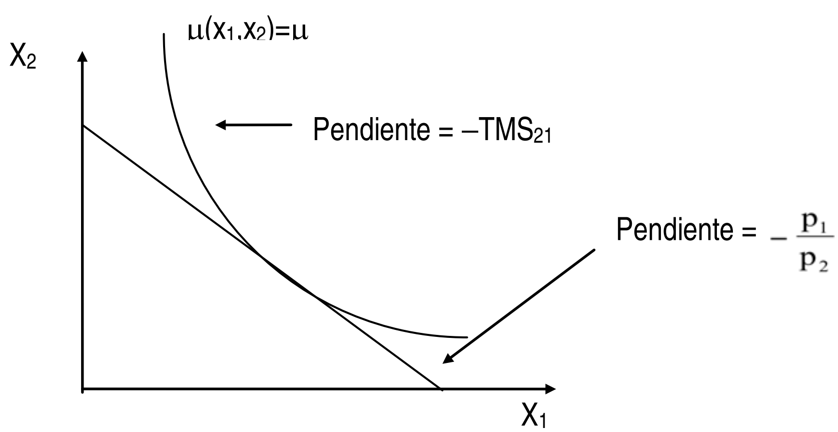
(2.17)
$$\frac{\partial u(x_1,x_2)}{\partial x_1} + \frac{\partial u(x_1,x_2)}{\partial x_2}\frac{\partial x_2}{\partial x_1} = 0$$
$$\text{de donde :} TMS_{21} = -\frac{\partial u(x_1,x_2)}{\partial x_1} / \frac{\partial u(x_1,x_2)}{\partial x_2}$$

De igual forma, la solución al problema básico del consumidor nos muestra

(2.18) $$\frac{\partial u(x_1,x_2)/\partial x_1}{\partial u(x_1,x_2)/\partial x_2} = \frac{p_1}{p_2}$$

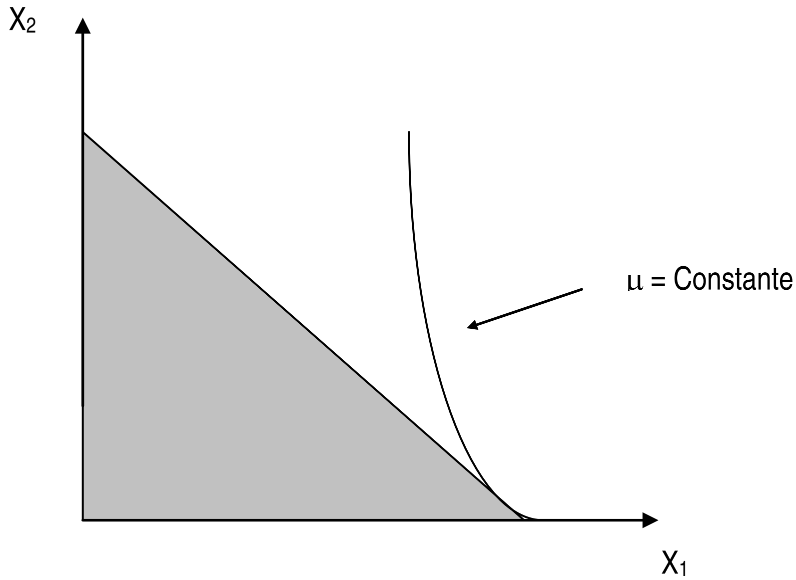
Donde (2.18) es la tasa de sustitución económica del bien 2 por el bien 1, y ésta es la pendiente de la línea de restricción presupuestaria.

Gráfica 2.9. Tasa marginal de sustitución



Se asume que la solución anterior ocurre en un punto x con x > 0, esto es posible incluso para uno o más componentes de x que sean cero, tal es el caso de las soluciones de esquina como en (2.9) veamos

Gráfica 2.10. Solución de esquina



La gráfica anterior en el caso de dos dimensiones nos muestra que si la solución ocurre en el punto $x_1 > 0$ y $x_2 = 0$ la $TMS_{21} > \frac{p_1}{p_2}$, este resultado se extiende a dimensiones mayores.

2.8 Elasticidad

Cuando se discute la sensibilidad de la demanda del consumidor ante cambios en variables como el precio o el ingreso, se puede medir directamente dicha sensibilidad, por ejemplo, $\partial x/\partial y$ en el caso de la sensibilidad en el ingreso. Una de las desventajas de esta medida, es la dependencia sobre las unidades usadas. Se ha vuelto común, usar la elasticidad ingreso de la demanda, cuya ecuación es

(2.19) $$\varepsilon_j = \frac{\partial x_j / x_j}{\partial y / y} = \frac{\partial Lnx_j}{\partial Lny}$$

Así (2.19) es el porcentaje de cambio en la demanda ante un cambio porcentual en el ingreso. Como un resultado adicional, se deberá observar que la elasticidad ingreso promedio deberá ser unitaria, esto es, $k_1\varepsilon_1+ ,\ldots\ldots , + k_1\varepsilon_1= 1$ donde $k_j = \frac{p_j x_j}{y}$ es la proporción del ingreso gastado en la bien j.

De igual forma, se puede derivar las elasticidades de las demandas de un bien x como

$$(2.19) \quad \varepsilon_j = \frac{\partial x_j / x_j}{\partial p_j / p_j} = \frac{\partial Ln x_j}{\partial Ln p_j}$$

2.9 Algunas formas funcionales

En este apartado, se desarrollará primero un ejemplo usual de maximización y luego se discutirán algunas formas funcionales tradicionales en la demanda. Suponga un consumidor, cuya función de utilidad viene definida por $u(x_1,x_2) = x_1^2+x_2^2$. La restricción presupuestaria viene dada por $p_1x_1+p_2x_2 = Y$. Para solucionar este problema se usará el Lagrangiano y deduciremos las condiciones de primer orden. El problema se plantea como

Max $\quad u(x_1,x_2) = x_1^2+x_2^2$

Sujeto a $\quad p_1x_1+p_2x_2 = Y$

$$\ell = x_1^2 + x_2^2 - \lambda (p_1x_1 + p_2x_2 - Y)$$

$$(2.20) \quad \frac{\partial \ell}{\partial x_1} = 2x_1 - p_1\lambda = 0$$

$$(2.21) \quad \frac{\partial \ell}{\partial x_2} = 2x_2 - p_2\lambda = 0$$

$$(2.22) \quad \frac{\partial \ell}{\partial \lambda} = p_1x_1 + p_2x_2 - Y = 0$$

¿Cuál es el significado económico del multiplicador Lagrangiano λ? Se puede observar, que el multiplicador no es más que la razón de cambio del máximo (o mínimo) valor de la función objetivo con respecto a un cambio paramétrico en el valor de la restricción, esto es, supongamos que la función objetivo para un individuo está determinada por el trabajo y el ocio, sujeta a la restricción del tiempo en algún nivel k: Si un incremento adicional del tiempo, ocurriese en Δk unidades, el ingreso se incrementaría en $\Delta y^* \equiv \lambda^*\Delta k$, en otras palabras λ^* será el valor marginal del tiempo o el costo de oportunidad del mismo y, en una economía competitiva, las empresas estarían dispuestas a pagar λ^* por cada incremento en el tiempo, esto significa que λ^* podría ser visto como el precio de reserva del tiempo, en últimas el salario de reserva.

Despejando (2.20) y (2.21) e igualando obtenemos

(2.23) $$x_1 = \frac{x_2 p_1}{p_2}$$

Sustituyendo (2.23) en (2.22) se obtiene

(2.23) $$p_1\left[\frac{x_2 p_1}{p_2}\right] + p_2 x_2 - Y = 0 \Rightarrow x_2\left[\frac{p_1^{\;2} + p_2^{\;2}}{p_2}\right] = Y$$

De donde se deduce que la demanda Marshalliana viene determinada por

(2.24) $$x_2^{\;m} = \frac{Yp_2}{p_1^{\;2} + p_2^{\;2}}$$

De igual forma, reemplazando esta solución en (2.22) se obtiene

(2.25) $$x_1^{\;m} = \frac{Yp_1}{p_1^{\;2} + p_2^{\;2}}$$

La función indirecta de utilidad se encuentra reemplazando (2.24) y (2.25) en la función de utilidad

$$U^*(x_1, x_2) = x_1^{\;2} + x_2^{\;2}$$

(2.26) $$= \left[\frac{Yp_1}{p_1^{\;2} + p_2^{\;2}}\right]^2 + \left[\frac{Yp_2}{p_1^{\;2} + p_2^{\;2}}\right]^2$$

$$= \frac{Y^2}{p_1^{\;2} + p_2^{\;2}}$$

A partir de (2.26) se puede usar la identidad de Roy, obteniendo

(2.27) $$-\frac{\partial U^*}{\partial p_1} = -\frac{0 - 2p_1Y^2}{\left(p_1^{\;2} + p_2^{\;2}\right)^2} = \frac{2p_1Y^2}{\left(p_1^{\;2} + p_2^{\;2}\right)^2}$$

(2.28) $$\frac{\partial U^*}{\partial Y} = \frac{2Y\left(p_1^{\;2} + p_2^{\;2}\right)}{\left(p_1^{\;2} + p_2^{\;2}\right)^2} = \frac{2Y}{\left(p_1^{\;2} + p_2^{\;2}\right)}$$

De donde se deduce

$$(2.29)\quad -\frac{\partial U^*/\partial p_1}{\partial U^*/\partial Y}=\frac{\dfrac{2p_1Y^2}{\left(p_1{}^2+p_2{}^2\right)^2}}{\dfrac{2Y}{\left(p_1{}^2+p_2{}^2\right)}}$$

Por lo cual la demanda será

$$(2.30)\qquad x_1{}^m=\frac{Yp_1}{p_1{}^2+p_2{}^2}$$

De acuerdo a la dualidad

Min $\quad p_1x_1+p_2x_2$
Sujeto a $\quad u(x_1,x_2) = x_1{}^2+x_2{}^2 = u^0$

Así, el Lagrangiano se plantea como

$$\ell = p_1x_1+p_2x_2+\lambda(x_1{}^2+x_1{}^2-u^o)$$

$$(2.31)\ \frac{\partial\ell}{\partial x_1}=p_1+\lambda 2x_1=0$$

$$(2.32)\ \frac{\partial\ell}{\partial x_2}=p_2+\lambda 2x_2=0$$

$$(2.33)\ \frac{\partial\ell}{\partial\lambda}=x_1{}^2+x_2{}^2-u^0=0$$

Despejando (2.31) y (2.32) e igualando

$$(2.34)\ x_2=\frac{x_1p_2}{p_1}$$

Sustituyendo (2.34) en (2.33) se obtiene

$$(2.35)\ x_1{}^2+\left[\frac{x_1{}^2p_2{}^2}{p_1{}^2}\right]=u^0 \Rightarrow x_1{}^2\left[\frac{p_1{}^2+p_2{}^2}{p_1{}^2}\right]=u^0$$

De donde se deduce que la demanda Hicksiana viene determinada por

(2.36) $$x_1^h = \frac{u^{\frac{1}{2}} p_1}{(p_1^2 + p_2^2)^{\frac{1}{2}}}$$

Usando el Lema de Sheppard, se obtiene

$$= \frac{p_1^2 u^{\frac{1}{2}}}{\left(p_1^2 + p_2^2\right)^{\frac{1}{2}}} + \frac{p_2^2 u^{\frac{1}{2}}}{\left(p_1^2 + p_2^2\right)^{\frac{1}{2}}}$$

(2.37)C(u,p_1,p_2) $$= \frac{p_1^2 u^{\frac{1}{2}} + p_2^2 u^{\frac{1}{2}}}{\left(p_1^2 + p_2^2\right)^{\frac{1}{2}}}$$

$$= u^{\frac{1}{2}} \left(p_1^2 + p_2^2\right)^{\frac{1}{2}}$$

Por lo cual

(2.38) $$\frac{\partial C^*(u, p_1, p_2)}{\partial p_1} = 0 + \frac{1}{2}\left(p_1^2 + p_2^2\right)^{-\frac{1}{2}} 2 p_1 u^{\frac{1}{2}}$$

$$x_1^h = \frac{u^{\frac{1}{2}} p_1}{\left(p_1^2 + p_2^2\right)^{\frac{1}{2}}}$$

; De igual forma (2.39) $$x_2^h = \frac{u^{\frac{1}{2}} p_2}{\left(p_1^2 + p_2^2\right)^{\frac{1}{2}}}$$

Veamos otro ejemplo. Considere la siguiente función de utilidad y la restricción presupuestaria

(2.20.1) $U(x_1,x_2) = x_1^{\alpha} x_2^{(1-\alpha)}$

(2.20.2) $p_1 x_1 + p_2 x_2 = y$

Solucionando el problema del consumidor tendremos

(2.20.3) $\ell = x_1^{\alpha} x_2^{(1-\alpha)} - \lambda(p_1 x_1 + p_2 x_2 - y)$

(2.20.4) $\frac{\partial \ell}{\partial x_1} = \alpha x_1^{\alpha-1} x_2^{1-\alpha} - \lambda p_1 = 0$ (2.20.5) $\frac{\partial \ell}{\partial x_2} = (1-) x_1^{\alpha} x_2^{-\alpha} - \lambda p_2$

(2.20.6) $$\frac{\partial \ell}{\partial \lambda} = p_1 x_1 + p_1 x_2 - y$$

Igualando (2.20.3) y(2.20.4) se obtiene

(2.20.5) $$\frac{\alpha x_1^{\alpha-1} x_2^{1-\alpha}}{p_1} = \frac{(1-\alpha) x_1^{\alpha} x_2^{-\alpha}}{p_2} = \frac{x_2}{x_1} = \frac{p_1(1-\alpha)}{p_2 \alpha}$$

Sustituyendo en (2.20.6) se obtienen las siguientes demandas Marshallianas

(2.20.6) $x_1^{m} = \dfrac{(1-\alpha)y}{p_2}$

(2.20.7) $x_2^{m} = \dfrac{\alpha y}{p_1}$

La función indirecta de utilidad será

$V(p_1,p_2,y) = \left(\dfrac{\alpha y}{p_1}\right)^{\alpha} \left(\dfrac{(1-\alpha)y}{p_2}\right)^{(1-\alpha)}$

2.9.1 El sistema Lineal de Gasto

El sistema lineal de gasto (Linear Expenditure System) es una generalización de la función de utilidad Cobb-Douglas, fue desarrollado por Klein y Rubin (1947-48) y Samuelson (1947-48). Investigada empíricamente por Stone (1954) y Geary (1950), por lo cual, se le da el nombre Stone-Geary. El sistema lineal de gasto es básicamente una Cobb-Douglas trasladada en el origen al punto (B_1 ,B_2), en el cuadrante positivo

(2.40) $\mu(X_1, X_2) = (X_1 - B_1)^{\alpha_1} (X_2 - B_2)^{\alpha_2}$

(2.41) $V(X_1, X_2) = \alpha_1 \mathrm{Ln}(X_1 - B_1) + \alpha_2 \mathrm{Ln}(X_2 - B_2)$

por la proposición 2.1(pág.15).

Supongamos la restricción

(2.42) $X_1 P_1 + X_2 P_2 = Y$

Entonces el problema de maximizar (2.41) sujeto a (2.42) se resuelve con

el Lagrangiano de la siguiente forma

(2.43) $\ell = \alpha_1 Ln(X_1 - B_1) + \alpha_2 Ln(X_2 - B_2) - \lambda(X_1 P_1 + X_2 P_2 - Y)$

(2.44) $\dfrac{\partial \ell}{\partial X_1} = \dfrac{\alpha_1}{(X_1 - B_1)} - \lambda P_1 = 0$ (2.45) $\dfrac{\partial \ell}{\partial X_2} = \dfrac{\alpha_2}{(X_2 - B_2)} - \lambda P_2 = 0$

(2.46) $$\frac{\partial \ell}{\partial \lambda} = X_1 P_1 + X_2 P_2 - Y = 0$$

Igualando (2.44) y (2.45) se obtiene

$$\frac{\alpha_1}{(X_1 - B_1)P_1} = \lambda \text{ ; } \frac{\alpha_2}{(X_2 - B_2)P_2} = \lambda \quad \Rightarrow \alpha_1(X_2 - B_2)P_2 = \alpha_2(X_1 - B_1)P_1$$

$$\Rightarrow \alpha_1 X_2 P_2 - \alpha_1 B_2 P_2 = \alpha_2 X_1 P_1 - \alpha_2 B_1 P_1$$

$$X_2 P_2 = \frac{\alpha_2}{\alpha_1} x_1 P_1 - \frac{\alpha_2}{\alpha_1} B_1 P_1 + B_2 P_2$$

Sustituyendo en (2.46) se obtiene

$$X_1 P_1 + \frac{\alpha_2}{\alpha_1} X_1 P_1 - \frac{\alpha_2}{\alpha_1} B_1 P_1 + B_2 P_2 = Y$$

$$X_1 P_1 + \frac{\alpha_2}{\alpha_1} X_1 P_1 = Y + \frac{\alpha_2}{\alpha_1} B_1 P_1 - B_2 P_2$$

$$X_1 P_1 (1 + \frac{\alpha_2}{\alpha_1}) = Y + \frac{\alpha_2}{\alpha_1} B_1 P_1 - B_2 P_2$$

$$X_1 P_1 (\frac{\alpha_1 + \alpha_2}{\alpha_1}) = Y + \frac{\alpha_2}{\alpha_1} B_1 P_1 - B_2 P_2 \qquad \text{Sí } \alpha_1 + \alpha_2 = 1$$

$$X_1 P_1 \frac{1}{\alpha_1} = Y + \frac{\alpha_2}{\alpha_1} B_1 P_1 - B_2 P_2 \qquad \Rightarrow$$

$$X_1 = Y \frac{\alpha_1}{P_1} + \frac{\alpha_2}{\alpha_1} \frac{\alpha_1}{P_1} B_1 P_1 - B_2 P_2 \frac{\alpha_1}{P_1} \qquad \Rightarrow$$

$$X_1 = Y \frac{\alpha_1}{P_1} - B_2 P_2 \frac{\alpha_1}{P_1} + \frac{(1-\alpha_1)}{\alpha_1} B_1 P_1 \frac{\alpha_1}{P_1} \qquad \Rightarrow$$

$$X_1 = Y \frac{\alpha_1}{P_1} - B_2 P_2 \frac{\alpha_1}{P_1} + \frac{B_1 P_1 \alpha_1}{\alpha_1 P_1} - B_1 P_1 \frac{\alpha_1}{P_1} \qquad \Rightarrow$$

(2.47) $$X_1{}^m = B_1 + \frac{\alpha_1}{P_1} (Y - B_2 P_2 - B_1 P_1)$$

(2.48) $X_2^{\,m} = B_2 + \dfrac{\alpha_2}{P_2}(Y - B_2P_2 - B_1P_1)$

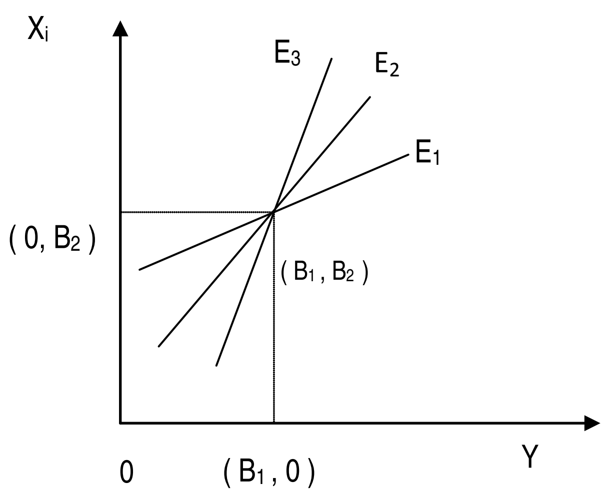
Las ecuaciones (2.47) y (2.48) son las demandas Marshallianas. Escribiendo las funciones de demanda en forma de gasto

(2.49) $X_1^{\,m} P_1 = P_1B_1 + \dfrac{\alpha_1 P_1}{P_1}(Y - B_2P_2 - B_1P_1)$

(2.50) $X_2^{\,m} P_2 = P_2B_2 + \alpha_2(Y - B_2P_2 - B_1P_1)$

De esta forma, el gasto en cada bien X_iP_i es lineal en precios e ingreso. El LES describe a unos consumidores comprando primero las cantidades de subsistencia de cada bien (B_i) y, dividiendo lo que queda del gasto, entre los bienes, en proporciones fijas (α_1, α_2). El gasto marginal en cada bien es constante y, el LES tiene una curva de Engel lineal, que no es homotética ya que todas las líneas de ingreso - gasto pasan a través del punto B_i

Gráfica 2.11. Sendas de expansión para una LES no homotética (E_i= Curva de Engel; Y= Ingreso)



Howe, Pollack y Wales(1979) estiman un LES usando los siguientes bienes: alimentos, ropa, abrigos y misceláneos denominados por f,c,s y m respectivamente, los datos fueron tomados del consumo en estados unidos entre 1929 y 1975 excluyendo los años de guerra(1942-1945), el resultado obtenido fue:

Parámetro	Valor	Desviación Estándar
α_f	0.38	0.022
α_c	0.24	0.015
α_s	0.17	0.02
α_m	0.21	-

Los α_i {f,c,s,m} son las partes asignadas del presupuesto marginalmente y, éstas son necesariamente independientes de los precios, del gasto y del consumo pasado. La restricción implícita como se puede observar, consiste en $\sum\alpha_i^4 = 1$, debido a que no se puede gastar más de lo que se tiene en cada bien y, que no tiene sentido gastar menos. De igual forma, para 1975 Howe et-al encontraron unos valores α_i {f,c,s,m} de 0.32, 0.12,0.36 y 0.20 respectivamente.

La concavidad en la función de gasto se satisface por el hecho de que los α_i son positivos y, x_i no es menor que $\sum_i p_i B_i$ ya que $x_i \geq B_i \ \forall$ i. Si dichas restricciones no se mantienen, la función de gasto no es cóncava. Deaton y Muelbauer(1981) indican que si observamos la cantidad $\sum_i p_i \alpha_i$ ésta nos muestra que no existen efectos de sustitución y entonces se deberá pensar en el LES como una función de utilidad "comprada" a un precio constante por unidad $\prod p_i^{\alpha_i}$. De igual forma, los α´s pueden ser pensados como la media geométrica de los precios y, de esta forma, como un índice de precios que representa el "costo marginal de vida".

Para obtener un indicador "real" de riqueza, deberemos partir de que si los B´s representan los requisitos de subsistencia, solamente Y - $\sum_i p_i B_i$ es lo que queda para asignarse en forma discrecional y, si deflactamos lo que queda por la media geométrica ponderada de los precios α´s, entonces obtendremos dicho indicador. Para una aplicación reciente a la demanda por cultura ver Mora, Bernat y Zuluaga (2012).

2.9.2 La función de Utilidad CES

La función de utilidad CES surge como una analogía directa a la teoría de la producción (Arrow et-al 1961) y tiene la forma

(2.51) $u(x_1,x_2) = (\alpha_1 x_1^{\rho} + \alpha_2 x_2^{\rho})^{1/\rho}$ con $\rho \leq 1$

Maximizando (2.51) sujeto a la restricción de presupuesto

(2.52) $X_1 P_1 + X_2 P_2 = Y$

Y, utilizando las condiciones de primer orden derivadas de usar el Lagrangiano obtenemos

(2.53) $\alpha_1 x_1^{\rho-1}(\alpha_1 x_1^{\rho} + \alpha_2 x_2^{\rho})^{\frac{(1-\rho)}{\rho}} - \lambda p_1 = 0$

(2.54) $\alpha_2 x_2^{\rho-1}(\alpha_1 x_1^{\rho} + \alpha_2 x_2^{\rho})^{\frac{(1-\rho)}{\rho}} - \lambda p_2 = 0$

(2.55) $Y - X_1 P_1 - X_2 P_2 = 0$

Igualando (2.53) y (2.54) se obtiene

(2.56) $\dfrac{x_1}{x_2} = \left(\dfrac{\alpha_2 p_1}{\alpha_1 p_2}\right)^{\frac{1}{\rho-1}}$

De donde la elasticidad de sustitución vendrá dada por

(2.57) $\sigma = \dfrac{-\partial \mathrm{Log}(x_1^* / x_2^*)}{\partial \mathrm{Log}(p_1 / p_2)} = \dfrac{1}{1-\rho}$

Entre mayor sea el valor del parámetro ρ mayor será el grado de sustitución entre los bienes. Puede observarse, que cuando $\rho = 0$ la CES será una Cobb-Douglas y cuando $\rho \to \infty$ será una Leontief. Las funciones de demandas Marshallianas correspondientes a la CES serán

(2.58) Y - $p_1\left(\dfrac{\alpha_2 p_1}{\alpha_1 p_2}\right)^{-\sigma} x_2 - p_2 x_2 = 0$

De donde se obtiene

(2.59)
$$x_1^m = \frac{(\alpha_1 p_2)^{\sigma} Y}{\alpha_1^{\sigma} p_1 p_2^{\sigma} + \alpha_2^{\sigma} p_2 p_1^{\sigma}}$$
$$x_2^m = \frac{(\alpha_2 p_1)^{\sigma} Y}{\alpha_1^{\sigma} p_1 p_2^{\sigma} + \alpha_2^{\sigma} p_2 p_1^{\sigma}}$$

Dado que las preferencias representadas por la función de utilidad CES son homotéticas, las funciones de demandas marshallianas son lineales en el Ingreso. Para encontrar la función de utilidad indirecta, se sustituye (2.59) en (2.55) y dado que $\rho\sigma = \dfrac{\rho}{(1-\rho)} = \sigma - 1$ obtenemos la función indirecta de utilidad

$$V^{*}=\left(\alpha_1\left[\frac{(\alpha_1 p_2)^{\sigma}Y}{\alpha_1{}^{\sigma}p_1p_2{}^{\sigma}+\alpha_2{}^{\sigma}p_2p_1{}^{\sigma}}\right]^{\rho}+\alpha_2\left[\frac{(\alpha_2 p_1)^{\sigma}Y}{\alpha_1{}^{\sigma}p_1p_2{}^{\sigma}+\alpha_2{}^{\sigma}p_2p_1{}^{\sigma}}\right]^{\rho}\right)^{\frac{1}{\rho}}$$

(2.60)
$$=\left((\alpha_1{}^{1+\rho\sigma}p_2{}^{\rho\sigma}+\alpha_2{}^{1+\rho\sigma}p_1{}^{\rho\sigma})\left[\frac{(\alpha_2 p_1)^{\sigma}Y}{\alpha_1{}^{\sigma}p_1p_2{}^{\sigma}+\alpha_2{}^{\sigma}p_2p_1{}^{\sigma}}\right]^{\rho}\right)^{\frac{1}{\rho}}$$

$$=\frac{Y[\alpha_1{}^{\sigma}p_2{}^{\sigma-1}+\alpha_2{}^{\sigma}p_1{}^{\sigma-1}]^{\frac{1}{(\sigma-1)}}}{p_1p_2}$$

La función de gasto, se obtiene directamente invirtiendo la función indirecta de utilidad

(2.61) $C(u,p_1,p_2)=Up_1p_2[\alpha_1{}^{\sigma}p_2{}^{\sigma-1}+\alpha_2{}^{\sigma}p_1{}^{\sigma-1}]^{\frac{1}{(1-\sigma)}}$

2.9.3 La función de Utilidad Indirecta Addilog

Como se ha podido observar, de los ejercicios anteriores, es posible derivar las funciones de demanda de los bienes a través de maximizar la utilidad, sujeta a la restricción de gasto. Sin embargo, la solución no siempre es estimable. La teoría de la dualidad sugiere, que una alternativa es especificar una función indirecta de utilidad, una función que es no decreciente en el ingreso, no decreciente y cuasi convexa en precios, continua y homogénea de grado cero en precios e ingreso. De esta forma, a partir de la función de utilidad indirecta que corresponda a algún tipo de preferencias del consumidor puede recuperarse la demanda con la identidad de Roy. Una forma funcional es la función "addilog" de utilidad indirecta introducida por Houthakker(1965)

(2.62) $U^{*}(p_1,p_2,Y)=\alpha_1\left(\frac{Y}{p_1}\right)^{\beta_1}+\alpha_2\left(\frac{Y}{p_2}\right)^{\beta_2}$

Las funciones de demandas obtenidas de una "addilog" usando la identidad de Roy serán

(2.63)
$$x_1{}^{m}=\frac{-\partial U^{*}/\partial p_1}{\partial U^{*}/\partial Y}$$
$$=\frac{\alpha_i\beta_i p_1{}^{-\beta_i-1}Y^{\beta_i}}{\alpha_1\beta_1p_1{}^{-\beta_1}Y^{\beta_1-1}+\alpha_2\beta_2p_2{}^{-\beta_2}Y^{\beta_2-1}}\quad ;\quad i=1,2$$

Si nosotros dividimos $x_1{}^{Y}$ por $x_2{}^{Y}$ y tomamos logaritmos el resultado es una logarítmica lineal en el ingreso y en los precios relativos de x_1, x_2

(2.64) $$Log\left(\frac{x_1^{\ Y}}{x_2^{\ Y}}\right) = Log\ \frac{\alpha_1\beta_1 p_1^{\ -\beta_1-1}Y^{\beta_1}}{\alpha_2\beta_2 p_2^{\ -\beta_2-1}Y^{\beta_1}}$$
$$= Log\left(\frac{\alpha_1\beta_1}{\alpha_2\beta_2}\right) - (\beta_1+1)Logp_1 + (\beta_2+1)Logp_2 + (\beta_1-\beta_2)LogY$$

2.9.4 Las especificaciones Translogarítmicas

La función de utilidad translogarítmica proviene de Christensen, Jorgenson y Lau (1971,1975). Esta ha sido la forma funcional más usada en análisis empíricos de demanda. Una de las ventajas de la translogarítmica es su forma funcional flexible; ya que puede ser aproximada de una función de segundo orden por Taylor a una función de utilidad indirecta arbitraria. La especificación translogarítmica básica viene dada por

(2.65) $$LogU^*(p_1,...,p_n,Y) = -\sum_j \alpha_j Log\ \frac{p_Y}{Y} - \frac{1}{2}\sum_k\ \sum_j \beta_{kj} Log\ \frac{p_k}{Y} Log\ \frac{p_Y}{Y}$$

Las restricciones teóricas nos indican que por adición $\sum_j\alpha_j=1$ y por simetría $\beta_{kj}=\beta_{jk}$ $\forall$ k y j. Cuando sea más conveniente trabajar con las ecuaciones de gasto, que las ecuaciones de demanda, entonces se usarán las especificaciones translogarítmica, podrá observarse entonces que

(2.66) $$-\frac{\partial LogU^* / \partial Logp_i}{\partial LogU^* / \partial LogY} = \left(\frac{\partial U^* / \partial p_i}{\partial U^* / \partial Y}\right)\left(\frac{p_i / U^*}{Y / U^*}\right) = \frac{p_i x_i}{Y}$$

La especificación translogarítmica puede denotarse también como

(2.67) $$LogU^* = LogY - \sum_j \alpha_j Logp_j - \frac{1}{2}\sum_k\ \sum_j \beta_{kj}(Logp_Y - LogY)(Logp_j - LogY)$$

Las ecuaciones de gasto se pueden obtener a través de la diferenciación logarítmica de (2.67) y se escriben como

(2.68)
$$x_i p_i = \frac{\alpha_i + \frac{1}{2}\sum_j \beta_{ji} Log(\frac{p_j}{Y}) + \frac{1}{2}\sum_k \beta_{ki} Log(\frac{p_k}{Y})}{1 + \frac{1}{2}\sum_k\ \sum_j \beta_{kj} Log(\frac{p_j}{Y}) + \frac{1}{2}\sum_k\sum_j\ \beta_{kj} Log(\frac{p_k}{Y})}$$
$$= \frac{\alpha_i + \sum_j \beta_{ji} Log(\frac{p_j}{Y})}{1 + \sum_k\ \sum_j \beta_{kj} Log(\frac{p_j}{Y})} ;\forall i = 1,...,n$$

Un caso especial de la translogarítmica, es la translogarítmica homotética, la cual se obtiene a través de imponer las siguientes restricciones

(2.69)$\sum_j \beta_{kj} = 0$, k = 1,…,n

que se conocen como restricciones de Homogeneidad. Dadas estas n restricciones, la función de utilidad indirecta y las ecuaciones de gasto serán

(2.70) $LogU^* = LogY - \sum_j \alpha_j Logp_j - \frac{1}{2}\sum_k \sum_j \beta_{kj} Logp_k Logp_j$

(2.80) $x_i p_i = \alpha_i + \sum_j \beta_{ij} Logp_j ; \forall i = 1,\ldots,n$

La ecuación (2.80) muestra que las partes del gasto son independientes del ingreso, lo cual confirma, que las preferencias son homotéticas. Observe también, que la función indirecta de utilidad (2.79) puede ser invertida para obtener una función de gasto translogarítmica homotética

(2.81) $LogY^*(p_1,\ldots,p_n,u) = Logu + \sum_j \alpha_j Logp_j + \frac{1}{2}\sum_k \sum_j \beta_{kj} Logp_k Logp_j$

La función de gasto (2.81), se usa frecuentemente en estudios empíricos sobre funciones de producción, ya que si se interpreta Y^* como el costo total, entonces las ecuaciones de intensidades factoriales se obtienen a través del Lema de Sheppard.

2.9.5 El sistema Casi - Ideal de Gasto AIDS

El sistema de ecuaciones de demanda puede ser derivado a partir de la función de gasto. Suponiendo que éste es continuo y no-decreciente precios y utilidad y, además cóncavo y homogéneo de grado cero, entonces

(2.82) $LogY^*(p_1,\ldots,p_n,u) = a(p_1,\ldots,p_n) + ub(p_1,\ldots,p_n)$

Donde
$$a(\bullet) = \alpha_0 + \sum_j \alpha_j Logp_j + \frac{1}{2}\sum_k \sum_j \gamma_{kj} Logp_k Logp_j$$
$$b(\bullet) = \beta_0 \prod_j p_j^{\beta_j} \quad ; (1) \sum_j \alpha_j = 0 , (2) \sum_j \beta_j = 0 , \sum_k \gamma_{kj} = 0 , (3)\ \gamma_{kj} = \gamma_{jk}$$

Usando el Lema de Sheppard, las ecuaciones de gasto serán

$$x_i p_i = \frac{\partial LogY^*}{\partial Logp_i}$$

(2.83)
$$\begin{aligned} &= \alpha_i + \sum_j \gamma_{ij} Logp_j + p_i u \frac{\beta_i}{p_i} b(\bullet) \\ &= \alpha_i + \sum_j \gamma_{ij} Logp_j + p_i \frac{LogY - a(\bullet)}{b(\bullet)} \frac{\beta_i}{p_i} b(\bullet) \\ &= \alpha_i + \sum_j \gamma_{ij} Logp_j + \beta_i Log \frac{Y}{p_i}, \forall i = 1,\ldots,n \end{aligned}$$

Deaton y Muellbauer arguyen que p puede ser considerado como un índice de precios y, aproximado por $\sum_j x_i p_i Logp_i$. Dada esta aproximación, el sistema de ecuaciones de demanda serán lineales en el logaritmo de precios e ingreso real.

El sistema AIDS (Almost Ideal Demand Sistem) cumple las restricciones de adición (1), homogeneidad (2) y simetría (3). Para satisfacer las condiciones de negatividad se requiere que la matriz de Slutzky sea semidefinida negativa, esto es

(2.84) $c_{ij} = \gamma_{ij} + \beta_i \beta_j \text{ Log } (Y/p) - x_i p_i \delta_{ij} + x_j p_j x_i p_i$

Donde δ es el producto de kronecker [7]que será igual a 1 sí i=j y 0 de lo contrario. Los resultados más importantes del AIDS consisten en que (2.84) es lineal y puede ser estimado por mínimos cuadrados ordinarios. Las restricciones sobre α y γ aseguran que p sea lineal, aunque en muchas estimaciones p pueda resultar colineal, si se usa algún índice de precios se elimina dicho problema. Los β´s del AIDS determinan cuando los bienes son de lujo o necesarios: si $\beta_i > 0$ el gasto($x_i p_i$) se incrementa con x, por lo cual, el bien i será de lujo, de forma similar $\beta_i < 0$ el bien i será necesario. Los γ_{ij} miden el cambio en la i-esíma parte del gasto siguiendo un cambio proporcional en p_j con (Y/ P) permaneciendo constante.

2.9.6 El modelo de Rotterdam

Propuesto inicialmente por Theil(1965) y Barten(1966). Este modelo es parecido al Stone-Geary, solo que en lugar de trabajar con los niveles de los logaritmos se usan las diferencias de los mismos, esto es, diferenciando (2.51) se obtiene

(2.85) $\partial Logx_i = e_i \partial LogY + \sum_j e_{ij} \partial Logp_j$

Se supone en (2.85) que las elasticidades e_i y e_{ij} permanecen constantes. Usando la descomposición de Slutzky y obteniendo $e_{ij} = e_{ij}^* - e_i x_i p_i$ donde e_{ij}^* es la elasticidad cruzada de los precios

[7] Sí **A** es(m×n) y **B** es (p×q) entonces se dice que el producto de Kronecker **A** $\otimes$ **B** = {b_{ij} **A** }$_{mp,nq}$

(2.86) $\partial Logx_i = e_i(\partial LogY - \sum x_k p_k \partial Logp_k) + \sum_j e_{ij}^* \partial Logp_j$

Las restricciones se efectúan también sobre las ecuaciones de gasto de esta forma se obtiene

(2.87) $x_i p_i \partial Logx_i = b_i \partial Log \bar{x} + \sum_j c_{ij} \partial Logp_j$

Donde:

(2.88) $\partial Log \ \bar{x} = \partial Log\ x - \sum_k x_k p_k Logp_k = \sum_k x_k p_k \partial Logx_k$

(2.89) $b_i = x_i p_i e_i = p_i \dfrac{\partial x_i}{\partial Y}$

(2.90) $c_{ij} = x_i p_i e_i^* = \dfrac{p_i p_j s_{ij}}{Y}$

Donde s_{ij} es el (i , j) término de la matriz de sustitución de Slutzky. Observe que (2.88) es un índice que representa el cambio proporcional en el gasto total real, mientras que (2.90) representa la demanda Hicksiana y (2.89) es la propensión marginal a gastar en el i- ésimo bien.

La propiedad de adición requiere que las propensiones marginales a gastar en cada bien sumen uno y que el efecto neto de un cambio de precio en el presupuesto sea cero. Algunas pruebas sobre homogeneidad en (2.88) han sido propuestas por Barten(1969) y Deaton (1974). Ver además Deaton et-al (pag. 71-73).

3 La demanda del consumidor

La demanda representa la cantidad que un consumidor desea comprar de una serie de bienes, ya sea expresada como una función de los precios y el ingreso o como una función de la utilidad y de los precios.

Debemos partir de que el comportamiento del consumidor es racional. Si las decisiones que toma el consumidor contradicen los supuestos, entonces dicho consumidor se considerará como irracional. De hecho, un estudio en sicóticos crónicos realizado en una institución mental en New York (USA) encontró que aquellas personas a quien la sociedad considera como "irracionales" siguen la famosa ley de la demanda "compran menos cuando aumentan los precios" (ver Battalio et-al 1973).

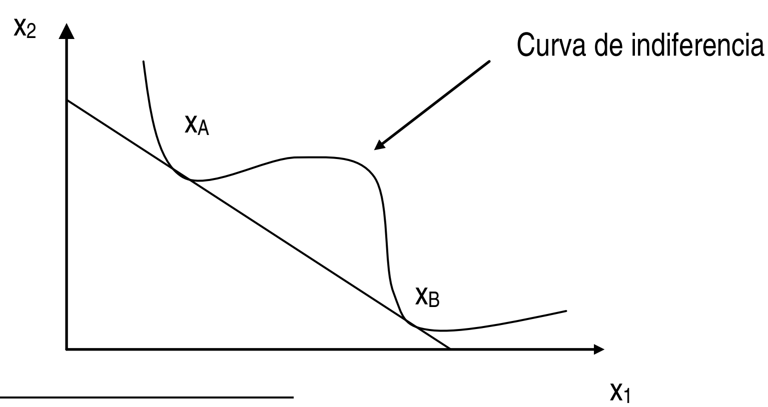
Obviamente la discusión sobre "el comportamiento racional", es más profunda de lo que se considera aquí, ya que no necesariamente se necesita estar en una institución mental para caracterizar a un individuo como "irracional" o que comprar menos, cuando aumentan los precios ratifique el comportamiento maximizador de los individuos [8].

3.1 Unicidad y continuidad

La demanda que corresponde a un vector de precios e ingreso podría no ser única, como se observa en la gráfica (3.1), allí existen dos soluciones x_A y x_B correspondientes a la restricción de presupuesto. Desde un punto de vista técnico, la condición que garantiza una función de demanda única consistirá en

"Si un orden de preferencias es continuo, satisface la insaciabilidad local, y es estrictamente convexo, entonces para todo $p \gg 0$, $y > 0$ la demanda $x(p, y)$ es única, define un valor singular, y es una función continua de (p, y)."

Gráfica 3.1. Soluciones no-únicas



[8] Una "buena" discusión sobre el problema de la racionalidad se presenta en la Parte III del libro "Rational Behaviour From An Experimental Approach" del libro de Arrow,K..,Colombato, E.,Perlman, M y Schmidt,Ch. (1996), y en McFfadden (1999).

3.2 El excedente del consumidor y disponibilidad a pagar

La teoría del consumidor, nos muestra un individuo eligiendo una canasta de bienes, dados unos precios e ingresos ¿Qué sucede cuando el entorno que rodea al consumidor cambia? Este será el objetivo de esta sección.

Consideremos a un consumidor con preferencias racionales, continuas, y localmente no saciadas. Asumiremos también que las funciones de gasto y utilidad indirectas son diferenciables, y concentraremos nuestro interés en los cambios de precios. Suponga que la riqueza del consumidor permanece constante a un nivel $y > 0$ y el vector de precios iniciales es p^0. Nosotros deseamos evaluar el impacto sobre la riqueza del consumidor, de un cambio de p^0 a un nuevo vector de precios p^1. Dicho cambio no debe parecernos extraño, si el gobierno decide aumentar los impuestos esto se traducirá directamente en los precios.

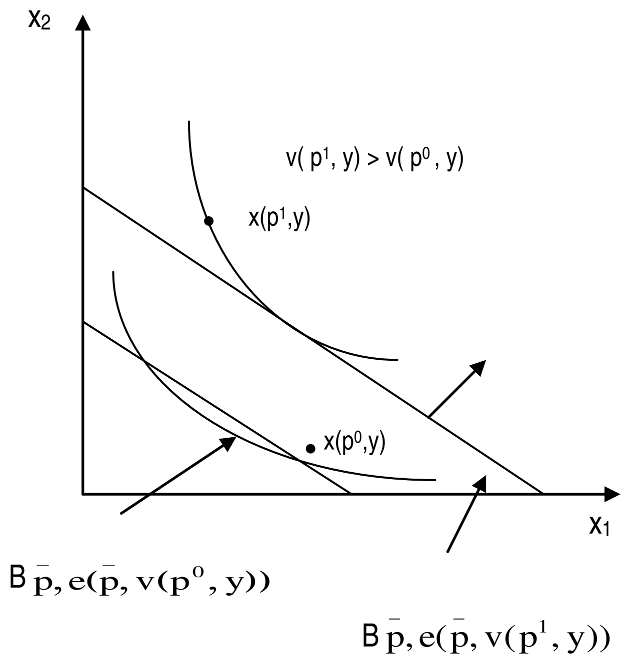
El problema se traduce en evaluar cuándo el consumidor estará mejor o peor, esto es, dada la función indirecta de utilidad derivada de un conjunto de preferencias bien definido el consumidor estará en una situación peor si $v(p^1,y) - v(p^0,y) < 0$.

La función de utilidad derivada de $\succsim$ es suficiente para realizar alguna comparación. Sin embargo existe una función de utilidad indirecta que lleva a una medida del cambio de la riqueza en unidades monetarias (pesos) que se puede denominar como *utilidad indirecta métrica monetaria* y que se construye a través de la función de gasto. En particular, se parte de la función de utilidad indirecta $v(\cdot, \cdot)$, eligiendo un vector de precios arbitrarios $\bar{p}$ estrictamente positivo y, a partir de la función $e(\bar{p}, v(p, y))$, se puede obtener la riqueza necesaria para alcanzar el nivel de utilidad $v(p, y)$ cuando los precios son $\bar{p}$. Observe también, que la función de gasto es estrictamente creciente, ya que depende del nivel de $v(p, y)$. Así, una medida del cambio de la riqueza expresada en pesos vendrá determinada por

$$e(\bar{p}, v(p^1, y)) - e(\bar{p}, v(p^0, y)) \tag{3.10}$$

Gráficamente lo podemos ver como

Gráfica 3.2. Función de utilidad métrica monetaria.



De esta forma, la utilidad indirecta métrica monetaria puede ser construida para algún vector de precios. Estas elecciones llevan a dos medidas en torno al cambio de la riqueza, la primera conocida como la variación equivalente (EV) y, la segunda como la variación compensatoria(CV).

Formalmente sea $u^0=v(p^0,y)$ y $u^1=v(p^1,y)$. De igual forma, haciendo $e(p^0,u^0)= e(p^1,u^1)= y$

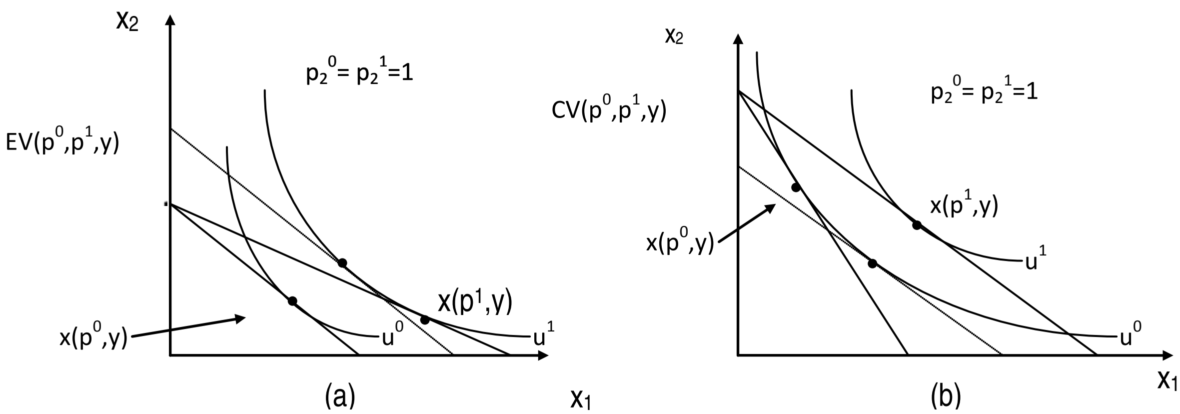
(3.11) $EV (p^0, p^1,y) = e(p^0,u^1)- e(p^0,u^0)= e(p^0,u^1)- y$
(3.12) $CV (p^0, p^1,y) = e(p^1,u^1)- e(p^1,u^0)= y - e(p^1,u^o)$

La variación equivalente, será la cantidad de pesos ante la cual el consumidor es indiferente, en lugar de aceptar un cambio en precios. Esto es, el cambio en su riqueza que es equivalente al cambio en precios en términos del impacto de riqueza (éste es negativo si el cambio en precios hace que el consumidor se encuentre peor). Deberá observarse, que $e(p^0,u^1)$ es el nivel de riqueza al cual el consumidor alcanza exactamente el nivel de utilidad u^1, es decir el nivel generado por el cambio en precios desde p^0. Además $e(p^0,u^1) - y$ es el cambio neto en la riqueza de tal forma que el consumidor alcanza la utilidad u^1 a precios p^0.

La variación compensatoria, medirá el ingreso neto que debe compensar al consumidor por el cambio en precios, una vez éste ha ocurrido, de tal forma que el consumidor recobre su nivel original de utilidad u^0. Como menciona Mascollel et-al " ésta puede ser pensada como la cantidad negativa que el consumidor justamente estaría dispuesto a aceptar del planeador que ha asignado el nuevo cambio de precios"(pag 82).

Gráficamente, estas medidas se pueden ver como

Gráfica 3.3. Medidas del cambio de riqueza: (a) Variación equivalente (b) Variación compensatoria.



Las variaciones equivalentes y compensatorias podrían diferir, en tanto, el vector de precios, al cual, la compensación se asume, difiera. Esto significa, como observa Azqueta(1994) que en el caso de una caída en los precios CV<EV y ante una elevación en el precio del bien CV > EV.

El cambio en la riqueza producido por un cambio en el precio del bien 1 puede ser medido por la curva de demanda marshalliana, es decir, si definimos la medida de variación como

(3.13) $$MV(p^0, p^1, y) = \int_{p_1^1}^{p_1^0} x_1(p_1, \bar{p}_2, \ldots, \bar{p}_n, y)\partial p_1$$

Y, si no existen efectos de riqueza en el bien 1, entonces ésta medida de variación es exactamente igual a las medidas equivalentes y compensatorias. Esta medida, será el cambio en el excedente del consumidor marshalliano. En tanto, se consideren las variaciones en la riqueza como efecto de variaciones en los precios, el excedente del consumidor se encuentra entre la variación equivalente y compensatoria. El error cometido de usar esta medida es del 2% como observa Willing(1976). Cuando las variaciones en la riqueza son debidas a cambios en las cantidades, Hanemman(1991) ha demostrado que estas dos medidas difieren ampliamente pues ya no sólo se deberá tener en cuenta el efecto renta, sino también efectos de sustitución.

3.2.1 La disponibilidad a pagar

Suponga que un consumidor tiene la oportunidad de comprar una cantidad x de un bien. Nosotros deseamos determinar cuánto de ésta oportunidad corresponde al “esfuerzo”, medido en unidades de gasto sobre otros ítems. Para determinar este valor, deberemos usar la gráfica (3.3), pero ahora, la curva muestra la curva inversa de la demanda del consumidor compensada. Esta curva, resulta de fijar la utilidad u^0 al valor original (x=0) y se usa su forma inversa. Imagine entonces una curva que se devuelve una unidad, comprando cada unidad al precio indicado. El valor de las unidades sucesivas son las áreas de las líneas verticales divididas bajo la curva. El consumidor estará dispuesto a pagar más por cada unidad y la utilidad permanecerá constante durante este proceso. De esta forma, la cantidad total que se estaría dispuesto a pagar será

(3.14) $DP(x)=\int_0^x p_c(\xi, u^0)\partial\xi$

Donde $p_c(\xi,u^0)$ es la demanda inversa compensada: el precio ajustado cuando los otros precios están fijos. De aquí se obtiene

(3.15) $p_c(x,u^0) = y'(x)$

Esto es, si el consumidor compra x unidades al precio p, el área bajo la curva de demanda compensada antes del precio p es la disponibilidad a pagar neta. En general, esta medida es diferente al excedente del consumidor, solo que si no existen efectos ingreso $\partial x(p,y)/\partial y=0$ las dos curvas de demandas serán iguales y la disponibilidad neta por pagar será igual al excedente del consumidor.

3.2.2 La compensación exigida

La compensación exigida, CE, refleja lo que se demandaría con el fin de aceptar un cambio que empeore su situación, o renunciar a un cambio que mejore su situación. Cuando el precio cae la CE es equivalente a la EV y cuando el precio aumenta la CE es equivalente a la CV. Por otro lado la CE no está limitada por la renta, por lo cual, su principal efecto será en términos de sustitución.

3.2.3 Comparación entre la disponibilidad a pagar y la compensación exigida

Aunque ambas medidas teóricamente representan los mismos resultados, los estudios de Hahneman (1991), Kahnemann,Knetsh y Thaler(1990) han mostrado que dichas medidas difieren. Por un lado, Hahneman muestra que existen diferencias cuando el cambio en la renta es debido a un cambio en las cantidades, sobre todo, en la provisión de bienes públicos. Por otro lado, existen asimetrías entre lo que un individuo está dispuesto a aceptar y, entre lo que un individuo estaría dispuesto a renunciar Kahnemann,Knetsh y Thaler(1990). En últimas, si existe un punto de referencia entre ambas medidas, las propiedades de la función de utilidad subyacente hacia dichas medidas diferirán en convexidad y dirección, esto significaría dependencia con respecto al punto de referencia(pérdidas y ganancias), aversión al riesgo(la pendiente de la función del ortante positivo tiene un valor mayor con relación al ortante negativo, lo que significa que las pérdidas tienen un valor superior que las ganancias) y, por último el valor marginal de las pérdidas y las ganancias disminuye con su tamaño.[Tversky y Kahneman(1991,pag1039)].

3.3 Integrabilidad de la función de utilidad

A partir de la función de demanda x(p,y) se puede recobrar la función de utilidad subyacente, en lo que se conoce como integrabilidad.

Suponga, que nosotros tenemos un función de demanda continua y diferenciable x(p,y). Si esta función se encuentra bien definida y, se cumple el supuesto de insaciabilidad local (y la ley de Walras) asegurando la igualdad, entonces como previamente se ha demostrado, deberán cumplirse las siguientes condiciones

1.No negatividad: $x(p,y) \geq 0 \ \forall$ p e y
2.Homogeneidad: $x(tp,ty)= x(p,y) \ \forall \ t > 0$
3.Insaciabilidad : $px(p,y)= y$
4.Simetría : La matrix de Slutzky es simétrica.
5.Semidefinida : S es semidefinida negativa

Usando el Lema de Shepard como una unión entre la demanda compensada y, la función de gasto, encontramos

(3.16) $$\frac{\partial e(p,u)}{\partial p_i} = x_i(p, e(p,u)) \ \forall i = 1,..,m$$

Donde (3.16) es un sistema de ecuaciones diferenciales parciales, y u se introduce como un parámetro. Especificando también las condiciones de contorno de la forma $e(p^*,u)= c$ donde p^* y c están dados, entonces se podrá recuperar la función de utilidad de e.

La solución que resulta es única y depende continuamente de c. Una vez que la función de gasto sea encontrada, la función de utilidad indirecta se puede hallar teniendo en cuenta

(3.17) $e(p,y) = y$

Donde e es estrictamente creciente, y puede ser invertida para encontrar v(p,y). Considere las siguientes funciones de demanda que provienen de una función de utilidad Cobb-Douglas

(3.18) $x_i(p,y) = \dfrac{\alpha_i y}{\alpha p_i}\ ;\ \forall\ i = 1,2,\ldots,m$

Siendo $\alpha = \sum_{i=1}^{m} \alpha_i$. De acuerdo a (3.16) se tiene

(3.19) $\dfrac{\partial e(p,u)}{\partial p_i} = \dfrac{\alpha_i e(p,u)}{\alpha p_i}\ ;\ \forall\ i = 1,2,..,m$

La i- ésima ecuación puede ser integrada con respecto a p_i obteniendo

(3.20) $\ln e(p,u) = \dfrac{\alpha_i}{\alpha} \ln p_i + c_i$

Donde c_i no depende de p_i(pero podría depender de p_j para algún $j \neq i$). Agregando se obtiene

(3.21) $\ln e(p,u) = \sum_{i=1}^{m} \dfrac{\alpha_i}{\alpha} \ln p_i + c$

Ahora c es independiente de todos los p_i´s. La constante c, representa la libertad en las condiciones de contorno. Para cada u se deberá hacer $p^*=(1,1,\ldots,1)$ y, usar la condición de contorno $e(p^*,u)=u$, por lo cual

(3.22) $\ln e(p,u) = \sum_{i=1}^{m} \dfrac{\alpha_i}{\alpha} \ln p_i + \ln u$

Entonces (3.22) es fácilmente invertible, obteniendo

(3.23) $\ln v(p,y) = -\sum_{i=1}^{m} \dfrac{\alpha_i}{\alpha} \ln p_i + \ln y$

La cual es una transformación monótona de la función de indirecta de utilidad, para una función de utilidad Cobb-Douglas.

3.4 Preferencias reveladas

Los axiomas básicos sobre las preferencias, son criticados por ser demasiado fuertes, ya sea en su preordenamiento o en su ordenamiento completo. Sin embargo, a menudo observamos cómo los individuos realizan elecciones, aunque las restricciones sobre el conjunto de preferencias no sean observables. Una forma de hacer compatibles los supuestos sobre las preferencias y las decisiones que observamos en el mercado, consiste, en lo que comúnmente se denomina como preferencias reveladas.

La "eficacia" de la teoría de la preferencia revelada, radica, en que el estado de las preferencias se construye a partir de las decisiones observables, esto es, de las elecciones actuales realizadas por un consumidor determinado.

Dado un vector de precios $p \geq 0$, un nivel de ingreso $y > 0$ y una canasta de bienes $x \in X$ se puede definir una restricción presupuestaria de la forma $px \leq y$. Asumiendo que la existencia de la restricción presupuestaria está relacionada con la elección de una casta posible x para el consumidor, donde x depende de p e y de la forma $x(p , y)$ cuando se realiza la elección[9] .

3.4.1 *Preferencia revelada directamente*

Suponga la existencia de un vector de precios p y un nivel de ingreso (y), de tal forma que $x \in x(p,y)$ y $px' \leq y$ para algún $x' \in X$. Entonces $u(x) \geq u(x')$. Se dice, que x se revela preferido directamente a x´, y se escribe esta relación como xR_px'.

La relación R_p es una clase parcial de relación de preferencias sobre x y puede ser usada para construir como se realizan las elecciones. Sin embargo, R_p tiene muy pocas propiedades: no se puede cumplir xR_px si x es la canasta elegida, lo que significa que no siempre se cumple la reflexividad, además xR_px no se cumple si x no es elegida. R_p no es completa y tampoco es transitiva ya que si se tiene xR_px' y $x'R_px''$ de aquí no se deriva que xR_px''.

3.4.2 *Preferencia revelada*

Una cesta x se dice que es revelada preferida a x´, $R^c{}_p$, si no se revela preferido x´a x , esto es, si existe un número finito de cestas $x_1,x_2,..,x_m$ en X tales que $xR^c{}_px_1$, $x_1R^c{}_px_2$, $x_2R^c{}_px_3,\ldots, x_mR^c{}_px'$, se dice que $x\ R^c{}_px'$.

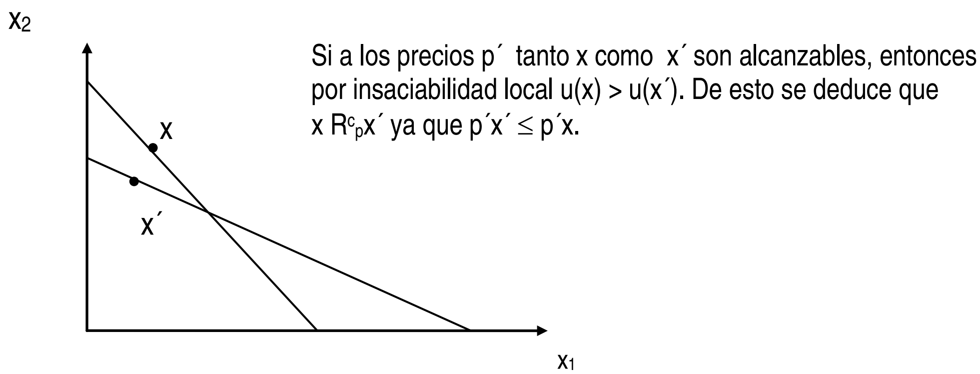
La relación anterior se ha construido teniendo en cuenta la existencia de $R^c{}_p$ relaciones y se denomina la clausura transitiva, esto es, nosotros tenemos que $x\ R^c{}_px'$ si x es elegido sobre x_1, x_1 sobre x_2, x_2,.., y finalmente x_m sobre x´ para algún punto intermedio. La relación de comparación no necesariamente deberá ser directa, pero la relación de hecho es transitiva: Si además de tener que $x\ R^c{}_px'$ tenemos que $x'R^c{}_px''$ entonces $x'R^c{}_px'_1$, $x'_1R^c{}_px'_2,\ldots, x'_nR^c{}_px''$ para algún x'_i´s, además si tenemos que $xR^c{}_px_1$, $x_1R^c{}_px_2$, $x_2\ R^c{}_px_3,\ldots, x_mR^c{}_px'$, $x'R^c{}_px'_1,\ldots, x'_nR^c{}_px''$ lo cual nos muestra que $x'R^c{}_px''$ cumple la transitividad.

---

[9] Esto implica que x(p,y) deberá ser homogénea de grado cero con respecto a p e y.

Adicionalmente, deberá observarse que si x $R^c{}_p x'$ entonces $p'x' \leq p'x$, lo cual se puede observar gráficamente como

Gráfica 3.4. Preferencias reveladas



3.4.2.1 El axioma débil de la preferencia revelada

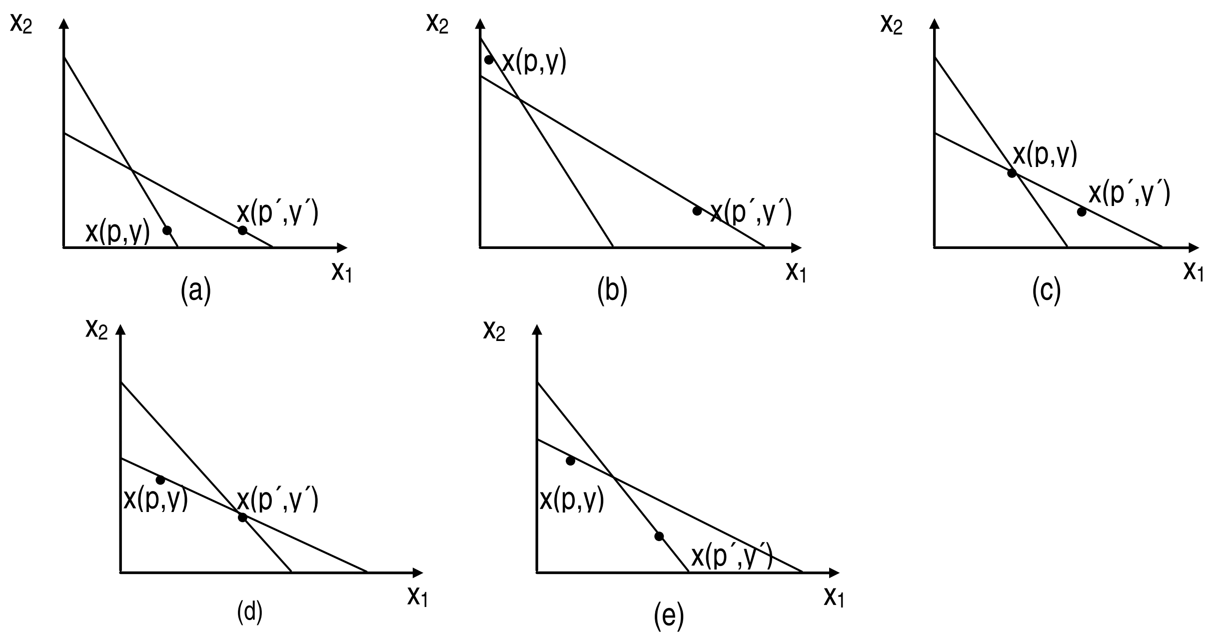
Si $xR^c{}_p x'$ y x no es igual a x´, entonces no es cierto que $x'R^c{}_p x$. Esto implica, que si $px(p',y') \leq y$, y sí $x(p',y') \neq x(p,y)$ entonces necesariamente $p'x(p,y) > y'$. De esta forma, dados los precios e ingreso, la canasta x(p,y) es elegida incluso cuando x(p´,y´) es también alcanzable. De aquí se deduce que la canasta x(p,y) no es alcanzable a la combinación de precios y riqueza (p´,y´) en la cual el consumidor ha elegido x(p´,y´) por lo cual, deberá cumplirse que $p'x(p,y) > y'$.

Para algún cambio compensado en precios de la situación inicial (p,y), a una nueva situación de riqueza $(p',y')=(p',p'x(p,y))$, se deberá cumplir

(3.24) $(p'-p)[x(p',y')-x(p,y)] \leq 0$

Con estricta desigualdad siempre que $x(p,y) \neq x(p',y')$. Como demuestra Mas-collel- et al(1995) la violación del débil axioma, implica la violación del cambio compensado en precios. A continuación, veamos gráficamente cuando se satisface el débil axioma y cuando lo viola

Gráfica 3.5. (a), (b) y (c) satisfacen el débil axioma y (d) y (e) no lo satisfacen.



3.4.2.2 El axioma fuerte de la preferencia revelada

Si $xR^c{}_px'$ y x no es igual a x´, entonces no es cierto que $x'R^c{}_px$. Esto implica que dada una lista de precios e ingreso (p,y),(p´,y´),…,(p^m,y^m) con $x(p^{m+1},y^{m+1}) \neq x(p^m,y^m)$ $\forall m \leq M-1$. Entonces p^M x(p,y)>y^M cuando p^M $x(p^{m+1},y^{m+1}) \leq y^M$ $\forall m \leq M-1$. En otras palabras, si x(p,y) es directamente(o indirectamente) revelado preferido a x(p^m,y^m), entonces x(p^m,y^m) no puede ser directa(o indirectamente) revelado preferido a x(p,y), esto es x(p,y) no puede ser alcanzado a (p^m,y^m). De lo anterior, se deriva que si una función de demanda Walrasiana x(p,y) satisface el axioma fuerte de la preferencia revelada, entonces existe una relación de preferencia racional($\succsim$) que racionaliza x(p,y), esto es, para todo (p,y), x(p,y)>x´ para cada x´ ≠x(p,y) donde x´ pertenece a la restricción presupuestaria en (p,y) y, que ($\succsim$) racionalice x(p,y), significa que la conducta observada alcanza su valor máximo en el conjunto presupuestario de las cestas elegidas.

3.4.3 Condición suficiente para maximizar la utilidad

Si las elecciones (p,x) fueron generadas por un consumidor que maximiza su utilidad y sus preferencias cumplen el supuesto de insaciabilidad, entonces, estas elecciones satisfacen las preferencias reveladas directamente. Formalmente, se requiere el cumplimiento del **teorema de Afriat.** Supongamos que (p^t,x^t), $\forall$t=1,…,T., sea un número finito de observaciones de vectores de precios y cestas de consumo, las siguientes condiciones son equivalentes

1- Existe una función de utilidad que cumple la insaciabilidad local y que racionaliza las elecciones.
2- Las elecciones (p,x) satisfacen $R^c{}_p$.
3- Existe una serie de números positivos $(u^t,\lambda^t), \forall t=1,\ldots,T$ que satisfacen las desigualdades de **Afriat,** esto implica que $u^1 \leq u^t + \lambda^t (x' - x^t)$ cualesquiera que sean (t) y (´).
4- Existe una función de utilidad monótona, cóncava, continua y no saciada que racionaliza las elecciones.

La existencia de una función de utilidad que racionalice las elecciones (p,x) implica que existe una función de utilidad monótona, cóncava, continua y no saciada que racionaliza las elecciones (p,x) y, si la función de utilidad, no estuviera definida con las propiedades anteriores, nunca se observaría que se tomara alguna decisión en dichas cestas, esto significa, que las elecciones extraídas del mercado no permiten rechazar la hipótesis de convexidad y monotonicidad en las preferencias.

3.5 Agregación

Un punto de singular importancia, consiste en la agregación sobre los individuos, ya que el comportamiento agregado de los consumidores, en muchas situaciones, es más importante que el comportamiento de un consumidor en particular. Y desde, un punto de vista econométrico, deberán existir restricciones en la agregación cuando se estimen las funciones de demanda.

En torno a la demanda agregada, se deberá discutir en primer lugar si esta puede ser expresada como una función de los precios y de la riqueza agregada. En segundo lugar, si las restricciones individuales sobre las preferencias se sostienen en el agregado y en tercer lugar, como se medirían dichos cambios agregados.

Una agregación perfecta en un modelo de un período, depende de que todos los precios sean los mismos para todos los individuos. Así, las variaciones provienen por parte de la riqueza que cada individuo posee. Por otro lado, en modelos de elecciones intertemporales, no sólo existen diferencias en el ingreso, también existen diferencias en la edad, y en las expectativas acerca de los precios futuros.

3.5.1 Agregación lineal

Supongamos, la existencia de I consumidores con preferencias racionales $\succsim$ y, funciones Walrasianas de demanda $x_i(p,W_i)$. Dados unos precios y unos niveles de riqueza $(W_1,\ldots,W_I)$ para I consumidores y m bienes, la demanda agregada podría escribirse como

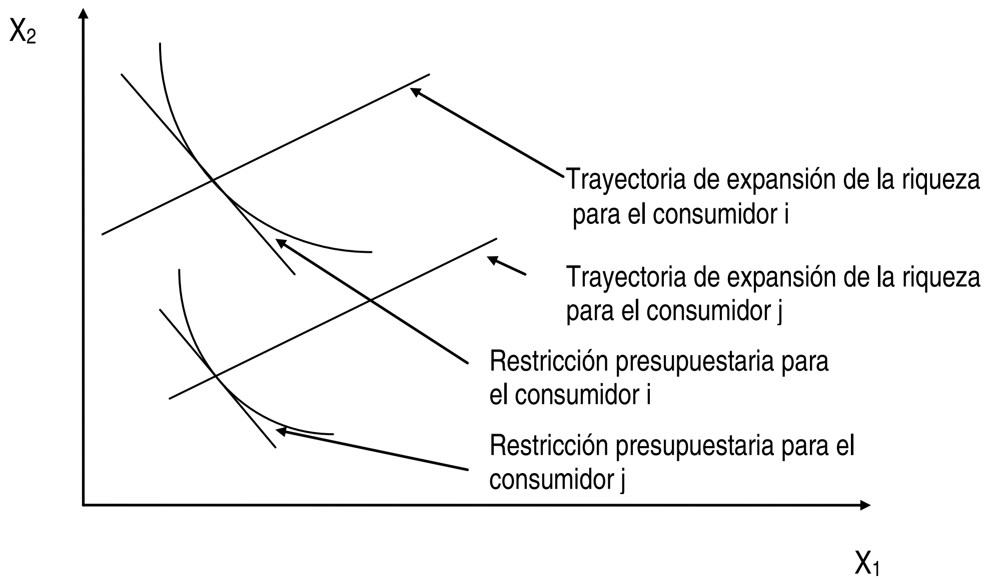
(3.26) $$x(p, W_1, \ldots . W_I) = \sum_{i=1}^{I} x_i(p, W_i)$$

Se puede observar, que la demanda agregada no depende solamente de los precios sino también de los niveles de ingreso de los consumidores. La función (3.26) se mantiene, en tanto la agregación sea igual en todas las posibles distribuciones de la riqueza de los consumidores. Suponga que existen $(W_1,\ldots,W_I)$ y $(W'_1,\ldots,W'_I)$ y $\sum_i W_i = \sum_i W_i'$, entonces

$$(3.27) \quad \sum_i x_i(p,W_i) = \sum_i x_i(p,W'_i)$$

La condición (3.27) implica invarianza en la demanda agregada ante cambios en la redistribución de la riqueza, gráficamente esto se puede observar como

Gráfica 3.6. Invarianza en la función de utilidad



La condición (3.27) muestra que para todos los consumidores las trayectorias de expansión de la riqueza son paralelas, como se observa en la gráfica (3.6).

Una condición necesaria y suficiente para que el conjunto de consumidores tenga trayectorias de expansión de riqueza paralelas consistirá en que las preferencias admitan funciones de utilidad indirectas tipo GORMAN:

$$(3.28) \quad V_i(p,W_i) = a_i(p) + b(p)W_i$$

En (3.28) b(p) deberá ser igual para todos los agentes y a_i podría diferir entre todos los consumidores[Luenberger(1995),Mas-Collel et-al(1995)]. Deaton (1980) considera que a_i puede interpretarse como el gasto de subsistencia, que debería ser igual para todos los agentes que pertenecen a una comunidad. A través de la identidad de Roy, se puede demostrar que las demandas vienen definidas como

$$(3.29) \quad X_i(p,W_i) = A_i(p) + B(p)W_i$$

Y la demanda agregada, sobre todos los consumidores viene definida por

(3.30) $X(p, W_1 , W_2 , W_3 ,.....,W_I) = \sum_{i=1}^{I} \{ A_i(p) + B(p)W_i\}$

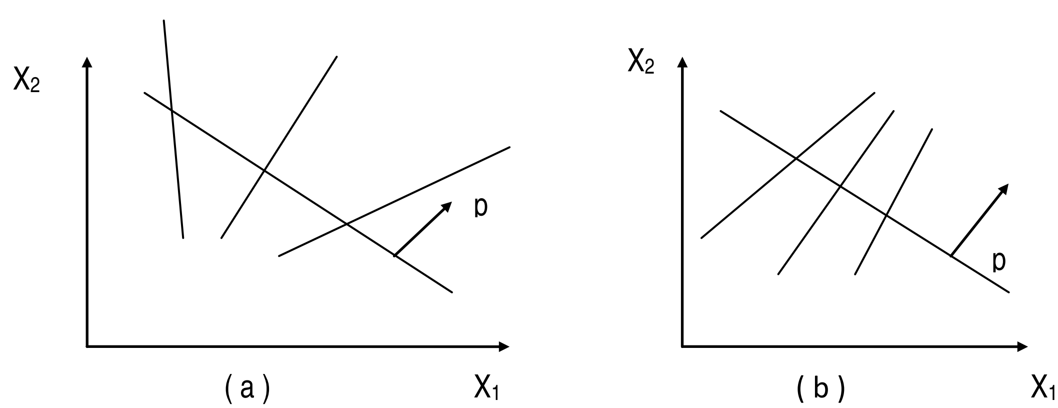
Lo cual muestra, que la demanda agregada depende de los niveles de ingreso solamente con relación a la riqueza total. Por otro lado, la demanda de un bien deberá ser cero, siempre que W_i sea cero. En el caso básico a_i debería ser cero, sin embargo todo depende del bien en cuestión. En cumplimiento de la homoteticidad, las curvas de Engel deberán provenir de una elasticidad de gasto unitaria, y las funciones de utilidad idénticas y homogéneas de grado 1. La condición anterior, no debe ser tomada a la ligera, pues implica que cualquier transformación monótona creciente deberá mantener la función de utilidad. Finalmente, se deberá exigir que la demanda agregada satisfaga el débil y el fuerte axioma de la preferencia revelada[Mas-Collel, et-al(1995)] y que sea dos veces continuamente diferenciable.

De esta forma, suponiendo que todos los consumidores tienen preferencias $\succsim$, funciones de demanda individuales $\bar{x}$ (p,W) y que la riqueza individual se distribuye uniformemente en el intervalo [0, $\bar{W}$], entonces existe un continuo de consumidores, cuya función de demanda agregada (el promedio) vendrá determinada por

(3.31) $X(p) = \int_0^{\bar{W}} \bar{x}(p, W) \partial W$

Suponiendo, que existe una distribución uniforme de la riqueza a través de los consumidores, en el caso de dos bienes, una relación positiva mostraría que aquellos consumidores con un consumo mayor que el promedio de un bien, gastan una fracción mayor que el promedio de la última unidad de riqueza sobre el bien

Gráfica 3.7. Trayectorias de expansión de la riqueza (a) Relación positiva (b) Relación negativa



Las ecuaciones (3.28) a (3.31) implican que se puede agregar una serie de individuos cuando ellos tienen unas preferencias consistentes con su comportamiento. Sin embargo, para agregar perfectamente, se deberá imponer como condición que el intercepto, que está reflejando las características del hogar como la edad, el sexo, la raza, la educación del padre, el número de hijos, etc.[Ver además Pollak(1970)], esté relacionado a través de los individuos con variables como el gasto total y las cantidades demandadas.

3.5.2 Agregación no lineal

Una agregación lineal exacta, requiere que la demanda promedio del mercado este en función de gasto total promedio [Deaton y Muelbauer (1980:154),Deaton(1989:53)]. La manipulación de esta condición, da como restricción curvas de Engel lineales y éstas tienen la misma pendiente para cada individuo. Supongamos que la riqueza agregada, $\bar{W}$, para el i- ésimo bien sea

$$(3.32)\quad \bar{W}_i = \frac{p_i \sum_h x^h{}_i}{\sum_h x^h} = \sum_h \frac{x^h}{\sum x^h} w^h{}_i$$

De tal forma, que el patrón de demanda es un patrón promedio de los patrones del hogar, y estos promedios son proporcionales al gasto de cada hogar. Si $\bar{W}_i$ está en función de los precios y de los gastos totales de cada hogar, una aproximación a la agregación consiste en restringir $\bar{W}_i$ de tal forma que dependa sobre los precios y que el nivel de los gastos x_0 este en función de la distribución de los gastos. Si esto se mantiene, la demanda agregada de mercado puede ser derivada del comportamiento de un individuo representativo, con un gasto total x_0 a unos precios p.

Formalmente, un consumidor representativo existe si nosotros podemos definir una función de utilidad $\mu(x, p)$ asociada a una función de gasto $g(\mu, p)$, de tal forma que para algún $\mu_0=\mu(x_0 , p)$ se cumpla

$$(3.33)\quad \bar{W}_i = W_i(\mu_0, p) = \frac{\partial \log g(\mu_0, p)}{\partial Log p_i} = \sum_h \frac{x^h}{\sum x^h} \frac{\partial \log g^h(\mu^h, p)}{\partial Log p_i}$$

Siendo $g^h(\mu^h, p)$ la función de gasto del hogar h, $\mu^h=\mu(x^h,p)$. Entonces, la ecuación (3.33) está definida como una agregación no lineal modificada a través de $\mu^h=\mu_0$. El conjunto de ecuaciones diferenciales parciales (3.33) pueden integrarse para encontrar la función de gasto, esto es, para el hogar h la función de gasto toma la forma

$$(3.33)\quad g^h(\mu^h, p) = \theta^h[\mu^h, a(p), b(p)] + \Omega^h(p)$$

Con a(p), b(p) y $\Omega^h(p)$ funciones homogéneas lineales de los precios y θ^h una función homogénea lineal en a(p) y b(p). Por otro lado, θ^h, a(p) y b(p) serán funciones crecientes en sus argumentos y θ^h cóncava en a(p), b(p) y p. Agregando en todos los consumidores Ω^h(p) deberá ser cero, por lo cual la función de gasto podría expresarse como

(3.35) $g(\mu_0,p)= \theta[\mu_0, a(p),b(p)]$

Por otro lado, (3.35) podría reescribirse como

$$\bar{W}_i = W_i(\mu_0, p) = \frac{\partial \log\theta}{\partial Loga}\frac{\partial \log a}{\partial Logp_i} + \frac{\partial \log\theta}{\partial Logb}\frac{\partial \log b}{\partial Logp_i} \qquad (3.36)$$

Dado que θ es homogénea de grado uno en a(p) y b(p), entonces $\frac{\partial \log\theta}{\partial Loga} = 1 - \frac{\partial \log\theta}{\partial Logb}$. De esta forma, (3.36) se puede escribir como

$$\bar{W}_i = (1-\lambda)\frac{\partial \log a}{\partial Logp_i} + \lambda\frac{\partial \log b}{\partial Logp_i} \qquad (3.37)$$

Con $\lambda = \frac{\partial \log\theta}{\partial Logb} = \lambda(x_0, p)$

Si los precios son constantes, entonces $\bar{W}_i$ puede ser estimada directamente. La restricción lineal impuesta a (3.35), se conoce también como Linealidad Generalizada(GL) y, para un consumidor representativo x_0 podría ser algún punto en la función de distribución del gasto, dicha posición esta determinada por el grado de no-linealidad en la curva de Engel y el vector de precios p.

Un caso especial, ocurre cuando el nivel de gasto para un consumidor representativo es independiente de los precios y depende solamente de la distribución de los gastos, este caso se conoce como Linealidad Generalizada Independiente de los Precios y ocurre cuando las funciones de gasto toman la forma

(3.38) $g^h(\mu^h , p) = k^h[a(p)^\alpha (1 - \mu^h)+b(p)^\alpha \mu^h]^{1/\alpha}$

La ecuación (3.38) representa la función de gasto, con una la utilidad promedio de orden α entre los dos índices. Cuando α tiende a 0 y tomando logaritmos nosotros tendremos

(3.39) $\text{Log } g(\mu^h , p) = (1 - \mu^h)\text{Log } a(p)+ \mu^h \text{ Log } b(p)$

Esta función logarítmica, es también conocida como PIGLOG. Como se puede observar, el parámetro α es importante en determinar la no-linealidad de las curvas de Engel, así cuando α es igual a 1, la función de costo y las curvas de Engel son lineales. Si α = –1 entonces, las curvas de Engel son cuadráticas.

4 Separabilidad

Cuando usted va a comprar alimentos, es natural que destine una parte de su ingreso en la compra de éstos. Al igual, como destina parte de su ingreso en comprar alimentos, destinará dinero para alquiler, servicios públicos, ropa, entretenimiento, etc. Observe que esto implica que en cada ítem se agrupe una serie de bienes (por ejemplo alquiler-apto; garaje, etc, ropa: camisas, zapatos, etc y así sucesivamente). Agrupar los bienes, requiere que las preferencias reflejen este agrupamiento. Hace algunos años Sono (1962), Goldman y Usawa(1964) y Pudney(1981), entre otros, observaron que es posible separar las decisiones de los individuos de asignar sus ingresos a una serie de bienes, manteniendo la estructura de preferencias, de las decisiones intertemporales de gastar.

De esta forma, el objetivo de este capítulo, consistirá en mostrar cómo manteniendo la estructura de las preferencias es posible separar las decisiones de gastar en cada grupo.

4.1 Estructura de las preferencias

Supongamos que los bienes son particionados en dos subgrupos, con un vector **x = (y,z)**, esto es, $X = Y \times Z$. Para un **z** fijo se define un orden condicional $\succsim_z$ sobre Y tal que la relación **y** $\succsim_z$ **y´** se mantiene si y solo si **(y,z)** $\succsim$ **(y´,z)**. De esta forma, $\succsim_z$, es una restricción sobre el orden original definiendo un **z** fijo. Deberá observar, que para algún **z** la relación $\succsim_z$ es de hecho un orden de preferencia sobre Y. Para una partición **x = (y,z)** si el orden de preferencias condicionado sobre Y es independiente de **z**, nosotros diremos que **y** es independiente de **z**.

4.1.1. Independencia

Suponga un orden de preferencia representado por una función de utilidad μ(**y,z**). Entonces, si **y** es independiente de **z** la función de utilidad (μ) puede ser escribirse como

(4.1) $\mu(\mathbf{y},\mathbf{z}) = U(\nu(\mathbf{y}),\mathbf{z})$

Donde $U(\nu,\mathbf{z})$ es estrictamente creciente en ν. Si μ es continua y fuertemente monótona, entonces ν y μ son continuas.

4.1.2. Débil y fuerte independencia

Suponga la existencia de n bienes particionados en g grupos. Una partición, se define como $\{n_1,n_2,\ldots,n_g\}$ con $n_i \cap n_j$ vacío para todo $i \neq j$ y $\bigcup_{i=1}^{g} n_i = N$. Para una canasta de bienes arbitraria $\mathbf{x} \in \Re_+^m$ particionada como $\mathbf{x} = (\mathbf{x}_1,\mathbf{x}_2,\ldots,\mathbf{x}_g)$ dado algún i=1,2,…,g .

Sea $\mathbf{x}_{-i}$ el vector de bienes en el complemento de n_i, de tal forma que $\mathbf{x}_{-i}=(\mathbf{x}_1,\mathbf{x}_2,\mathbf{x}_3,.....,\mathbf{x}_{i-1},\mathbf{x}_{i+1},...,\mathbf{x}_g)$, entonces

A) Un orden de preferencias es débilmente independiente, con respecto a una partición $\{n_1,n_2,....,n_g\}$, si para cada i=1,2,..,g el vector $\mathbf{x}_i$ es independiente de su complemento.

B) Un orden de preferencias es fuertemente independiente, con respecto una partición $\{n_1,n_2,....,n_g\}$, si este es débilmente independiente con respecto a la partición $\{n_1,n_2,....,n_g\}$ y, con respecto a las particiones que consisten de todas las uniones de $n_1,....,n_g$ y a los subconjuntos propios de n.

Dado que ν es continua y creciente, $\nu^{-1}u$ es aditiva y representará el mismo orden de preferencias[10]. En general, no se requieren más supuestos excepto cuando $x_1,...x_n$ son consumos en el *tiempo* 1,...,t. En tal caso, se debe usar el principio de Stroz, esto es, la consistencia dinámica de las preferencias.

4.1.1 Separabilidad débil y fuerte

Sea N= $\{n_j\}_{j=1}^{k}$ una partición del conjunto del conjunto { 1,....,g} y, asuma el conjunto de consumo X = $S_1\times\cdots\times S_k$. Tal partición, es apenas natural si el consumo es considerado en varios lugares. La separabilidad implica que las preferencias sobre las canastas en cada elemento de la partición, en la fecha y en el lugar será independiente del nivel de consumo[Barten, A y Böhn, V(1982)].

4.1.2 Separabilidad débil

Una función de utilidad $\mu:\prod_{j\in J} S_j \to \Re$ es *débilmente separable*, si existe una función continua $\mu_j : S_j \to \Re$, j $\in$ J y $\nu : \Re^k \to \Re$ tal que $\mu(x)=\nu(\mu_1(x_1),\ldots,\mu_1(x_1))$.

4.1.3 Separabilidad fuerte

Una función de utilidad $\mu:\prod_{j\in J} S_j \to \Re$ es *fuertemente separable*, si existe una función continua $\mu_j : S_j \to \Re$, j $\in$ J y $\nu : \Re \to \Re$ tal que $\mu(x)=\nu(\sum_{j\in J}\mu_j(x_j))$.

[10] Ver Barten, A y Böhn, V. (1982).

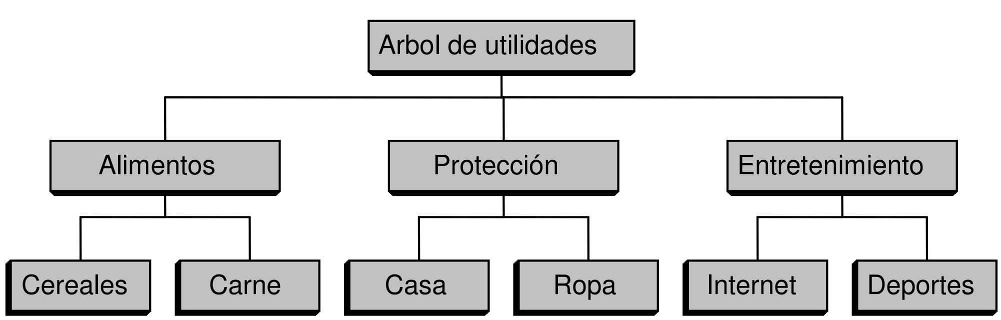
4.2 Separabilidad de las preferencias

Para definir grupos de bienes o estructuras de bienes, deberemos partir de la definición de separabilidad en torno a las preferencias. Si esto es plausible, los bienes pueden ser particionados en grupos donde las cantidades en un grupo son independientes de las cantidades en otros grupos. Si los alimentos pertenecen a un grupo, el consumidor puede ordenar diferentes canastas de alimentos en un orden bien definido, el cual, es independiente del consumo en gasolina, entretenimiento, arrendamientos, y cualquier bien por fuera del grupo. Esto significa que nosotros tendríamos funciones de subutilidades para cada grupo y que los valores de cada subgrupo de utilidades se combinan, de tal forma, que se puede obtener una utilidad total.

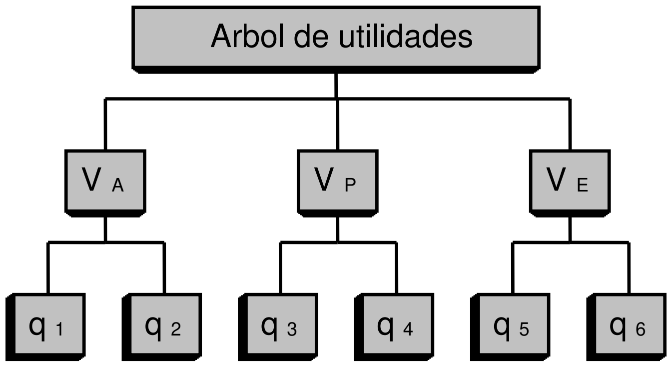
Para una definición más formal, considere J = {1,…,k} y para algún $j \in J$ y $x \in X$, sea $\mathbf{x}_{-i} = (\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \ldots, \mathbf{x}_{i-1}, \mathbf{x}_{i+1}, \ldots, \mathbf{x}_g)$ el vector de componentes diferentes de x_j. Para algún $\bar{x}_{-i}$ fijo, el orden de preferencias $\succsim$, induce un orden de preferencias sobre S_j, el cual está definido por $x_j \succsim_{\bar{x}_{-i}} x'_j$ sí y solo sí $(\bar{x}_{-i}, x_j) \succsim (\bar{x}_{-i}, x'_j)$ para algún x_j y x'_j en S_j. La primera noción de separabilidad, nos indica que los órdenes de preferencias para un elemento particular j de la partición, son idénticos para todos los $\bar{x}_{-i}$.

Débil separabilidad de las preferencias: Un orden de preferencias $\succsim$ sobre $\prod_{j \in J} S_j$ es llamado débilmente separable, si para cada $j \in J$, $x_j \succsim_{\bar{x}_{-i}} x'_j$ implica $x_j \succsim_{\bar{x}_{-i}} x'_j$ para todo $\bar{x}_{-i} \in \prod_{j \neq J} S_j$. Supongamos el siguiente esquema

Gráfica 4.1. Separabilidad de la función de utilidad



Gráfica 4.2. Presupuesto en dos etapas



A partir de la gráfica (4.2), se puede plantear la función de utilidad como

(4.2) $U = \nu(q_1, q_2, q_3, q_4, q_5, q_6) = F\,[\, \nu_A\,(q_1, q_2), \nu_P\,(q_3, q_4), \nu_E\,(q_5, q_6)\,]$

Donde, F(.) es una función creciente y ν_A, ν_P, ν_E son funciones de subutilidades asociadas con alimentos, protección y entretenimiento.

Este diagrama de utilidad nos muestra un presupuesto en dos estados, esto ocurre cuando el consumidor puede asignar el gasto total en las dos etapas. La primera etapa, podría caracterizarse como un mayor estado del gasto total en un grupo extenso de bienes como alimentos, protección, entretenimiento y, un segundo estado, más bajo, de bienes individuales como, cereales, carne, casa, ropa, internet y deportes. La separabilidad de las preferencias y el presupuesto en dos etapas, están íntimamente relacionadas, pero esto no significa que la una implique la otra, lo que sí es cierto, es que la separabilidad en (4.2) es necesaria y suficiente para el segundo estado de presupuesto.

Si algún grupo de bienes aparece solamente en una sub-función de utilidad separable, entonces las cantidades compradas en el grupo, pueden descomponerse como una función del gasto del grupo y precios con el grupo solamente. Esto se puede ver de la gráfica (4.2). La maximización de la utilidad en (4.2) deberá implicar que ν_A, ν_P y ν_E se maximizan cada una, sujeta a la restricción de cuanto se gasta en alimentos, protección y entretenimiento, si esto no fuese así ν_A, ν_P y ν_E serán crecientes, violando la restricción presupuestaria.

Los gastos sobre q_1 y q_2 son el resultado de maximizar $\nu_A(q_1, q_2)$ sujeto a $p_1q_1+p_2q_2=x_A$ el gasto total sobre el ingreso, así el gasto total sobre alimentos puede escribirse como

(4.3) $q_i = g_{Fi}(x_F, p_1, p_2)$; i =1,2

Donde q_i es la demanda marshalliana para un subgrupo. ¿Qué tipo de relación existe entre la débil separabilidad y el segundo estado? En primer lugar, la separabilidad sobre las preferencias impone restricciones sobre el comportamiento y limita los posibles efectos sustitución entre los bienes en diferentes grupos. En segundo lugar, aparte de los efectos ingreso, un cambio en el precio de los cereales podría afectar las cantidades de gasolina o deporte. Para mantener la consistencia, diremos que los siguientes lemas deben cumplirse

Lema 1: Las preferencias son débilmente separables intertemporalmente. Este supuesto depende del período de tiempo; es decir el supuesto implica que la demanda de cada bien, en cada período, es una función del gasto total y precios en cada período.

Lema 2: Si el ocio es débilmente separable de los bienes, la asignación del gasto total es independiente de las decisiones sobre las horas, lo cual es claramente imposible para bienes que sean complementarios totalmente como el ocio y los bienes recreacionales: la television, etc..; o entre bienes sustitutos como viajar al trabajo y el ocio, sin embargo podría ser aceptable para el tamaño de los gastos de los consumidores. En un segundo estado del presupuesto, se deberá usar la débil separabilidad. La asignación de todo el gasto, sin considerar las horas trabajadas será válida si todos los bienes son separables del ocio.

Lema 3: El gasto sobre el bien se relaciona con el gasto en el grupo y precios del grupo solamente.

Lema 4: Aplicación de racionamiento: Si algún bien o grupo de bienes es racionado y, si otro grupo de bienes es separable del bien racionado o el grupo racionado, entonces el efecto del racionamiento sobre los bienes en el grupo separable se hace a través del gasto total en el grupo. Si uno de los bienes es racionado y los otros bienes son separables de éste, el gasto necesario para comprar la ración se deduce simplemente del gasto total y, lo que queda se asigna entre los otros bienes independientemente de la ración. Como ejemplo, tomaremos el ocio: Si un consumidor no tiene elección sobre el número de horas trabajadas y si los bienes son separables débilmente del ocio, lo que se gasta como resultado del ingreso es explicable sin necesidad de usar el número de horas actualmente trabajadas, lo inverso es también importante: Si los bienes no son débilmente separables del ocio y si el consumidor es restringido en las horas trabajadas, las horas trabajadas aparecerán como un argumento exógeno en la asignación del gasto.

4.3 Separabilidad y sustitución intergrupal

La separabilidad débil, implica restricción sobre el grado de sustituibilidad entre los bienes, en grupos diferentes. Suponga que las preferencias separables son representadas por una función de utilidad de la forma

(4.4) $U=F[\nu_1(q_1), \nu_2(q_2), ..., \nu_G(q_G)]$

Con los subvectores $q_1,...,q_G$ y una función F(.) creciente en sus argumentos. La anterior función de utilidad implica un subgrupo de demanda o demandas condicionales de la forma

(4.5) $q_i = g_{Gi}\left(x_G, p_G\right)$; $\forall$ i $\in$G

Con $x_G = \sum_k p_{Gk} q_{Gk} =$ la cantidad total gastada sobre el grupo G. Ahora suponga, que existe un i $\in$ G y un j $\in$ H donde G $\neq$ H son subgrupos. Si diferenciamos (4.5) con respecto a j, manteniendo constante la utilidad, el efecto será a través de x_G

(4.6) $S_{ij} = \left.\frac{\partial q_i}{\partial x_G}\frac{\partial x_G}{\partial p_j}\right|_{\text{U constante}} \qquad S_{ji} = \left.\frac{\partial q_j}{\partial x_H}\frac{\partial x_H}{\partial p_i}\right|_{\text{U constante}} = S_{ij}$

Por simetría

(4.7) $\frac{\partial q_i}{\partial x_G}\frac{\partial x_G}{\partial p_j} = \frac{\partial q_j}{\partial x_H}\frac{\partial x_H}{\partial p_i} \Rightarrow \frac{\frac{\partial x_G}{\partial p_j}}{\frac{\partial q_j}{\partial x_H}} = \frac{\frac{\partial x_H}{\partial p_i}}{\frac{\partial q_i}{\partial x_G}}$

Y, como $\frac{\frac{\partial x_H}{\partial p_i}}{\frac{\partial q_i}{\partial x_G}}$ es independiente de j, nosotros podemos representar esta cantidad por λ_{GH}. Haciendo $\frac{\partial x_G}{\partial p_j} = \lambda_{GH}\frac{\partial q_j}{\partial x_H}$ y por (4.6) conocemos que $S_{ij} = \frac{\partial q_i}{\partial x_G}\frac{\partial x_G}{\partial p_j}$, de donde se deduce

(4.8) $S_{ij} = \frac{\partial q_i}{\partial x_j}\lambda_{GH}\frac{\partial q_j}{\partial x_H}$

Y, haciendo $\mu_{GH}\frac{\partial x_G}{\partial x}\frac{\partial x_H}{\partial x}=\lambda_{GH}\Rightarrow S_{ij}=\partial q_i\partial q_j\frac{\mu_{GH}\frac{\partial x_G}{\partial x}\frac{\partial x_H}{\partial x}}{\partial x_j\partial x_H}$, entonces

$$\Rightarrow S_{ij}=\partial q_i\partial q_j\mu_{GH}\frac{\partial x_G\partial x_H}{\partial x\partial x\partial x_j\partial x_H}\Rightarrow$$

(4.10) $S_{ij}=\frac{\partial q_i}{\partial x}\frac{\partial q_j}{\partial x}\mu_{GH}$

La ecuación (4.10) muestra, la condición necesaria y suficiente para separabilidad débil. Esta ecuación, resume las implicaciones empíricas de la separabilidad. La sustituibilidad entre bienes en grupos diferentes está limitada como es natural. La cantidad μ_{GH} resume la interrelación entre los grupos y es independiente de i y de j.

4.4 Separabilidad y aditividad

La hipótesis de separabilidad [Sono(1962), Leontief(1947)] implica que la utilidad puede ser aditiva o separable. Una función de utilidad es aditiva si

(4.11) $U=F\lfloor\nu_1(q_1)+\nu_2(q_2)+\ldots+\nu_g(q_g)\rfloor$

Donde, las utilidades individuales ν_n están en función de las cantidades consumidas y F(•) es una función aditiva de n utilidades. Adicionalmente, una función será aditiva grupalmente si cada función de utilidad está definida por (4.2). Una función es separable si toma la forma (4.4). Pearce (1964) ha mostrado como la ecuación de Slutzky, correspondiente a una función de utilidad grupalmente aditiva tiene la forma

(4.12) $K_{ij}=\frac{\partial q_i}{\partial p_j}+q_j\frac{\partial q_i}{\partial Y}=\phi\frac{\partial q_i}{\partial Y}\frac{\partial q_j}{\partial Y}=K_{ji}$

Para algún i $\in$ I, j $\in$ J, I$\neq$J. Donde K_{ij} son los términos de sustitución del ingreso compensado, q son las cantidades p los precios e Y el ingreso. Expresando (4.12) en elasticidades, se obtiene

(4.13) $\frac{e_{ij}}{w_j}+E_i=\frac{\phi}{Y}E_iE_j=\theta E_iE_j=\frac{e_{ji}}{w_i}+E_j$

Cuando se expresa (4.13) como una restricción no lineal, obtenemos

(4.14) $e_{ij}=w_j\theta E_iE_j-w_jE_i$

Siendo e_{ij} es la elasticidad de la demanda para el bien (i) con respecto al precio del bien (**y** $\theta = \frac{\phi}{y}$). E_i es la elasticidad ingreso de la demanda para el bien (i) y $w_j = p_j q_j / Y$ es la proporción del gasto total usada en el bien (j). El coeficiente de elasticidad ($-Y/\phi$) es denominado por Frish(1959) como la "flexibilidad de la utilidad marginal del dinero" y por Barten(1964) como "la elasticidad de la utilidad marginal del dinero". La ecuación de Slutzky para una función de utilidad separable, puede escribirse como:

(4.15) $$\frac{e_{ij}}{w_j} + E_i = \theta_{IJ} E_i E_j = \frac{e_{ji}}{w_i} + E_j = \theta_{JI} E_i E_j$$

Para Pearce(1964) los coeficientes de separabilidad entre grupos θ_{IJ} pueden ser interpretados como la medida del nivel general de sustitución entre diferentes grupos que representan diferentes deseos. En torno a los trabajos empíricos, se ha llegado a la distinción, entre: aditividad, aditividad grupal, y separabilidad. En torno, a presentar un conjunto de restricciones, deberá adicionarse la homogeneidad, la agregación de Engel, la simetría de Slutzky, lo que se puede expresarse como

(4.16) Homogeneidad $$e_{ij} = \left(-\sum_{\substack{j=1 \\ i \neq j}}^{n} e_{ij} \right) - E_i$$

(4.17) Agregación de Engel $$E_n = \frac{1}{w_n} - \sum_{j=1}^{n-1} \frac{w_j}{w_n} E_j$$

(4.18) Simetría Slutzky $$e_{ij} = \frac{w_j}{w_i} e_{ji} + w_j E_j - w_j E_i$$

(4.19) Separabilidad (o Separabilidad de Pearce) $$e_{ij} = w_j \theta_{IJ} E_i E_j - w_j E_i$$

Claramente, la restricción de separabilidad depende sobre la hipótesis de aditividad y de la agrupación elegida. Deberá observarse, que las condiciones locales para una Slutzky o para la separabilidad, implican cambios pequeños bajo el supuesto de que los gustos no cambian.

4.5 Pruebas de separabilidad

La mayoría de las pruebas de separabilidad son desarrolladas por Byron (1969), Jorgenson-Lau (1975) y Pudney (1981), quienes han usado esta técnica, para encontrar patrones de separabilidad entre bienes con cierto grado de separabilidad en un período determinado.

Barten, ha comprobado la hipótesis de la restricción de separabilidad entre bienes y ocio usando series de tiempo para datos en U.S.A y ha rechazado la separabilidad. Los resultados en últimas podrán sugerir una considerable especificación errónea de los estudios tradicionales.

Es también posible usar un test de separabilidad de corte transversal entre bienes y ocio como lo hace Muellbauer(1981). Sea la siguiente función de gasto

(4.20) $C(\eta, w, p) = d(p) + b(p)w + \left[a(p)^{1-\delta} w^{1-\delta} \mu\right]$

Donde w es el salario, d(p), b(p) y a(p) son función de p homogéneas de grado 1, 0, 1 respectivamente. A través del Lema de Sheppard, Muellbauer muestra que

(4.21) $q_i = \alpha_i + \beta_i w + \gamma_i \eta$
(4.22) $w_h = \alpha_0 + \beta_0 w + \gamma_0 \eta$

Para un ingreso transferido de η, una cantidad de horas trabajadas h y, una serie de parámetros α, β, γ constantes en una serie de corte transversal. Se puede observar, que (4.20) satisface (4.4) para el ocio a través de los bienes solamente cuando b(p) es constante, esto implica en (4.21) y (4.22) que $\frac{\beta_i}{\lambda_i}$ sea independiente para todo i=1,..,n.

Se estima el sistema (4.20) por mínimos cuadrados ordinarios y se usa la prueba de Wald para los (n-1) restricciones, i= 1,....,(n-1) esto significa probar la restricción

(4.23) $B_i\gamma_n - \gamma_i B_n = 0$

Deaton(1981) sugiere que existe poco conflicto con la separabilidad. Blundell y Walker (1982) usando una variación de (4.20) rechazan la hipótesis de que el ocio de las esposas sea separable de los bienes. Deberemos observar, que probar la separabilidad entre diferentes períodos de tiempo es muy difícil, ya que es imposible obtener estimadores no restringidos de los efectos sustitución entre los bienes individuales a través de los diferentes períodos.

5 La función de producción de hogares

Entre 1965 y 1966 los artículos de Gary Becker y Kevin Lancaster, introducen el concepto de Función de Producción de Hogares (household production function). De esta forma, los consumidores en lugar de obtener la utilidad directamente de los bienes comprados en el mercado, derivan ésta de los atributos que poseen los bienes; por ejemplo, aunque el consumidor compre alimentos crudos en el mercado, la utilidad se deriva de consumir una comida que ha sido producida a través de combinar alimentos crudos con trabajo, tiempo y otros insumos.

Muchos bienes parecen ser producidos de la forma anterior. Al igual que los alimentos, la ropa y gran parte de los bienes parecen exhibir una gama de variedades y cualidades. Los consumidores, parecen seleccionar una o pocas de estas cualidades y privarse completamente del consumo de otras. Por ejemplo, Becker (1965) propone que "ver una ópera" depende de una serie de insumos como el tiempo, los actores, etc. Y por ejemplo, "dormir" depende del insumo cama, del hogar y tiempo. De igual forma, "el jugo de naranja" se produce con un vector de características tales como calorías y vitamina c y tiempo.

El álgebra de maximización, a la cual estamos acostumbrados, indica que debemos clasificar a un bien como X_1 y otro bien como X_2 aplicando este análisis de igual forma a naranjas, kiwi, peras, manzanas o autos.

Esta forma de clasificar los bienes hace que nosotros consideremos la carne de res y la carne de cerdos como sustitutos o un disquete y un programa de computadora como complementarios. Pero esta idea, tiene su fundamento en la tecnología de usar dichos bienes particularmente. Esto es, la vía a través de la cual se combina una serie de insumos y tiempo en orden a producir alguna utilidad.

Lancaster (1966) postula que el vector de bienes X, comprado en el mercado al vector de precios P se transforma por alguna función Z=g(X), en la cual, los atributos Z producen alguna utilidad. En forma general, el problema se puede plantear como

(5.1) $\text{Maximizar}_z \quad \mu = \mu(z)$

$\text{Sujeto a} \quad Z = g(X)$

$P X = Y$

Siendo Y el Ingreso total del consumidor. Combinando la función de transformación y la función de utilidad se puede plantear el problema como

(5.2) Maximizar $\mu = \mu(g(X)) = \nu(X)$; Sujeto a $P X - Y = 0$

A este nivel de generalidad, el modelo de Lancaster es equivalente al modelo de utilidad estándar asumiendo que $\nu(X)$ exhibe las mismas propiedades de una función de utilidad. Sin embargo, las curvas de demandas compensadas $X = X^*(P, \nu)$ definidas como la solución de

(5.3) Minimizar $PX = Y$; Sujeto a $\nu(X) = \nu_0$

Muestran que las derivadas parciales no tienen realmente efectos puros de sustitución en el sentido tradicional, ya que cambios en la producción, cambios en Z a través de g(X), podrán tomar lugar como cambios en precios solamente aunque la matriz $\partial(X^*) / \partial P$ sea semidefinida negativa, lo cual al nivel de generalidad propuesta lo hace indistinguible del modelo de utilidad estándar.

En orden a conservar una estructura observable, Lancaster impone algunas restricciones sobre la función de transformación g(X). Para Lancaster g(X) deberá ser lineal Z = b.X siendo b una matriz de coeficientes tecnológicos constantes.

Por otro lado, b deberá ser constante entre los consumidores, esto significa que la tecnología por medio de la cual se convierten las X's bienes en los atributos Z's es la misma para todos los consumidores. Si la matriz b difiere de cada consumidor existe poca probabilidad de que el modelo sea operacional. Normalmente una de las X's es tiempo y puede estar en función del salario.

Adicionalmente, la función de producción deberá ser separable, de esta forma si el conjunto de producción es convexo, es fácil mostrar que la cuasiconcavidad en $\mu(Z)$ implica la cuasiconcavidad de ν (X). De otra forma, existirá una función de producción conjunta y la tecnología subyacente en t(Z,X) no podrá expresarse en términos de dos funciones de producción separadas y existirán rendimientos crecientes a escala. Si existen rendimientos crecientes, el conjunto de producción no es convexo, y esta no-convexidad mostrará unos bienes cuya función de utilidad no es cuasicóncava. Suponga $\mu(Z_1,Z_2) = Z_1^{1/2} Z_2^{1/2}$ y la tecnología en el hogar está dada por las funciones de producción $Z_1 = X_1^4$ y $Z_2 = X_2^4$ la función de utilidad será entonces $\nu(X_1,X_2)= X_1^2 X_2^2$ la cual claramente no es cuasicóncava. Esto es apenas evidente, si hacemos que una de las X's sea la fuerza de trabajo y existen indivisibilidades en el uso de la misma el resultado será el antes mencionado[11]. Por ejemplo, si X es el tiempo de un individuo y con este tiempo se puede comer y ver simultáneamente televisión, existirán entonces rendimientos crecientes en el uso de X. Por lo tanto, no se puede asumir producción conjunta en t y la función de costos conjunta deberá tener como restricción t(Z,X) = 0. En general, la restricción podría no ser lineal en los Z's y los costos marginales de los bienes podrían ser solamente funciones del bien elegido.

Para alcanzar el máximo de utilidad de Z, Z^*, el consumidor necesariamente deberá comprar en el mercado los bienes X's que producen Z^* al menor costo. De esta forma el consumidor deberá también solucionar el problema lineal

(5.4) $\text{Min} \{ P.X \mid t(Z,X) = 0 \}$; Sujeto a $bX = Z^*$

En el caso de n bienes que satisfacen Z, el problema se plantea como

(5.5) $\text{Max } \mu(g(x))$

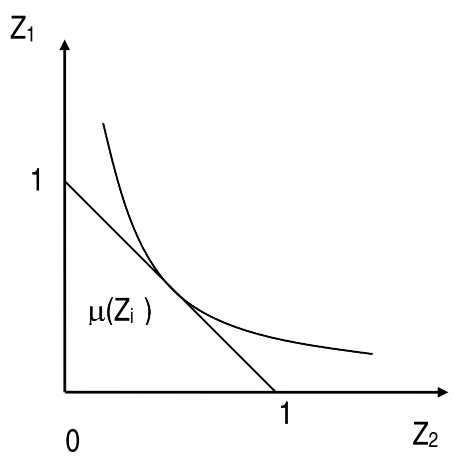
Sujeto a $bX_1 \geq Z^*$

[11] Aplicando el teorema de B.K Ross y M Star con indivisibilidades suficientes y disminución de los costes de coordinación el resultado será rendimientos crecientes.

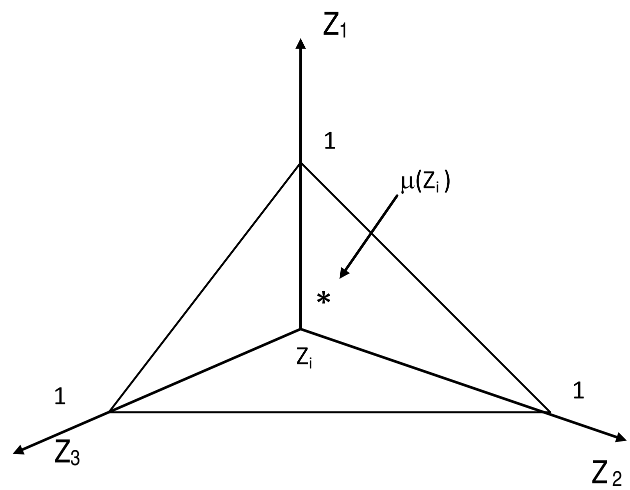
$$bX_n \geq Z^*$$
$$(X_1,....,X_n) \geq 0$$
$$(P_1,.....,P_n) \geq 0$$

Una condición suficiente, es que el conjunto de los Z's alcanzables consista ahora en un poliedro convexo n-dimensional con esquinas y caras en Z_i donde $\mu(Z_i) \in Z$, como se puede observar a continuación

Gráfica 5.1. Dos atributos en Z_i



Gráfica 5.2. Tres atributos en Z_i



Si los cambios en los costes de la tecnología, son menores que el costo de producir algún atributo Z_i, el cambio en algún bien podría seguirse realizando al menor costo de producción y al mismo tiempo maximizar los atributos Z_i de la utilidad.

Por ejemplo, si para producir un artículo usted usa una computadora y dado que los productores de computadores van introduciendo mejoras con el pasar de un Pentium I a un Pentium II 286 a 386 a 486 a Pentium I, II, III. Aun cuando un individuo cambie para producir el artículo, la utilidad derivada de éste no ha cambiado. La idea de una computadora que usa Pentium como un nuevo bien es lo que el análisis tradicional nos indica. Sin embargo, esta idea puede replantearse ya que la invención de una nueva "computadora" no debe generar una reorganización en el conjunto de preferencias sino una nueva solución al problema de la minimización de los costes que involucra el atributo "computadora".

Como podrá observarse, identificar y medir los atributos puede ser más difícil que medir los bienes de mercado. Incluso si existiesen pocas variables, medir y predecir los cambios en los coeficientes tecnológicos es algo complejo. En general, el modelo funciona bien en aquellos bienes que tienen atributos aditivos y no conflictivos (Silberberg 1985), en el caso de las computadoras el principio será aplicable para aquel que pase de 286 a 486 comprando las tarjetas necesarias e incluso cambiando a Pentium, en tanto, no modificará el artículo por usted producido ni la utilidad derivada de éste. A diferencia de Lancaster, Gary Becker incorpora decisiones respecto al uso del tiempo en el modelo estándar de la utilidad al considerar el costo del tiempo en términos de un uso gastado en producir ingreso. De esta forma, Becker provee las bases para explicar los cambios en el consumo como cambios en el ingreso laboral en términos del efecto sustitución, el cual tiene un signo conocido, así en lugar de tener que asumir supuestos ad-hoc sobre el efecto ingreso, ya conocemos su signo.

Si el incremento en el ingreso, es resultado de un incremento en el salario entonces este cambio se verá representado en el valor marginal del ocio. Deberá, por lo tanto, esperarse que el consumidor sustituya bienes intensivos en tiempo, aquellos bienes cuyo consumo involucra usar más tiempo, por bienes que tengan un menor costo de tiempo. De esta forma, los cambios en el consumo ya no son ad-hoc dado que un cambio en los gustos, o un cambio en el signo del efecto ingreso, son el resultado de la ley de la demanda.

Becker, asume que la utilidad es una función de un vector de atributos Z_i, $\mu = \mu(z)$ pero Becker adopta una estructura para la producción de atributos, para cada Z_i tendríamos

(5.6) $T_i = t_i Z_i$; $X_i = b_i Z_i$

Donde t_i es un parámetro que indica, por unidad de tiempo, el consumo de tiempo para cada Z_i consumido. De esta forma, el gasto total de tiempo consumido en alguna cantidad Z_i es T_i. Por otro lado, b_i es un parámetro que indica la cantidad del bien de mercado X_i requerido por unidad de Z_i. Así, aquellos atributos que tengan mayores valores de t_i serán intensivos en tiempo.

Los consumidores maximizarán la utilidad de los atributos consumidos sujetos a las restricciones de presupuesto y a las restricciones de tiempo.

Supongamos, que T represente el tiempo total disponible para todas las actividades (por ejemplo 24 Horas), denotando T_w como la cantidad de tiempo gastado trabajando a alguna tasa de salario constante y asumiendo que los individuos tienen un ingreso individual no salarial (herencias, arriendos, etc) en la cantidad η, entonces el problema para el consumidor vendrá dado como

(5.7) $$\begin{aligned} \text{Max} \quad & \mu = \mu(Z_1,\ldots, Z_n) \\ \text{Sujeto a} \quad & \textstyle\sum p_i x_i = wT_w + \eta \\ & \textstyle\sum T_i = T - T_w \end{aligned}$$

Dado que el tiempo y el mercado de bienes se encuentran relacionados por las ecuaciones de producción (5.6), las dos restricciones pueden ser combinadas reemplazando T_w en la restricción del ingreso por $T - \sum T_i$ de la segunda restricción lo cual da como resultado la siguiente restricción

(5.8) $\sum p_i x_i = w(T - \sum T_i) + \eta$

Al despejar obtenemos

(5.9) $\sum p_i x_i + w \sum T_i = w \; T + \eta$

Y, al reemplazar (5. 6) en (5.9) obtenemos

(5.10) $\sum p_i b_i Z_i + w \sum t_i Z_i = wT + \eta$

Así, finalmente el problema para el consumidor se puede plantear como

(5.11) Max $\mu = \mu(Z1,...., Zn)$
Sujeto a $\sum(p_i b_i + w t_i) Z_i = w T + \eta$

Además, puede interpretarse $\alpha_i = p_i b_i + w t_i$ como el precio total de consumir Z_i. El precio implícito de Z_i es independiente de la elección final de Z_i y solamente depende de la tecnología de la función de producción de hogares, específicamente, debemos asumir que $Z_i = g(X_i)$ tiene rendimientos constantes a escala(Pollack y Watcher) como ya se había mencionado.

Cuando una unidad de algún atributo se consume, el gasto monetario será $p_i b_i$ más el gasto en tiempo de t_i horas, este tiempo puede ser usado en la producción de ingreso en la cantidad $w t_i$ que será el costo de oportunidad de consumir Z_i.

En este modelo, el tiempo para “holgazanear” y dormir son atributos y parte del conjunto de los Z_i’s. Dado que no involucran algún gasto monetario el valor de b_i para dichos bienes será de cero. Todo el tiempo, deberá ser valorado a alguna tasa constante de salario w, asumiendo entonces que el individuo tiene disponible mucho trabajo que el o ella desea a la tasa de salario presente.

Como el tiempo total gastado en consumir todos los atributos es $T_c= T- T_w = \sum T_i$ y, si las condiciones de segundo orden se mantienen, la solución de las condiciones de primer orden producirá el conjunto de demandas Marshallianas

(5.12) $Z_i=Z_i^Y(\alpha_1,...., \alpha_n, w,Y) = Z_i^Y(p,b,t,w, \eta)$

Donde p,b y t son vectores de precios, coeficientes tecnológicos e intensidades de tiempo respectivamente.

5.1 Estática comparativa

Bajo el análisis de estática comparativa, nos interesan las respuestas de los consumidores ante un cambio en el salario y los coeficientes tecnológicos. Como cualquier modelo de maximización de la utilidad, todos los parámetros del modelo de Becker entran en la restricción, y las implicaciones usuales pueden ser derivadas de la maximización solamente. Considerando los efectos de sustitución puros la demandas Hicksianas se obtienen de la siguiente forma

(5.13) Min $\sum(p_i\, b_i + w\, t_i\,)\, Z_i\; - w\;\, T$
Sujeto a $\mu\,(Z_1,...., Z_n) = \mu^{\,0}$

Si las condiciones de primer y segundo orden se mantienen, las demandas Hicksianas serán

(5.14) $Z_i^{\,h} = Z_i^{\,\mu}\,(\alpha_1,...., \alpha_n, w, \mu^{\,0}\,) = Z_i^{\,\mu}\,(p,b,t,w, \mu^{\,0})$

La estructura del modelo en α_i y Z_i es formalmente idéntica al modelo estándar de minimización del gasto, esto es $\partial Z_i^{\,\mu} / \partial\, \alpha_i < 0$. Dados los cambios paramétricos en p_i, b_i y t_i se incrementará en una cantidad proporcional α_i, de esto se sigue que $\partial Z_i^{\,\mu} / \partial\, p_i < 0$, $\partial Z_i^{\,\mu} / \partial\, b_i < 0$ y $\partial Z_i^{\,\mu} / \partial\, t_i < 0$. Dadas las restricciones tecnológicas y definiendo las demandas Hicksianas para los bienes de mercado y el tiempo gastado en cada bien como $X_i^{\,\mu}$ y $T_i^{\,\mu}$ podemos observar

$$
(5.15)\quad
\begin{aligned}
\frac{\partial X_i^{\,\mu}}{\partial p_i} &= \frac{b_i \partial Z_i^{\,\mu}}{\partial p_i} < 0 \\
\frac{\partial X_i^{\,\mu}}{\partial t_i} &= \frac{b_i \partial Z_i^{\,\mu}}{\partial t_i} < 0 \\
\frac{\partial T_i^{\,\mu}}{\partial p_i} &= \frac{t_i \partial Z_i^{\,\mu}}{\partial p_i} < 0 \\
\frac{\partial T_i^{\,\mu}}{\partial b_i} &= \frac{t_i \partial Z_i^{\,\mu}}{\partial b_i} < 0
\end{aligned}
$$

Sin embargo, los efectos de un cambio en b_i sobre X_i y de t_i sobre T_i son ambiguos

$$
(5.16)\quad \frac{\partial X_i^{\,\mu}}{\partial b_i} = \frac{\partial (b_i Z_i^{\,\mu})}{\partial b_i} = Z_i^{\,\mu} + b_i \frac{\partial Z_i^{\,\mu}}{\partial b_i} \frac{Z_i^{\,\mu}}{Z_i^{\,\mu}} = Z_i^{\,\mu}(1+\varepsilon_b)
$$

Donde ε_b es la elasticidad del consumo en Z_i, con respecto a los cambios en el coeficiente b_i uniendo X_i con Z_i. La elasticidad es negativa, sin embargo, el signo de la magnitud entera depende de la proporción relativa en la unidad que tenga cada Z_i. De igual forma, para T_i tendremos

$$
(5.17)\quad \frac{\partial T_i^{\,\mu}}{\partial t_i} = T_i^{\,\mu}(1+\varepsilon_t)
$$

Donde ε_t es la elasticidad consumo de Z_i con respecto a los cambios en el coeficiente t_i. Esto se puede interpretar de la siguiente forma: suponga que t_i se incrementa, un aumento en t_i significa que el consumo de alguna cantidad Z_i requiere ahora de un mayor tiempo, lo cual aumenta el precio total de Z_i y además reduce las cantidades consumidas. Solamente cuando la reducción en Z_i es una proporción mayor que el incremento en t_i se reducirá la cantidad total de tiempo gastada sobre el atributo, sin embargo, como $X_i = b_i Z_i$ una disminución en Z_i deberá llevar a consumir menos del bien X_i del cual se deriva.

El análisis sobre los cambios en salarios es todavía más problemático. El parámetro w entra en el precio total de cada uno de los Z_i para el cual el tiempo es consumido. Un cambio en w además necesariamente deberá cambiar muchos precios simultáneamente (recreación, ocio, etc.) siendo complicado un análisis de demanda. Dado que w aparece en todas las ecuaciones de primer orden (primeras derivadas) es imposible establecer ecuaciones de demanda compensadas. Por esta razón, Becker arguye que si el salario se incrementa, el consumo podría cambiar de bienes que son más intensivos en tiempo a aquellos menos intensivos en tiempo. Esto parece plausible, pero debe hacerse supuestos adicionales sobre los valores de varios de los parámetros en el modelo para obtener un resultado riguroso.

El efecto sustitución, surge de la consideración sobre el número total de horas trabajadas, sin embargo no existe un signo determinado. Si usamos la relación $\sum T_i = T - T_w$ entonces el modelo de minimización del gasto en todas las n+1 variables, las $Z_1, \ldots, Z_n$, T_w y las dos restricciones, se expresa como

$$
\begin{aligned}
(5.18)\quad \text{Min} \quad & \sum p_i\, b_i\, Z_i - w\, T_w \\
\text{Sujeto a} \quad & \mu\,(Z_1, \ldots, Z_n) = \mu^0 \\
& \sum T_i + T_w = T
\end{aligned}
$$

El parámetro w no aparece en las restricciones, éste entra en la función objetivo en una forma simple $-wT_w$ como un precio de T_w, esto es, $\partial(-wT_w^{\mu})/\partial w < 0$ lo cual indica que la función de gasto es cóncava en w y T_w^{μ} muestra la demanda compensada de las horas trabajadas.

Dado que el ocio en este tipo de modelos, significa el tiempo total gastado en consumir los Z_i's, entonces $\partial(T_w^{\mu})/\partial w = \partial(T - T_w^{\mu})/\partial w < 0$.

En un modelo simple de ocio y trabajo, un incremento compensado en los salarios representa un incremento en el coste de oportunidad del ocio y lleva a una caída en el ocio consumido y a un incremento en el número de horas trabajadas.

Las teorías económicas de la familia, de las tasas de nacimiento, del número de hijos óptimos, de la participación en el mercado de trabajo, de la diferenciación entre grupos de hombres y mujeres e incluso el reciente auge en los modelos medioambientales del coste de viaje, se derivan de aquí. Mayores salarios en el mercado para las mujeres, por ejemplo, aumentan el coste de oportunidad de los niños y de otras tareas que deberán realizar las mujeres en el hogar, y si incluso pensáramos en un "niño" como un bien inferior, un mayor ingreso está asociado con familias pequeñas. De esta forma, el incremento en el consumo de "bienes convenientes" por familias con dos trabajadores puede ser atribuido a salarios de mercado más altos y mayores salarios compraran ítems con "mayores cualidades" donde la cualidad del atributo reduce la cantidad de tiempo dedicado a las tareas en el hogar (reparaciones, atención de los niños, etc.).

La teoría de la producción de hogares nos da para pensar más rigurosamente sobre la importancia de las elecciones y provee un marco para reemplazar las explicaciones basadas en los gustos, por aquella basada en el cambio en las oportunidades.

5.2 Análisis de la riqueza en el mercado de bienes

La función de producción de Hogares es usada también "para analizar el daño realizado por la contaminación del aire, o los beneficios derivados de actividades recreativas, o proyectos de evaluación social" (Pollack 1978, pág. 28). Esta aproximación depende de la distinción entre bienes comprados y bienes consumidos, y en particular, del uso de la medición de los beneficios derivados de los bienes públicos.

Los trabajos de Willig(1976) y Hausman(1981) emplean el teorema de la dualidad para demostrar que dada la unión entre el gasto y las funciones de utilidad, la demanda compensada no observada (debido a los atributos Z_i) puede ser encontrada a partir de la función de demanda Marshalliana que sí es observada.

Bockstael y MacConell(1983) por su parte tienen serios reparos en los trabajos anteriores. Como ellos mencionan es imposible derivar la curva de demanda Marshalliana de la compensada dada la ausencia de precios exógenos, esto es, la utilidad y la función de gasto existen, pero la ausencia de precios para los atributos impide directamente usar la identidad de Roy para recuperar la marshalliana de la función de utilidad indirecta. Deberá observarse también, que es imposible moverse de una función de demanda compensada a una única función de gasto debido a las no linealidades en la función de gasto cuando existen diferentes tecnologías en la producción de los Z_i's y, además, diferentes funciones de gastos podrán ser asociadas con los mismos valores de los costes marginales. Las medidas de riqueza pueden ser derivadas en un espacio de bienes pero de una forma diferente. Por simplicidad, se usará Z_1 pensando que algún Z_i podrá ser elegido. Supongamos la siguiente partición de bienes

$Z=(Z_1, \bar{Z})$ donde $\bar{Z} = (Z_2,...,Z_n)$. Al derivar la función de gasto condicionada sobre el atributo Z_1 encontramos

$$(5.19)\ E(Z_1, p, \mu^0) = \underset{Z}{Min}[C(Z_1, \bar{Z}, p) | \mu^0 = \mu(Z_1, \bar{Z})]$$

Por el teorema de la envolvente

$$(5.20)\ \frac{\partial E(Z_1, p, \mu^0)}{\partial Z_1} = -(\lambda\mu_1(Z_1, \bar{Z}^*, \mu^0) - C_1(Z_1, \bar{Z}^*, p))$$

Donde $\bar{Z}^{*} = \bar{Z}^{*}(Z_1, p, \mu^0)$ es ajustado óptimamente cuando Z_1, p y μ^0 cambian. El primer término a la derecha en (5.20), es el valor marginal compensado para Z_1 y el segundo término es el costo marginal de producir Z_1. De esta forma (5.20) refleja el cambio en el gasto necesario para mantener el nivel de utilidad μ^0 cuando Z_1 se incrementa. Esta expresión podrá ser negativa para $Z_1 < Z_1^*$(la cantidad optima de Z_1) .

A través, de $E(Z_1, p, \mu^0)$, se puede calcular la variación compensada asociada con el consumo de Z_1^*. Esta medida refleja el cambio en el ingreso, el cual, podrá alcanzar el consumidor a su nivel de utilidad y se halla integrando entre 0 y Z_1^* como se puede observar

$$(5.21) \quad \int_0^{Z_1^*} [\frac{\partial E(Z_1, p, \mu^0)}{\partial Z_1}] \partial Z_1$$

$$= E(Z_1,p, \mu^0) - E(0,p, \mu^0)$$

Esta expresión, es la parte negativa del área entre el valor marginal compensado y el costo marginal para Z_1. En últimas, dicha área puede ser expresada como

$$(5.22) \int_0^{Z_1^*} [\lambda\mu_1(Z_1, \bar{Z}^*, \mu, \bar{Z}^*, \mu^0) - C_1(Z_1, \bar{Z}^*, p)] \partial Z_1$$

5.3 Bienes Públicos

Supongamos que α sea un bien medioambiental, tal como la calidad del aire, un lago o un paisaje. Entonces, α entra en la función de utilidad directamente y es complementario con algún bien denota como Z_1, por ejemplo recreación. De esta forma, la utilidad es una función de α y Z. Observe también, como α entra en la función de transformación $t(z,x, \alpha)$ y la función de gasto dependerá de α. Cuando μ^0 es el nivel de riqueza inicial, la variación compensada de un cambio en el vector de parámetro de α^0 a α' esta dado por

(5.23) $VC= C(p, \mu^0, \alpha') - C(p, \mu^0, \alpha^0)$

Donde C es la función de gasto. Si α es un bien público la variación compensada es negativa para un incremento en α y positiva cuando ésta cae. La medida dada por (5.23) no es directamente observable como observan Bockstael y MacConnell para evaluar los efectos de riqueza cuando existe un cambio en α dada la información sobre la producción y el consumo asociado a la del bien Z_1.

La diferencia entre el valor marginal y el costo marginal evaluado de α^0 a α' es equivalente a

$$(5.24)\quad \int_0^{Z_1^*(\alpha')} [\frac{\partial E(Z_1, p, \mu^0, \alpha')}{\partial Z_1}]\partial Z_1 - \int_0^{Z_1^*(\alpha^0)} [\frac{\partial E(Z_1, p, \mu^0, \alpha^0)}{\partial Z_1}]\partial Z_1$$

Esta expresión, puede escribirse como

(5.25) $E(Z_1^*(\alpha'),p, \mu^0, \alpha') - E(0,p, \mu^0, \alpha') - E(Z_1^*(\alpha^0),p, \mu^0, \alpha^0) - E(0,p, \mu^0, \alpha^0)$

Análogamente $E(Z_1^*(p, \mu^0)p, \mu^0)= C(p, \mu^0)$ de donde se deduce

(5.26) $C(p, \mu^0, \alpha') - C(p, \mu^0, \mu^0) - E(Z_1^*,p, \mu^0, \alpha') + E(0,p, \mu^0, \alpha^0)$

5.3.1 De igual forma, se puede observar

(5.27) $E(Z_1^*,p, \mu^0, \alpha') = E(0,p, \mu^0, \alpha^0)$

De esta forma, la variación compensada asociada con un cambio en α puede ser medida a través de los cambios en el área perteneciente a la demanda compensada y la curva de costo marginal para Z_1. Además (5.26) es la medida correcta de los cambios en la riqueza dados por (5.23). Por otro lado, α deberá entrar en la función de preferencias del hogar directamente como una condición suficiente para que (5.27) se mantenga, y deberá también satisfacer que sea débilmente complementaria con Z_1.

Karl-Göran Mäler define la ***débil complementariedad*** de la forma siguiente: "Si la demanda para un bien privado es cero, entonces la demanda para algún servicio medioambiental [bien público] podría también ser cero" (pag 183). Esta débil complementariedad consiste en $\frac{\partial \mu(0, \bar{Z}, \alpha)}{\partial \alpha} = 0$ lo cual implica que el individuo es indiferente a variación en los niveles del bien exógeno cuando él no consume Z_1. Alternativamente cuando α entra como un insumo en el proceso productivo la condición (5.27) se mantiene si α es solamente un insumo en la producción de Z_1.

6 Variables dependientes discretas y limitadas

Existen muchos fenómenos en la actividad económica que responden a elecciones discretas como la decisión de trabajar, la decisión de comprar una bien, la decisión de votar por un candidato, etc.

Para Amemiya (1981), existen dos factores que explican el interés en los modelos de respuesta cualitativa: Primero, los economistas trabajan con modelos que involucran más de una variable discreta, más de dos respuestas, y por supuesto, más de una variable independiente. Segundo, el creciente número de encuestas que se realizan y la posibilidad de trabajar los datos que éstas producen.

A continuación, se desarrollarán algunos modelos estadísticos cuyo objetivo consiste en facilitar la contrastación empírica de la teoría del consumidor. Estos modelos son el de probabilidad lineal, el Logit, el Probit y, el Tobit en sus diferentes versiones. Luego se presentará una versión del modelo de autoselección de Heckman y, finalmente, el modelo de variables latentes.

6.1 Especificación del modelo

Suponga que usted desea considerar la ocurrencia de un evento como "comprar un carro" para describir este evento, definiremos una variable aleatoria dicotómica Y, la cual tomará el valor de 1 si el evento ocurre y 0 si no ocurre. De igual forma, deberemos asumir que la probabilidad del evento depende sobre un vector de variables independientes x* y un vector de parámetros desconocidos θ. El subíndice i denota el i-ésimo individuo. De esta forma, un modelo general dicotómico univariado, se puede expresar como

(6.1) $p_i = p(Y_i = 1) = G(x^*_i, \theta)$; $i = 1,2,...., n$.

Los Y_i son distribuidos independientemente. Por otro lado, dado que Y_i es la probabilidad de comprar un auto, x^*_i estaría representando aquellas variables que explican Y, como el ingreso, el sexo, la edad, el estatus, la educación del individuo i, así como los precios del auto. Ya que (6.1) es muy general, el investigador deberá escoger alguna función $H(x^*_i, \theta)$, la cual, se conoce sobre un vector de parámetros θ

(6.2) $p(Y_i = 1) = F[H(x^*_i, \theta)]$

6.2 Formas comunes de las funciones de probabilidad

Considere el siguiente modelo de consumo de automóviles: el consumidor responderá Y=1 si compra el automóvil y Y=0 si no lo compra. Dado que nosotros consideraremos que los factores x^*_i, explican la decisión que toma el consumidor

(6.3) $\text{Prob}(Y = 1) = F(\beta' x)$
$\text{Prob}(Y = 0) = 1 - F(\beta' x)$

El conjunto de parámetros β refleja el impacto de los cambios en x sobre la probabilidad. Una primera forma de representar este evento es considerar la siguiente regresión lineal

(6.4) $F(x, \beta) = \beta' x$

Dado que $E[Y] = F(x, \beta)$, el modelo de regresión será

(6.4.1) $Y_i = E[Y_i] + (Y_i - E[Y_i])$
$= \beta' x + \varepsilon$, con $\varepsilon = Y_i - E[Y_i]$

Que se conoce como el modelo de Probabilidad Lineal. El modelo de Probabilidad Lineal tiene como "debilidad" que $\beta' x_i$ no está restringido a pertenecer entre 0 y 1 como debería ser en términos de la probabilidad. Esto significa, que no se satisface en el modelo de probabilidad lineal la condición $0 < \beta' x_i < 1$.

Amemiya (1981) nos indica que este defecto podrá corregirse si definimos F=1 si $F(\beta' x_i) > 1$ y F=0 si $F(\beta' x_i) <1$, lo cual produce puntos de truncamiento no reales. Este procedimiento, fue utilizado en los primeros años de investigación por su simplicidad computacional.

Igualmente, el modelo de probabilidad lineal presenta problemas de heterocedasticidad, ya que ε deberá ser igual a uno o cero. Entonces ε deberá ser igual tanto a $-\beta' x$ como a $1- \beta' x$ con probabilidades 1-F y F por lo cual $V[\varepsilon|x] = V(Y_i) = \beta' x_i (1-\beta' x_i)$.

Como menciona Green(1999) restringir $B' x_i$ al intervalo (0, 1) produciría probabilidades y varianzas negativas. Dadas las desventajas del modelo de Probabilidad Lineal, su interés ha ido decayendo lo cual ha originado que modelos como el Logit o Probit se usen más frecuentemente. Veamos en qué consisten estos modelos.

Suponga que para un vector de regresores se cumple que

(6.5) $\lim_{\beta' x \to +\infty} \text{Prob}(Y_i = 1) = 1$

$\lim_{\beta' x \to +\infty} \text{Prob}(Y_i = 0) = 0$

Asumiendo una distribución normal estándar Φ, obtenemos el modelo Probit

(6.6) Prob (Y=1) = $\int_{-\infty}^{\beta' x} \phi(t)dt$

$$= \Phi(\beta' x) = \int_{-\infty}^{\beta' x} \frac{1}{\sqrt{2\pi}} e^{-t^2/2} dt$$

De igual forma, se obtiene una distribución logística

(6.7) Prob (Y=1) = $\frac{e^{\beta' x}}{1+e^{\beta' x}}$ = $\Lambda(\beta' x)$ = $L(\beta' x)$

Donde $\Lambda(\beta' x)$ es la función de distribución acumulativa logística. Las funciones de distribución Logit y Probit están acotadas entre 0 y 1. La distribución logística tiene varianza igual a $\pi^2/3$. Considere ahora

(6.8) $L_\lambda(\beta' x) = \frac{e^{\lambda\beta' x}}{1+e^{\lambda\beta' x}}$

Para un valor apropiado de λ se puede aproximar (6.8) a una normal estándar. Esto es, cuando $\lambda = \pi/\sqrt{3}$ se tendría una distribución con media cero y varianza unitaria. Amemiya (1981) muestra los diferentes valores, para los cuales (6.6) y (6.7) difieren a partir de (0,5). Debido a la similitud entre (6.6) y (6.7), es difícil distinguir estadísticamente entre estas funciones, a menos que uno tenga una gran cantidad de datos [Chambers y Cox(1967)].

De esta forma, en modelos univariados dicotómicos no es posible distinguir cuándo usar Logit o Probit, a menos, que exista una concentración en la cola dadas las características del problema estudiado. Debido a las anotaciones anteriores, la distribución logística tiende a dar mayores probabilidades a Y= 0 que la distribución normal cuando $\beta' x$ es muy pequeño y probabilidades menores que la distribución normal a Y = 0 cuando $\beta' x$ es muy grande. De esta forma, el tamaño de la muestra podría influenciar nuestros resultados aunque asintóticamente ambas distribuciones no difieren significativamente[12].

[12] Greene (1999) menciona que los modelos generan predicciones diferentes si una muestra contiene pocas respuestas afirmativas(Y=1) o pocas respuestas negativas (Y=0) y cuando existe una gran variación en la variable independiente de importancia.

Amemiya (1981) muestra las equivalencias en los parámetros entre ambos modelos. Si $\hat{\beta}$ es el parámetro estimado en ambos modelos

$$(6.9)\qquad \begin{cases} \hat{\beta}_{(6.7)} \approx 0.4\,\hat{\beta}_{(6.6)} & excepto \quad para \quad el \quad termino \quad constante \\ \hat{\beta}_{(6.7)} \approx 0.4\,\hat{\beta}_{(6.6)} + 0.5 & para \quad el \quad termino \quad constante \end{cases}$$

$$\begin{cases} \hat{\beta}_{(6.7)} \approx 0.25\,\hat{\beta}_{(6.4)} & excepto \quad para \quad el \quad termino \quad constante \\ \hat{\beta}_{(6.7)} \approx 0.25\,\hat{\beta}_{(6.4)} + 0.5 & para \quad el \quad termino \quad constante \end{cases}$$

Cuando se desea comparar modelos con diferentes funciones de probabilidad, es mejor comparar directamente las probabilidades que los coeficientes estimados. De esta forma, observaremos las derivadas de las probabilidades con respecto a una variable independiente particular. Sea $x_{i\,k}$ el k elemento del vector x_i y sea $\beta_{i\,k}$ el k elemento del vector de parámetros β_i, entonces, las derivadas para (6.4), (6.6) y (6.7) serán

$$(6.10)\qquad \begin{cases} \dfrac{\partial}{\partial x_{ik}} \beta' x_i = \beta_k \\ \dfrac{\partial}{\partial x_{ik}} \Phi(\beta' x_i) = \phi(\beta' x_i)\beta_k \\ \dfrac{\partial}{\partial x_{ik}} L(\beta' x_i) = \dfrac{e^{\beta' x_i}}{(1+e^{\beta' x_i})^2}\beta_k \end{cases}$$

Donde ϕ es la función de densidad normal estándar. En el límite, como muestran Amemiya (1981) y Green (1999) la diferencia no es sustancial.

6.3 Estimación

A excepción del modelo de Probabilidad Lineal, los modelos Probit y Logit se estiman por máxima verosimilitud donde cada observación es extraída de una distribución de Bernoulli. El modelo con una probabilidad de suceso f(β'x) y observaciones independientes lleva a una probabilidad conjunta o a una función de verosimilitud de la forma

$$(6.11)\quad \Pr ob\left(Y_1 = y_1, Y_2 = y_2, Y_3 = y_3, \ldots, Y_n = y_n\right) = \prod_{y_i=0}\left[1 - F\left(\beta' x_i\right)\right]\prod_{y_i=1} F\left(\beta' x_i\right)$$

Reacomodando términos, (6.11) puede escribirse como

(6.12) $L = \prod_i \left[F\left(\beta' x_i\right)\right]^{y_i} \left[1 - F\left(\beta' x_i\right)\right]^{1-y_i}$

Que es la probabilidad para una muestra de n observaciones. Tomando logaritmos

(6.13) $Ln\ L = \sum_i y_i\ LnF\left(\beta' x_i\right) + \left(1 - y_i\right) Ln\left[1 - F\left(\beta' x_i\right)\right]$

Al derivar con respecto al vector de parámetros se obtiene

(6.14) $\frac{\partial Ln\ L}{\partial \beta} = \sum_i \left[\frac{y_i F(\beta' x_i)}{F(\beta' x_i)} + (1 - y_i) - \frac{F(\beta' x_i)}{1 - F(\beta' x_i)}\right] x_i = 0$

La elección de una F(•) particular lleva a un modelo empírico. Entre las formas disponibles para calcular (6.13) se encuentra el método de algoritmos de Newton, Newton-Rampson, Máxima verosimilitud. Hoy día, calcular un Logit o un Probit es bastante sencillo, pues estos métodos se encuentran en paquetes estadísticos como el RATS, SAS, SPSS, GAUSS, LIMDEP, E-Views y el STATA debiendo solamente especificarse que algoritmo se desea.

6.4 Algunos modelos aplicados

En economía, la tradición de usar modelos Logit y Probit es extensa, aquí menciono tan sólo algunos modelos, quedando en deuda con el resto.

6.4.1 Domencich y McFadden

Considérese a un individuo que toma la decisión entre conducir o usar un método alternativo para ir al trabajo (autobús, metro, etc.). La utilidad que se asocia a cada forma de transporte, está en función de las características Z (principalmente el tiempo y el costo en que se incurre en cada elección) y las características individuales socioeconómicas w, más un término aleatorio de error ε. Definiendo μ_{i1} y μ_{i0} como la utilidad indirecta[13] asociada a la i- ésima persona cuando conduce o toma otro transporte y cuando esta función es lineal

(6.15) $$\begin{aligned} \mu_{i0} &= \alpha_0 + \beta' Z_{i0} + \gamma'_0\, w_i + \varepsilon_{i0} \\ \mu_{i1} &= \alpha_0 + \beta' Z_{i1} + \gamma'_0\, w_i + \varepsilon_{i1} \end{aligned}$$

De esta forma, la i- ésima persona conduce el automóvil si $\mu_{i1} > \mu_{i0}$ y viaja en otro tipo de transporte si $\mu_{i1} < \mu_{i0}$. El individuo estaría indeciso si $\mu_{i1} = \mu_{i0}$ pero esto sucederá

[13] En el capítulo siguiente, se discutirá con mayor detalle los modelos de utilidad aleatoria.

con una probabilidad de cero si ε_{i1} y ε_{i0} son variables continuas aleatorias. Definiendo y_i = 1 si la i- ésima persona conduce un automóvil y cero de otra forma, se tendrán las probabilidades

$$(6.16) \quad \Pr ob\left(y_i = 1\right) = \Pr ob\left(\mu_{i1} > \mu_{i0}\right) = \Pr ob\left[\varepsilon_{i0} - \varepsilon_{i1} < \alpha_1 - \alpha_0 + \beta'\left(Z_{i1} - Z_{i0}\right) + \left(\gamma_1 - \gamma_0\right)' w_i\right]$$

$$= F\left[\left(\alpha_1 - \alpha_0\right) + \beta'\left(Z_{i1} - Z_{i0}\right) + \left(\gamma_1 - \gamma_0\right)' w_i\right]$$

Donde F es la función de ε_{i0} - $\varepsilon_{i1.}$ Si se asume una distribución normal para ε_{i1} - ε_{i0} se obtendrá un Probit.

Domencich y McFadden sustentan que el término aleatorio de error está determinado por el tipo de transporte, que a su vez, vendrá determinado por una serie de características socioeconómicas que no son observadas por el investigador. Convenientemente debe asumirse que las utilidades de individuos diferentes son distribuidas independientemente, es decir, que la correlación entre ε_{i1} y ε_{i0} es cero.

Por otro lado, la idea de que las constantes α_o y α_1 en (6.15) sean diferentes no es muy consistente con la idea "abstracta de una forma de transporte" ya que esto significa diferencias en el efecto de transportar en una forma especifica más que diferencias en los atributos que son los que explican uno u otro. Por esta razón, los autores deberán asumir que α_o= α_1 y excluir el término constante. Sin embargo, las constantes se incluyeron con el fin de ajustar mejor los datos.

El modelo utilizado fue un Logit para una encuesta con 115 individuos en Pisttburg(U.S.A) en 1967, los resultados fueron

$$(6.17) \quad \hat{y} = -\underset{(0.51)}{3.82} + \underset{(0.05)}{0.158}Tw - \underset{(0.25)}{0.382}(Aiv - Tss) - \underset{(0.58)}{2.56}(Ac - F) + \underset{(1.07)}{4.94}A/w - \underset{(1.37)}{2.91}R - \underset{(1.17)}{2.36}Z$$

Desviaciones estándar entre paréntesis, y Tw es el tiempo de tránsito a pie (en minutos), Aiv es el tiempo en automóvil, Tss es el tiempo de transporte entre estaciones, Ac es el costo de parqueadero más costos de operar tal vehículo (en dólares), F es el precio del pasaje en autobús, A / w es el número de autos por trabajador en el hogar, R es la raza (0 si es blanco, 1 si no es blanco), Z es la ocupación (0 si no es trabajador profesional, 1 si es trabajador profesional). El ingreso y otras variables socioeconómicas se excluyen dado que son comunes en ambas funciones de utilidad.

En el modelo de elección de transporte, se puede también definir un índice que nos muestre la propensión a preferir un carro con relación a transitar en algún otro transporte para la i- ésima persona. Sea y^*_i = $\beta' x_i - \xi_i$ donde x_i es el vector de variables que aparece a la derecha de (6.17) y ξ_i corresponde a ε_{i0}-$\varepsilon_{i1.}$ Entonces, el modelo de transporte vendrá determinado por la distribución específica de ξ_i. De esta forma si definimos y_i = 1 sí y solo sí $y^*_i > 0$ encontraremos un índice continuo.

6.4.2 Lee, L.F.

Lee define la propensión del i- ésimo trabajador de unirse a un sindicato como

(6.18) $$y_i^* = \beta_1 + \beta_2 \left[\frac{w_{i1} - w_{i0}}{w_{i0}} \right] + \beta'_3\, x_i - \xi_i$$

Donde w_{i1} y w_{i0} será el salario cuando el trabajador pertenece un sindicato y cuando no respectivamente, x_i es un vector de características del i- ésimo trabajador así como los atributos en la industria en la cual está empleado. El trabajador se une a un sindicato ($y_i = 1$) sí y solo sí $y^*_i > 0$. Lee asume una distribución normal para ξ_i y estima los parámetros por máxima verosimilitud para un Probit.

6.4.3 Pencavel

Pencavel estudia cómo inciden en las decisiones de trabajar de la esposa y el esposo la ayuda económica brindada por el gobierno de los Estados Unidos en Seattle y Denver. De esta forma, estima la probabilidad de trabajar de la esposa usando 1657 familias durante 2 años. Las variables que el autor usa son: F que es igual a uno si la familia pertenece al experimento y cero de lo contrario, L que es igual a uno si el esposo trabaja durante el año anterior al experimento y cero de lo contrario, Y que es igual a uno si la observación es extraída del segundo año de experimento y cero si es extraída del primer año, U que es igual a uno si el esposo estuvo desempleado durante el año. Se estimaron dos modelos: uno que reporta estimaciones bajo un modelo de probabilidad lineal $\hat{y}_{PL}$ y otro que estima un modelo Logit $\hat{y}_L$, como se observa a continuación

(6.19) $$\hat{y}_{PL} = -0.069F + 0.497L + 0.0055Y + 0.043U + 0.036F.L - 0.08F.Y - 0.041F.U$$

(6.20) $$\hat{y}_L = -\underset{(0.15)}{0.327}F + \underset{(0.126)}{2.305}L + \underset{(0.121)}{0.309}Y + \underset{(0.131)}{0.25}U + \underset{(0.169)}{0.097}FL - \underset{(0.167)}{0.425}FY - \underset{(0.175)}{0.23}FU$$

Entre paréntesis, los errores estándar. Por otro lado, los términos constantes fueron incluidos en el modelo pero no fueron reportados. Como observa Pencavel las probabilidades estimadas por los dos modelos son parecidas.

6.5 Modelo de efectos fijos y aleatorios en datos de panel

Considérese el modelo estructural Probit para datos de panel [12]

(6.21) $y_{it}^{*} = \beta' x_{it} + u_{it} \qquad u_i \sim N(0,1)$, i = 1,....,n ; t = 1, ...,T_i

$y_{it}^{*} = 1 \text{ si } y_{it} > 0$ y cero de otra forma

Si las u_{it} son variables estándar independientes, la naturaleza de los datos de panel es irrelevante. Como el modelo Probit en sí mismo no tiene efectos fijos, esto es, $u_{it} = \alpha_i$, remover la heterogeneidad presente en los datos, sobre todo, en corte transversal es bastante complicado, por esto, idealmente nosotros debemos especificar que u_{it} y u_{is} sean libremente correlacionados (en el grupo y no entre grupos) lo cual involucrará calcular la probabilidad conjunta de una distribución normal bivariada de dimensión T, lo cual también es problemático.

En contraposición al Probit, el Logit no lleva en sí mismo a tratamientos de efectos fijos, así que obtendremos

(6.22) $$\Pr ob(y_{it} = 1) = \frac{e^{\alpha_i + \beta' x_{it}}}{1 + e^{\alpha_i + \beta' x_{it}}}$$

El cual es un modelo no lineal. Chamberlain (1980) sugiere que cuando el conjunto de datos tiene un gran n y el tiempo es corto, la probabilidad no condicionada para nT observaciones vendrá dada por

(6.23) $$L = \prod_i \prod_t (F_{it})^{y_{it}} (1 - F_{it})^{1-y_{it}}$$

Chamberlain, sugiere maximizar la función condicional de verosimilitud

(6.24) $$L^c = \prod_i \text{prob}\left(Y_{i1} = y_{i1}, Y_{i2} = y_{i2}, \ldots Y_{iT} = y_{it} \middle| \sum_t y_{it} \right)$$

De esta forma, (6.24) muestra la probabilidad para un conjunto de T observaciones condicionadas sobre el número de unos en el conjunto. Por ejemplo, suponga que la

[12] Son datos obtenidos a través de un seguimiento a lo largo del tiempo de secciones cruzadas; por ejemplo, un seguimiento a las familias en diferentes periodos de tiempo (Greene, pág. 256).

muestra consiste en un gran número de unidades de corte transversal y que cada observación se da solamente en los períodos 1 Y 2. La probabilidad no condicionada será

(6.25) $$L = \prod_{i} prob\left(Y_{i1} = y_{i1}\right) prob\left(Y_{i2} = y_{i2}\right)$$

Si las observaciones son independientes, la función de verosimilitud es el producto de las probabilidades y, para cada par de observaciones se tienen las siguientes posibilidades

- $y_{i1} = 0$ y $y_{i2} = 0 \Rightarrow$ prob (0,0|suma=0) = 1
- $y_{i1} = 0$ y $y_{i2} = 1 \Rightarrow$ prob (1,1|suma=2) = 1

Como podrá observarse el i- ésimo término en la función de verosimilitud condicionada será uno en el caso de que y_{i1}=0 y que y_{i2}=1. Por lo cual, tendremos

- prob((0,1) | suma = 1) = $\dfrac{prob(0,1)}{prob(0,1) + prob(1,0)}$

En términos generales, la probabilidad condicionada para un par de observaciones vendrá dada por

(6.26) $$\frac{e^{\beta' x_{i2}}}{e^{\beta' x_{i1}} + e^{\beta' x_{i2}}}$$

Al condicionar las probabilidades a la suma en las dos observaciones se remueve la heterogeneidad existente. Por lo tanto, la función de verosimilitud condicional será el producto de los conjuntos de observaciones para los que la suma no es cero ni T.

6.6 El modelo Logit condicionado

Esta es una versión reciente, para incluir los atributos presentes en los bienes. Suponga que exista un modelo de elección no ordenada que provenga de una utilidad aleatoria para el i- ésimo consumo en j elecciones. De esta forma, la utilidad de la elección j es

(8.27) $$U_{ij} = \beta' Z_{ij} + \varepsilon_{ij}$$

Si el consumidor realiza una elección j en particular y, asumiendo que U_{ij} es el máximo entre j utilidades, el modelo estadístico que depende de la elección j será

(8.27.1) Prob ($U_{ij} > U_{ik}$) $\forall$ $k \neq j$

Sea y_i una variable aleatoria indicando la elección realizada, McFadden muestra que sí y solo sí las j perturbaciones son independientes e idénticamente distribuidas siguiendo una distribución Weibull

(6.28) $$F\left(\varepsilon_{ij}\right) = \mathbf{exp}\left(e^{-\varepsilon_{ij}}\right)$$

(6.28.1) $$prob\left(y_i = j\right) = \frac{e^{\beta' Z_{ij}}}{\sum_j e^{\beta' Z_{ij}}}$$

El cual se denomina Logit Condicionado. Usualmente la elección depende de los atributos de Z_{ij}. Así, tendremos a x_{ij}, que varía entre los individuos, además de las w_i características individuales. De esta forma, el modelo puede plantearse como

(6.28.2) $$prob\left(y_i = j\right) = \frac{e^{\beta' Z_{ij} + \alpha' w_i}}{\sum_j e^{\beta' Z_{ij} + \alpha' w_i}}$$

Schmidt y Strauss (1975), estiman un modelo de ocupación basado en una muestra de 1000 observaciones cuya variable dependiente es la: Ocupación que es igual a 1 si es empleado doméstico, 2 si es obrero no especializado, 3 si Artesano(trabajador manual), 4 si es oficinista y 5 si es trabajador profesional. En el conjunto de variables independientes se incluyeron la constante, la educación, la experiencia, la raza y el sexo. El modelo incluyendo los estratos sociales será

(6.28.2) $$prob(y_i = 1) = \frac{e^{\beta'_j x_i}}{\sum_{k=0}^{4} e^{\beta'_k x_i}}$$

Debe observarse que las probabilidades estimadas dependen del estrato 1, 2, 3, 4 y 5. De esta forma, el modelo condicional Logit computa las probabilidades relativas a cada estrato, el estrato podrá contener pocos casos o muchos casos. Por lo tanto, la probabilidad de que la i- ésima observación pertenezca a un estrato social S_i, viene dada por

(6.28.3) $$P(Y_i = 1) = \frac{e^{\sum_{j=0}^{p} \beta'_j x_{ij}}}{\sum_{m \in S_i} e^{\sum_{j=0}^{p} \beta'_j x_{mj}}}$$

6.7 Modelos multinomiales

Un modelo multinomial de respuesta cualitativa se define de la siguiente forma. Asuma que la variable dependiente yi toma mi + 1 valores { 0, 1, 2, . . ., mi }, entonces el modelo multinomial vendrá dado

(6.29) $P(y_i = j) = F_{ij}(\mathbf{x^*}, \theta)$; i = 1, 2, . . ., n y j = 1, 2, . . .,m_i.

Donde **x*** y θ son vectores de variables independientes y parámetros respectivamente. m_i depende de un i en particular cuando los individuos tienen diferentes conjuntos de elección. Para definir el estimador de θ en el modelo (6.29) usualmente se definen $\sum^{n}_{i=1} (m_i + 1)$ variables binarias, de la forma

(6.29.1) y_{ij} = 1 si $y_i = j$
= 0 si $y_i \neq j$; i = 1, 2, . . ., n y j = 1, 2, . . .,m_i.

La función de verosimilitud viene definida como

(6.29.2) $$Ln \quad L = \sum_{i=1}^{n} \sum_{j=0}^{m_i} y_{ij} LnF_{ij}$$

Donde el estimador insesgado $\hat{\theta}$ de θ se define como una solución a la ecuación $\frac{\partial LnL}{\partial \theta}$ =0.

6.7.1 Modelos ordenados

Los modelos multinomiales de respuestas cualitativas se pueden clasificar en modelos ordenados y no ordenados. Un modelo ordenado se define como

(6.30) $P(y = j \mid x, \theta) = p(S_j)$

Para alguna medida de probabilidad p, dados sobre x y θ y, una secuencia finita de intervalos sucesivos { S_j } que depende sobre x y θ tal que $\bigcup_j S_j = \Re$.

En los modelos ordenados, los valores que **y** toma, corresponden a una partición sobre la línea real. A diferencia de un modelo no ordenado, donde la partición correspondería a particiones no sucesivas sobre la línea real o a particiones de dimensiones mayores sobre el espacio euclideano. En la mayoría de las aplicaciones, el modelo ordenado toma la forma

(6.31) $P(y = j \mid x, \alpha, \beta) = F(\alpha_{j+1} - \mathbf{x}'\boldsymbol{\beta}) - F(\alpha_j - \mathbf{x}'\boldsymbol{\beta})$; $j = 0, 1,\ldots,m;\alpha_0=-\infty$; $\alpha_j \leq \alpha_{j+1}$;$\alpha_{m+1} = \infty$

Para alguna distribución F, se puede definir un modelo Probit ordenado o un modelo Logit ordenado.

6.7.2 Modelo Logit multinomial

El modelo Logit multinomial se define como

(6.32) $P_{ij} = \left[\sum_{k=0}^{m_i} \exp(x'_{ij}\beta) \right]^{-1} \exp(x'_{ij}\beta)$; i =1,2,...,n y j= 0,1,....,m_i.

McFadden(1974)[13] considera el siguiente modelo multinomial derivado del problema del consumidor. Considere a un individuo (i) cuyas utilidades están asociadas con tres alternativas, de la forma siguiente

(6.33) $U_{ij} = \mu_{ij} + \varepsilon_{ij}$, con j = 0, 1, 2

Siendo μ_{ij} una función no estocástica de variables explicatorias y parámetros ε_{ij}. El individuo elige aquella alternativa en la que obtiene la mayor utilidad. El multinomial Logit, se puede derivar del problema de maximizar la utilidad sí y solo sí los ε_{ij} son independientes y la función de distribución de ε_{ij} viene dada por $\exp[-\exp(\varepsilon_{ij})]$. De esta forma, la probabilidad de que el i individuo elija una alternativa j, será

(6.34) $P(y_i=2) = P (U_{i2} > U_{i1}, U_{i2} > U_{i0})$

$= P (\varepsilon_2 + \mu_2 - \mu_1 > \varepsilon_1, \varepsilon_2 + \mu_2 - \mu_0 > \varepsilon_0$

$$= \int_{-\infty}^{\infty} f(\varepsilon_2) \left[\int_{-\infty}^{\varepsilon_2+\mu_2-\mu_1} f(\varepsilon_1)\partial\varepsilon_1 \int_{-\infty}^{\varepsilon_2+\mu_2-\mu_0} f(\varepsilon_0)\partial\varepsilon_0 \right] \partial\varepsilon_2$$

[13] En el capítulo siguiente se profundizara en el modelo de McFadden.

$$= \int_{-\infty}^{\infty} \exp(-\varepsilon_2)\exp[-\exp(-\varepsilon_2)] \times \exp[-\exp(-\varepsilon_2 - \mu_2 + \mu_1)] \times \exp[-\exp(-\varepsilon_2 - \mu_2 + \mu_0)]\partial\varepsilon_2$$

$$= \frac{\exp(\mu_{i2})}{\exp(\mu_{i0}) + \exp(\mu_{i1}) + \exp(\mu_{i2})}$$

Y tomará una forma parecida a (6.32) si nosotros hacemos $\mu_{i2} - \mu_{i0} = x'_{i2}\beta$ y $\mu_{i1} - \mu_{i0} = x'_{i1}\beta$.

6.8 Variables dependientes limitadas

Existe un gran número de datos cuya observación nos muestra que están limitados o acotados de alguna forma. Este fenómeno lleva a dos tipos de efectos: el truncamiento y la censura.

El efecto de truncación ocurre cuando la muestra de datos es extraída aleatoriamente de una gran población de interés, por ejemplo, cuando se estudia el ingreso y pobreza se establece un valor sobre el cual, el ingreso, se encuentra por encima o por debajo del mismo. De esta forma, algunos individuos podrán no ser tenidos en cuenta.

Por otro lado censurar, es un procedimiento en el cual los rangos de una variable son limitados a priori por el investigador, este procedimiento produce una distorsión estadística similar al proceso de truncación.

6.8.1 Truncamiento

Una distribución truncada es la parte de una distribución no-truncada antes o después de un valor específico, imagínese por ejemplo que nosotros deseamos conocer la distribución de los ingresos anteriores a 100000 o el número de viajes a una zona mayores de 2, ésta será tan sólo una parte de la distribución total.

6.8.1.1 Densidad de una variable aleatoria truncada

Si una variable continua aleatoria x, tiene una función de densidad de probabilidades, pdf(x), y a es una constante, entonces

(6.35) $$f(x|x > a) = \frac{f(x)}{\text{prob}(x > a)}$$

Si x tiene una distribución normal con media μ y desviación estándar σ

(6.36) $$\text{prob}(x > a) = 1 - \Phi\left(\frac{a-\mu}{\sigma}\right)$$

$$= 1 - \Phi(\alpha)$$

Donde $\alpha = \dfrac{a-\mu}{\sigma}$ y $\Phi(\alpha)$ es función de densidad acumulativa, entonces la distribución normal truncada será

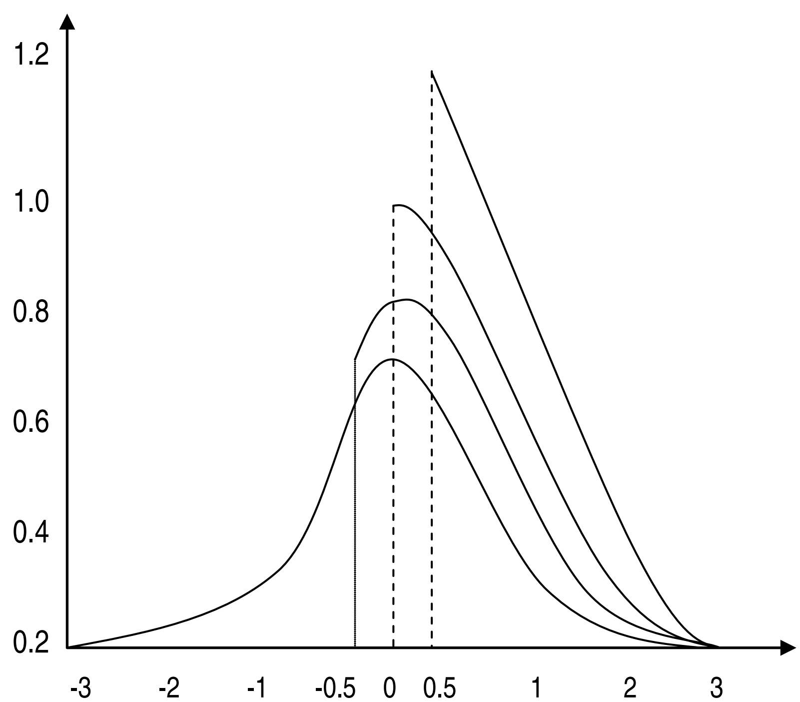
(6.37) $$f(x|x > a) = \frac{f(x)}{1-\Phi(\alpha)}$$

$$= \frac{\left(2\pi\sigma^2\right)^{-1/2} e^{-(-x-\mu)^2/\left(2\sigma^2\right)}}{1-\Phi(\alpha)}$$

$$= \frac{\left(1/\sigma\right)\phi^{(x-\mu)/\sigma}}{1-\Phi(\alpha)}$$

Donde ϕ será la normal estándar, pdf. La distribución normal estándar truncada con $\mu = 0$ y $\sigma = 1$ para a igual a -0.5, 0 y 0.5, será

Gráfica 6.1. Distribuciones normales truncadas



6.8.1.2 Momentos de una Distribución Truncada

Los momentos de mayor interés son la media y la varianza, sobre todo son muy convenientes para hallar la razón inversa de Mills. Si $x \sim N[\mu,\sigma^2]$ con μ constante, entonces la media vendrá dada por E[x|truncamiento] = $\mu + \sigma\lambda(\alpha)$, y la varianza por var[x|truncamiento]= $\sigma^2(1-\delta(\alpha))$ donde $\alpha=(a-\mu)/\sigma$. Por otro lado, nosotros observamos que

(6.38) $$\lambda(\alpha) = \frac{\phi(\alpha)}{1-\Phi(\alpha)}$$ Si el truncamiento ocurre en $x \geq a$

$$\lambda(\alpha) = \frac{-\phi(\alpha)}{\Phi(\alpha)}$$ Si el truncamiento ocurre en $x < a$

Como $\delta(\alpha) = \lambda(\alpha)(\lambda(\alpha) - \alpha) \quad donde\ 0 < \delta(\alpha) < 1 \quad \forall \quad \alpha$, y $\dfrac{\partial\phi(\alpha)}{\partial(\alpha)} = -\alpha(\alpha)$.

Donde $\lambda(\alpha)$ se conoce también como la razón inversa de Mills.

6.8.1.2.1 Estimación por máxima verosimilitud

Tomando el logaritmo de (6.37) y, al realizar la suma de los logaritmos de estas densidades nosotros obtenemos

(6.39) $$Ln\ L = \frac{-n}{2}\left(Ln(2\pi) + Ln\sigma^2\right) - \frac{1}{2\sigma^2}\sum_i \left(y_i - \beta' x_i\right)^2 - \sum_i Ln\left[1 - \Phi\left(\frac{a - \beta' x_i}{\sigma}\right)\right]$$

Las condiciones necesarias para maximizar (6.39) serán

$$\frac{\partial LnL}{\partial \beta} = \sum_i \left[\frac{y_i - \beta' x_i}{\sigma^2} - \frac{\lambda_i}{\sigma} \right] x_i = 0$$

(6.40) $$\frac{\partial LnL}{\partial \sigma^2} = \sum_i \left[\frac{-1}{2\sigma^2} + \frac{(y_i - \beta' x_i)^2}{2\sigma^4} - \frac{\alpha_i x_i}{2\sigma^2} \right] = 0$$

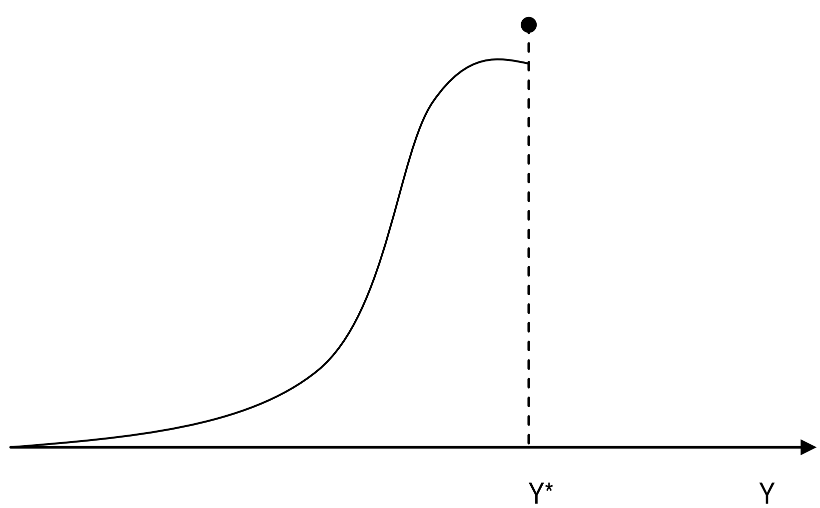
$$donde \alpha_i = \frac{a - \beta' x_i}{\sigma} y \lambda_i = \frac{\phi(\alpha_i)}{1 - \Phi(\alpha_i)}$$

6.8.2 Censuramiento

Un procedimiento normal con datos microeconómicos, consiste en censurar la variable dependiente. Cuando la variable dependiente es censurada, los valores en un determinado rango son todos transformados a un valor singular. De esta forma, si definimos una variable aleatoria y transformada de la variable original como

(6.41) y = 0 Si y* $\leq$ 0 y = y* Si y* > 0

Gráfica 6.2. Distribución censurada



La distribución correspondiente a y* ~ N[μ,σ^2] será

prob (y=0) = prob(y* $\leq$ 0)

$$= \Phi\left(\frac{-\mu}{\sigma}\right)$$
$$= 1 - \Phi\left(\frac{\mu}{\sigma}\right)$$

Si $y^* > 0$ y tiene la densidad de y^*, entonces, la distribución tiene partes discretas y continuas, donde la probabilidad total será de 1 como se requiere. Para lograr esto, se asigna la probabilidad total en la región censurada al punto de censuramiento.

6.8.2.1 Momentos

La media de una variable censurada vendrá dada por

$$E(y) = \Phi a + (1 - \Phi)(\mu + \sigma\lambda) \tag{6.42}$$

Y, la varianza

$$\mathrm{Var}(y) = \sigma^2 (1 - \Phi) \quad \text{donde}$$
$$\Phi\left[\frac{a - \mu}{\sigma}\right] = \Phi(\alpha) = \mathrm{prob}(y^* \leq a) = \Phi; \tag{6.43}$$
$$\lambda = \frac{\phi}{1 - \Phi}; \delta = \lambda^2 - \lambda\alpha$$

6.8.3 Modelos Tobit

Los modelos Tobit se refiere a modelos censurados o truncados donde el rango de la variable dependiente se restringe de alguna forma. El problema puede verse como

(6.44) $y = y^*$ Si $y^* > y_0$ $\qquad$ $y = 0$ Si $y^* \leq y_0$

Asumiendo que $y^* = \beta_1 + \beta_2 x + \mu$, con μ el término de perturbación aleatoria, y que y_0 varía entre los hogares pero que dicha variación es conocida, el modelo que genera los datos para una función de verosimilitud con n observaciones independientes será

(6.45) $$L = \prod_0 F_i\left(y_{0i}\right) + \prod_1 f_i\left(y_i\right)$$

Donde F_i y f_i son las funciones de distribución y de densidad de y_i^*. Por otro lado, Π_0 significa el producto sobre aquellos y para los cuales $y_i^* < y_{0i}$ y Π_1 significa el producto sobre aquellos y para los cuales $y_i^* < y_{1i}$. Por esta razón, el modelo estándar puede definirse como

(6.46) $$\begin{aligned} y_i^* &= \beta' x_i + \mu_i \,, \quad i = 1,2,3,\ldots,n \\ y_i &= y_i^* \,, \quad Si \quad y_i^* > 0 \\ y_i &= 0 \,, \quad Si \quad y_i^* \le 0 \end{aligned}$$

Asumiendo que μ_i ~ N[0, σ^2], la función de verosimilitud para el modelo Tobit estándar será

(6.47) $$L = \prod_0 \left[1 - \Phi\left(\frac{\beta' x_i}{\sigma}\right)\right] \prod_1 \sigma^{-1} \phi\left[\frac{y_i - B' x_i}{\sigma}\right]$$

Donde Φ y ϕ son las funciones de distribución y densidad de normal estándar. La función de verosimilitud de la versión truncada del Tobit puede escribirse como

(6.48) $$L = \prod_1 \Phi\left(\frac{B' x_i}{\sigma}\right)^{-1} \sigma^{-1} \phi\left[\frac{y_i - B' x_i}{\sigma}\right]$$

De igual forma se puede plantear la función

(6.48.1) $$f\left(y_{ij}\right) = \begin{cases} \dfrac{\frac{1}{\sigma}\phi\left[y_{ij} - B'_j\, x_{ij}\right]/\sigma}{1 - \Phi\left[-B' x_{ij}\right]/\sigma} & , Si \; y_{ij} > 0 \\ 0 & , de \; otra \;\; forma \end{cases}$$

Dado el creciente uso de los modelo tipo Tobit, Amemiya (1984) se ha dado la laboriosa tarea de clasificar de acuerdo a similitudes en la función de verosimilitud. La caracterización de los tipos de modelos Tobit es la siguiente

Tipo	**Variable Dependiente**		
	y_1	y_2	y_3
1	censurado	-	-
2	binario	Censurado	-
3	censurado	Censurado	-

4	censurado	Censurado	censurado
5	binario	Censurado	censurado

Las funciones de verosimilitud serán

Tipo	Función de Verosimilitud
1	$p(y_1<0) \,.\, p(y_1)$
2	$p(y_1<0) \,.\, p(y_1>0, y_2)$
3	$p(y_1<0) \,.\, p(y_1, y_2)$
4	$p(y_1<0, y_3) \,.\, p(y_1, y_2)$
5	$p(y_1<0, y_3) \,.\, p(y_1>0, y_2)$

Cada y_j con j =1,2,3 se asume que tiene una distribución del tipo $N(\beta'_j x_j, \sigma^2_j)$ y p denota una probabilidad o densidad o una combinación. Amemiya toma el producto de cada p sobre la observación que pertenece a una categoría particular determinada por el signo de y_1. De esta forma, el tipo 1, modelo Tobit estándar, la función de verosimilitud $[p(y_1<0).p(y_1)]$, será una notación abreviada para $\prod_0 p\left(y^*_{1i}<0\right)\prod_1 f_{1i}\left(y_1, y\right)$ donde f_{1i} es la densidad de $N(\beta'_1 x_{1i}, \sigma^2_1)$. De igual forma, en cada tipo de modelo, el signo de y_i determina cada una de las dos posibles categorías para las observaciones y la variable censurada se observa en una categoría y no en la otra.

6.8.3.1 Modelo Tobit tipo 2: $\{ p(y_1<0) \,.\, p(y_1>0, y_2) \}$

Este modelo se define como

$$
\begin{aligned}
y^*_{1i} &= \beta'_1 x_{1i} + \varepsilon_{1i} \\
y^*_{2i} &= \beta'_2 x_{2i} + \varepsilon_{2i} \\
y_{2i} &= \begin{cases} y^*_{2i} & Si \quad y^*_{1i} > 0 \\ 0 & Si \quad y^*_{1i} \leq 0 \end{cases} \quad ; \text{ i = 1, 2, 3,........., n}
\end{aligned}
\tag{6.49}
$$

Donde $\{\varepsilon_{1i}, \varepsilon_{2y}\}$ son i.i.d extraídos de una distribución normal bivariada con media cero, varianzas σ^2_1 y σ^2_2 y covarianza σ_{12}. Se asume que el signo de y^*_{1i} se observa así como y^*_{2i} solamente cuando $y^*_{1i}>0$. De igual forma, x_{1i} se observa para todos los i, sin embargo x_{2i} no necesariamente se observa para aquellos i donde $y^*_{1i} \leq 0$. La función de verosimilitud para este modelo será

(6.50) $$L = \prod_0 p\left(y_{1i}^* \le 0\right)\prod_1 f\left(y_{2i} \mid y_{1i}^* > 0\right)p\left(y_{1i}^* > 0\right)$$

Donde Π_0 y Π_1 es el producto para aquel i en el cual $y_{21} = 0$ y $y_{21} \neq 0$ respectivamente y $f(. \mid y^*_{1i}>0)$ es la densidad condicional de y^*_{2i} dado $y^*_{1i} > 0$. Este tipo de modelos, fueron originalmente usados por Nelson(1977), Dudley y Montmarquette(1976), Weslin y Gillen(1978). Veamos de modelo de Gronau(1976).

Gronau (1976) asume que el salario ofrecido W^0, a cada ama de casa es independiente de las horas trabajadas H. En este sentido, no existe un menú de salarios $W^0(H)$. Dado W^0, una ama de casa solucionará el siguiente problema

(6.50.1) Maximizar $\mu(C,X)$

Sujeto a $W^0H + V = Y$ [Restricción de ingresos]
$C + H = T$ [Restricción de tiempo]

Donde C es el tiempo gastado en el hogar para cuidar los niños, x es un vector de dos bienes, T es tiempo total disponible y V otros ingresos. De esta forma una ama de casa no trabajará si

(6.50.2) $$\left[\frac{\partial\mu}{\partial C}\left(\frac{\partial\mu}{\partial X}\right)^{-1}\right]_{H=0} > W^0$$

Y, trabajará si la desigualdad contraria se mantiene. Si ella trabaja H y la tasa actual de salario es W^0, entonces

(6.50.3) $$\frac{\partial\mu}{\partial C}\left(\frac{\partial\mu}{\partial X}\right)^{-1} = W^0$$

La parte derecha de (6.50.2) y (6.50.3) se conoce como el salario de reserva y, en adelante, se denotará como W^r. Asumiendo que W^0 y W^r son combinaciones lineales de variables independientes más el término de perturbación aleatoria ε_i, el modelo estadístico puede plantearse como

$$W_i^0 = \beta'_2 \, x_{21} + \varepsilon_{2i}$$

(6.50.4) $$W_i^r = \alpha' Z_i + \varepsilon_i$$

$$W_i = \begin{cases} W_i^0 & Si \; W_i^0 > W_i^r \\ 0 & Si \; W_i^0 = W_i^r \end{cases} ; \quad \forall \; i = 1,2,\ldots.,n$$

Donde los $\{\varepsilon_{2i}, \varepsilon_i\}$ son i.i.d extraídas de una distribución normal bivariada con media cero, varianzas $\sigma^2_{\varepsilon 2}$ y σ^2_{ε} y covarianza $\sigma_{\varepsilon 2 \varepsilon}$. Note que si se hace $W_i^0 - W_i^r = y^*_{2i}$ y $W_i^0 = y^*_{2i}$ tendremos el modelo (6.49).

6.8.3.2 Modelo Tobit tipo 3: { $p(y_1<0) \,.\, p(y_1, y_2)$ }

El tipo 3 se define como

$$y^*_{1i} = \beta'_1 \, x_{1i} + \varepsilon_{1i}$$

$$y^*_{2i} = \beta'_2 \, x_{2i} + \varepsilon_{2i}$$

(6.51) $$y_{1i} = \begin{cases} y^*_{1i} & Si \quad y^*_{1i} > 0 \\ 0 & Si \quad y^*_{1i} \le 0 \end{cases} ; \quad \forall \; i = 1,2,3,\ldots,n$$

$$y_{2i} = \begin{cases} y^*_{2i} & Si \quad y^*_{1i} > 0 \\ 0 & Si \quad y^*_{1i} \le 0 \end{cases}$$

Donde los $\{\varepsilon_{1i}, \varepsilon_{2i}\}$ son extraídos de una distribución normal bivariada con media cero, varianza σ^2_1 y σ^2_2 y covarianza $\sigma_{12.}$ Este modelo difiere del anterior en que y^*_{1i} también se observa cuando este es positivo. La función de verosimilitud vendrá dada por

(6.52) $$L = \prod_0 p\left(y^*_{1i} \le 0\right) \prod_1 f\left(y_{1i}, y_{2i}\right)$$

Donde f(.,.) es la densidad conjunta de y^*_{1i} y y^*_{2i}. Dado que y^*_{1i} se observa cuando es positivo, todos los parámetros son identificables incluyendo σ^2_1.

El modelo tipo (3) también se conoce como de autoselección tipo Heckman, a continuación se expondrá en que consiste.

Heckman(1974) propone un modelo diferente al de Gronau en el sentido de que Heckman incluye la determinación del número de horas trabajadas [H] en el modelo. Al igual que Gronau, Heckman asume que el salario ofrecido W^0 es independiente de las horas trabajadas, además la ecuación de W^0 es la misma de la ecuación de Gronau

(6.52.1) $W^0_i = \beta'_2 X_{2i} + \varepsilon_{2i}$

Y define

(6.52.2) $\frac{\partial \mu}{\partial C}\left(\frac{\partial \mu}{\partial X}\right)^{-1} = W^r$

Por lo tanto

(6.52.3) $W^r_i = \gamma H_i + \alpha' Z_i + \nu_i$

De esta forma, un individuo trabajará sí

(6.52.4) $W^r_i (H_i = 0) = \alpha' Z_i + \nu_i < W^0_i$

El salario W_i y las horas trabajadas H_i se determinan cuando se soluciona simultáneamente (6.52.1) y (6.52.3) entonces $W^0_i = W^r_i = W_i$. El modelo de Heckman, se puede definir como

(6.52.5) $W_i = \beta'_2 X_{2i} + \varepsilon_{2i}$

(6.52.6) $W_i = \gamma H_i + \alpha' Z_i + \nu_i$

Y, para el i que cumple la restricción de las horas trabajadas

(6.52.7) $H^*_i = \beta'_1 X_{1i} + \varepsilon_{1i} > 0$

Donde $\beta'_1 X_{1i} = \gamma^{-1} (\beta'_2 X_{2i} - \alpha' Z_i)$ y $\varepsilon_{1i} = \gamma^{-1} (\varepsilon_{2i} - \nu_i)$. Se puede observar que (6.52.4) y (6.52.7) son equivalentes dado que $\gamma > 0$.

6.8.4 Modelo Tobit tipo 4: { $p(y_1<0, y_3) . p(y_1, y_2)$}

El modelo Tobit tipo 4 se define como

$$
\begin{aligned}
&y^*_{1i} = \beta'_1 x_{1i} + \varepsilon_{1i} \\
&y^*_{2i} = \beta'_2 x_{2i} + \varepsilon_{2i} \\
&y^*_{3i} = \beta'_3 x_{3i} + \varepsilon_{3i} \\
&y_{1i} = \begin{cases} y^*_{1i} & Si \quad y^*_{1i} > 0 \\ 0 & Si \quad y^*_{1i} \le 0 \end{cases} ; \quad \forall\, i = 1,2,3,\ldots,n \\
&y_{2i} = \begin{cases} y^*_{2i} & Si \quad y^*_{1i} > 0 \\ 0 & Si \quad y^*_{1i} \le 0 \end{cases} \\
&y_{3i} = \begin{cases} y^*_{3i} & Si \quad y^*_{1i} \le 0 \\ 0 & Si \quad y^*_{1i} > 0 \end{cases}
\end{aligned} \tag{6.53}
$$

Donde los $\{\varepsilon_{1i}, \varepsilon_{2i}, \varepsilon_{3i}\}$ son i.i.d extraídos de una distribución normal trivariada. Este modelo difiere del tipo 3 en la adición de y^*_{3i}, el cual se observa solamente si $y^*_{1i} \le 0$. La función de verosimilitud vendrá dada por

$$
L = \prod_0 \int_{-\infty}^{0} f_3\left(y^*_{1i}, y_{3i}\right) dy^*_{1i} \prod_1 f_2\left(y^*_{1i}, y_{2i}\right) \tag{6.54}
$$

Donde $f_3(.\,,.)$ será la densidad conjunta de y^*_{1i} e y^*_{3i} y $f_2(.\,,.)$ la densidad conjunta de y^*_{1i} e y^*_{2i}. Nelson y Olson(1978) proveen el siguiente ejemplo: Sea y^*_1 el tiempo que se gasta en entrenamiento escolar vocacional, que es completamente observado si $y^*_1 > 0$. Si esta condición no se cumple entonces pertenecerá al intervalo $(-\infty, 0]$. El tiempo que se gasta en educación y^*_2 se observa que pertenece a uno de los intervalos $(-\infty, 0]$, $(0, 1]$ y $(1, \infty)$. El salario y^*_3 se observa completamente. De igual forma, y^*_4 son las horas trabajadas y son observables completamente.

Dado que cada variable pertenece a un conjunto de ecuaciones simultáneas se puede estimar solamente la forma reducida, el modelo se resume entonces, al siguiente modelo de 2 ecuaciones simultáneas

$$y^*_{1i} = \zeta_1 y_{2i} + \alpha_1' x_{1i} + \varepsilon_{1i} \tag{6.54.1}$$
$$y_{2i} = \zeta_2 y^*_{1i} + \alpha_2' x_{2i} + \varepsilon_{2i} \tag{6.54.2}$$

Donde y_{2i} se observa siempre. Por otro lado, y^*_{1i} se observa como y_{1i} si $y^*_{1i} > 0$. Los parámetros estructurales de este modelo, se estiman de la siguiente forma

- Paso 1: Estime los parámetros de la ecuación en la forma reducida para y^*_1 con un Tobit de máxima verosimilitud y para la ecuación reducida y_2 estime los parámetros por mínimos cuadrados ordinarios.

- Paso 2: Reemplace y $_{2i}$ al lado derecho de (6.54.1) por la estimación de los mínimos cuadrados ordinarios obtenido en el paso anterior y estime los parámetros de (6.54.1) a través de un Tobit de máxima verosimilitud.
- Paso 3: Reemplace y^*_{1i} en el lado derecho de (6.54.2) por su estimación obtenido en el paso 1 y estime los parámetros de (6.54.2) por mínimos cuadrados ordinarios.

6.8.4.1 Modelo Tobit tipo 5: { $p(y_1<0, y_3)$. $p(y_1>0, y_2)$}

El modelo Tobit tipo 5 se deriva del tipo 4, y se omite la ecuación para y_{1i}. Dado que solamente observamos el signo de y^*_{1i}, tendremos

$$
\begin{aligned}
y_{1i}^* &= \beta'_1 x_{1i} + \varepsilon_{1i} \\
y_{2i}^* &= \beta'_2 x_{2i} + \varepsilon_{2i} \\
y_{3i}^* &= \beta'_3 x_{3i} + \varepsilon_{3i}
\end{aligned}
$$

(6.55)
$$
y_{2i} = \begin{cases} y_{2i}^* & Si \quad y_{1i}^* > 0 \\ 0 & Si \quad y_{1i}^* \le 0 \end{cases} \qquad \forall\, i = 1,2,3,\ldots,n
$$
$$
y_{3i} = \begin{cases} y_{3i}^* & Si \quad y_{1i}^* \le 0 \\ 0 & Si \quad y_{1i}^* > 0 \end{cases}
$$

Donde los {ε_{1i}, ε_{2i}, ε_{3i}} son i.i.d extraídas de una distribución normal trivariada. La función de verosimilitud para este modelo será

(6.56)
$$
L = \prod_{0} \int_{-\infty}^{0} f_3\left(y_{1i}^*, y_{3i}\right) dy_{1i}^* \prod_{0} \int_{0}^{\infty} f_3\left(y_{1i}^*, y_{2i}\right) dy_{1i}^*
$$

Donde $f_3(.,.)$ y $f_2(.,.)$ se define como la densidad conjunta de y^*_{1i} y y_{3i} . Por otro lado, $f_2(.,.)$ es la densidad conjunta de y^*_{1i} y y^*_{2i}. Lee (1978) desarrolla el siguiente modelo : Sea y^*_{2i} el logaritmo de la tasa de salario del i- ésimo trabajador en el caso de que él o ella se unan a un sindicato. Cuando el trabajador decide unirse o no al sindicato, este evento estará determinado por el signo de la variable (y^*_{1i})

(6.57) $$y_{1i}^* = y_{2i}^* - y_{3i} + \alpha' Z_i + \varepsilon_i$$

Dado que nosotros observamos solamente si el trabajador se une al sindicato (y_{2i}) y si no se une al sindicato (y^*_{3i}). El logaritmo del salario observado, denotado por y_i, se define como

(6.57.1)
$$
\begin{aligned}
y_i &= y_{2i} \text{ si } y_{1i}^* > 0 \\
y_i &= y_{3i}^* \text{ si } y_{1i}^* \le 0
\end{aligned}
$$

Lee asume que x_2 y x_3, las variables independientes en (y_{2i}) y (y^*_{3i}), incluyen características individuales de las firmas y trabajadores tales como localización regional, tamaño ciudad, educación, experiencia, raza, sexo y riqueza. Por otro lado, Z incluye características individuales y variables que representan el costo monetario y no-monetario de ser miembro del sindicato. Dado que y^*_1 no es observado a excepción del signo, la varianza de (y^*_1) puede asumirse como unitaria. Lee estima el modelo por dos etapas tipo Heckman aplicado separadamente a (y^*_2) y (y^*_3). Amemiya (1994) define también el Tobit tipo 5 como

$$\begin{aligned} y^*_{ji} &= B'_j\, x_{ji} + \varepsilon_{ji} \\ Z^*_{ji} &= \gamma'_j\, S_{ji} + V_{ji} \\ \gamma_i = y^*_{ki} \;\; si \;\; Z^*_{ki} &= \max_j Z^*_{ji} \;\; ; \quad j = 1,2,3,\ldots \quad i = 1,2,3\ldots \end{aligned} \tag{6.57.2}$$

Donde cada y_i, x_{ji}, S_{ji} son observados. Se asume que $\{\varepsilon_{ji}, V_{ji}\}$ son i.i.d a través de i, pero correlacionado a través de j, y para cada i y j las dos variables pueden estar correlacionadas. La forma de verosimilitud vendrá dada por

$$\text{(6.57.3)}\; L = \prod_1 f\left(y^*_{1i} | Z_{1i} \text{ es el max}\right) p_{1i} \times \prod_2 f\left(y^*_{2i} | Z_{2i} \text{ es el max}\right) p_{2i} \times \prod_j f\left(y^*_{ji} | Z_{ji} \text{ es el max}\right) p_{ji}$$

Donde Π_j es el producto sobre aquel i para el cual Z^*_{ji} es el máximo(max) y p_{ji} = p (Z^*_{ji} máximo).

6.9 Contrastes de especificación

De la mano, con el desarrollo de las formas de estimación de los modelos, la literatura ha venido ofreciéndonos una serie de contrastes para conocer la "bondad" de los modelos estimados. El origen de estos contrastes, se remonta a los trabajos de Rao(1947) en lo que se conoce como " contraste Score " o "contraste de puntuación". Posteriormente Silvey(1959) propone el contraste de multiplicadores de Lagrange que no es otra cosa que el mismo contraste de Rao.

El contraste de multiplicadores de Lagrange no es el único que se pueda usar, pues están el de Hausman(1978) y el contraste de momentos condicionales[Newey(1985) y Tauchen(1985)]. Para Pagan y Vella (1989) el uso del contraste de especificación en variables dependientes limitadas no es muy común debido a la dificultad computacional de los mismos.

Los contrastes de especificación que se desarrollarán serán: El contraste de Rao "contraste Score", el contraste de especificación de Hausman, el cual parte de los trabajos de Durbin(1954) y por lo tanto se conoce también como Durbin-Hausman o Durbin-Wu-Hausman debido a los trabajos de Wu(1973), el contraste de la matriz de información de

White(1982) y el contraste de momentos condicionales sugerido por Newey(1985) y Tauchen(1985).

6.9.1 Contraste de Rao "contraste Score"

Suponga que existen n observaciones independientes y_1, y_2, y_3,......, y_n con funciones de densidad idénticas f(y, θ) donde θ es un vector p $\times$ 1 de p parámetros. Entonces L(θ), el vector *Score* d(θ), y la matriz de información I(θ) vienen definidas como

$$L(\theta) = \sum_{i=1}^{n} \ln f(y_i, \theta) = \sum_{i=1}^{n} L_i(\theta)$$

$d(\theta) = \dfrac{\partial L(\theta)}{\partial \theta}$, Para algún θ, E[d(θ)] = 0

$$I(\theta) = \text{cov}[d(\theta)] = E[d(\theta)\, d(\theta)'] = E\left[\frac{\partial^2 L}{\partial\theta\partial\theta'}\right]$$

El estimador máximo verosímil $\hat{\theta}$ viene dado por la ecuación d($\hat{\theta}$) = 0. La hipótesis a probar será H_0= h(θ)= 0, donde h(θ) es una vector de dimensión r de θ (r $\leq$ p) con una constante dada c. Rao(1948), propone el siguiente estadístico

(6.58) $dL(\bar{\theta})[I(\bar{\theta})]^{-1} d(\bar{\theta})$

Donde ($\bar{\theta}$) es el estimador máximo verosímil restringido de (θ) bajo Ho. Si Ho es cierto, entonces d($\bar{\theta}$) se espera que tienda a cero. Rao muestra que el estadístico Score tiene una distribución chi-cuadrada(χ^2) con r grados de libertad bajo Ho.

Breusch y Pagan(1980) sugieren usar este estadístico como un contraste de especificación. La ventaja del contraste Score consiste en que depende solamente de los estimadores máximos verosímiles del modelo restringido, ya que tanto el contraste Score como la matriz de información se basan en el modelo total. Una extensión del contraste Score consiste en un estimador general $\sqrt{n}$ más que en la restricción máximo verosímil, a esta extensión se le denomina el contraste Neyman-Rao[Hall y Mathiason(1990)].

Para modelos de elección binarios, Davidson y Mackinon(1989) muestran que el contraste Score con base en la matriz de información I($\hat{\beta}$) puede ser computado fácilmente y Orme(1992) muestra que las propiedades se conservan en los modelos Tobit [14].

[14] Como existen diferentes formas del contraste Score, que dependen de los estimadores de la matriz de información, en datos pequeños, las propiedades del estimador son bastante malas con respecto a la matriz

6.9.2 *El contraste Durbin-Hausman*

El contraste de especificación sugerido por Hausman se basa en la comparación de dos

conjuntos de parámetros estimados: Sea $\bar{\theta}$ un estimador de θ el cual es eficiente bajo Ho, pero inconsistente bajo H_1, y $\overset{+}{\theta}$ un estimador de θ el cual es consistente bajo Ho y H_1 pero ineficiente bajo Ho. Sea d= $\bar{\theta}$ - $\overset{+}{\theta}$ y var($\sqrt{n}\ \bar{\theta}$) = V_1 y var($\sqrt{n}\ \overset{+}{\theta}$) = V_0. El contraste de Hausman se basa en el resultado de var($\sqrt{n}$ d)=$V_1 - V_0$ bajo Ho(Rao 1973,pp.317). El test estadístico consiste en

($\sqrt{n}$ d)´($V_1 - V_0$)$^{-1}$($\sqrt{n}$ d) que tiene una distribución chi-cuadrada con p grados de libertad, que es la dimensión de θ. En el caso de que ($V_1 - V_0$)$^{-1}$ no exista, se puede usar la inversa generalizada [Rao y Mitra(1971)], pero la chi-cuadrada tendrá menores grados de libertad. Existe una variación del tipo Hausman, en el que ambos estimadores son inconsistentes bajo H_1, aunque pLim$\bar{\theta}\neq$pLim$\overset{+}{\theta}$ bajo H_1[ver Ruud(1984)].

6.9.3 *El contraste de la matriz de información de White*

La matriz de información de White(1982) se basa en el hecho de que en un modelo especificado correctamente, tendremos

(6.59) E[d(θ) d(θ)´] = $E\left[-\dfrac{\partial^2 L}{\partial\theta\partial\theta'}\right]$

Si nosotros consideramos el estadístico d(θ) d(θ)´+ $\left[\dfrac{\partial^2 L}{\partial\theta\partial\theta'}\right]$ éste deberá tener media cero para un modelo especificado correctamente. Como problema de esta constante se encuentra el hecho de que al tratar de obtener la varianza de este contraste se requieren derivadas de orden superior de al de L. Lancaster (1984) ha mostrado que éste estadístico se puede obtener a partir de la regresión

(6.60) r = $GC_1 + ZC_2 + \varepsilon_i$

Donde r es un vector de unos, G es una matriz n $\times$ p cuyos i ,j elementos son

---

de información cuando se usa el método del gradiente del producto externo (outer product gradient) [Davidson y Mackinon(1989), Orme(1990)]. Una alternativa consiste en tomar el hesiano como lo observa Taylor(1991).

$G_{i,j}(\theta) = \frac{\partial L_i(\theta)}{\partial \theta_j}$, i=1,2,3,...,n; j=1,2,3,...,p; Z es una matriz cuyos elementos típicos son

$\frac{\partial^2 L_i(\theta)}{\partial \theta_j \partial \theta_k} + \frac{\partial L_i(\theta)}{\partial \theta_j} \frac{\partial L_i(\theta)}{\partial \theta_k} \; \forall$ j=1,2,...,p; k=1,2,...,j.

Z y G son evaluados en el estimador restringido máximo verosímil. El número de columnas en Z será ½(k^2 + k). Cox(1983) y Chesher (1984) demuestran que el contraste de la matriz de información puede ser interpretado como un contraste Score para rechazar heterogeneidad o una variación de los parámetros en θ.

6.9.4 El contraste de momentos condicionados (CM)

El contraste de momentos condicionados fue sugerido por Newey(1985) y Tauchen(1985), y se basa, en la premisa de que bajo una especificación correcta, no solamente se tiene E[d(θ)] = 0 sino también condiciones de sobreidentificación como E[m (y, θ)] = 0. El contraste de momentos condicionados puede reducirse a la regresión

(6.61) i= $\hat{G}$ C + b$\hat{m}$+ε_i

Donde G es la matriz { $G_{ij} = \frac{\partial L_i(\theta)}{\partial \theta_j}$ } e i es un vector de unos. $\hat{G}$ es ortogonal a i dado

que i´$\hat{G}$ = 0 son las ecuaciones de solución máximo verosímil $\hat{\theta}$ y $\hat{m}$ será ortogonal a i si la condición de momentos se satisface. De esta forma, un contraste para la condición de momentos consiste en la hipótesis b = 0 en la ecuación (6.61). En el caso de que existan r condiciones de momentos, se define la matriz $\hat{m}$ = m ($\hat{\theta}$) de n $\times$ r, y se contrasta la hipótesis b= 0 en la regresión (6.61). Aún cuando i´$\hat{G}$ = 0 es importante incluir $\hat{G}$ en la regresión [Mackinon(1992,pp,132)]. La regresión artificial se basa en la igualdad informacional (6.59) la cual es válida solamente cuando la densidad completa de las observaciones es especificada correctamente.

Un último contraste de especificación proviene de Pregibon(1980), este contraste consiste en que si la regresión está bien especificada, no se deberían encontrar variables independientes significativas. De esta forma, una clase de error, consistirá en que la variable dependiente necesitase algún tipo de transformación en su relación con las variables dependientes, en lo que se conoce como una función de unión (LINK FUNCTION). El contraste consiste en la regresión

(6.61.1) y = $\beta' x_i + \varepsilon_i$

Donde β´ serán los paramentos estimados, entonces $\hat{\beta}' x_i$ deberá ser significativo y $(\hat{\beta}' x_i)^2$ no deberá serlo. El modelo también se usa para especificación en transformaciones de las variables independientes, de esta forma, si en un modelo inicial Logit o Probit $(\hat{\beta}' x_i)^2$ da significativo entonces alguna variable independiente podría necesitar algún tipo de transformación, una vez realizada esta $(\hat{\beta}' x_i)^2$ no deberá ser significativo.

6.9.5 Contrastes de heterocedasticidad

Entre los primeros trabajos sobre Heterocedasticidad realizados por Maddala y Nelson(1975) se argumentan que una regresión con Heterocedasticidad en los errores, los estimadores son consistentes pero ineficientes. En el caso del Tobit, el estimador máximo verosímil (ML) es inconsistente en la presencia de Heterocedasticidad[Brannas y Laitila(1989)].

En un modelo de regresión, la comprobación de Heterocedasticidad se realiza con base en los residuos del modelo de mínimos cuadrados. Pagan y Park(1993) sugieren que los contrastes existentes para probar heterocedasticidad pueden ser considerados como un contraste de momentos condicionados(CM). La condición de momentos para un contraste CM, será

(6.61) $$\frac{1}{n}\sum E[Z_i(\mu^2{}_i - \sigma^2)] = 0$$

Siendo μ_i el error con varianza σ^2, bajo Homocedasticidad, y Z_i es indicador mal especificado. Por ejemplo, si asumimos que $Var(\mu_i) = \sigma^2 (1 + Z_i \gamma)^2$ se deberá mostrar que ésta condición se sigue del contraste Score en caso del Tobit. En el caso de un Probit o Logit, sea $y^*_i = \beta' x_i + \mu_i$; $\mu_i \sim$ i.i.d$(0, \sigma^2)$. El cual es observado solamente cuando $y_i = \begin{cases} 1\ Si \quad y^*_i > 0 \\ 0 \quad De \quad otra \quad forma \end{cases}$. Dado que y^*_i es observado como una variable dicotómica, solamente β/σ es estimable. Asumiendo que $\sigma = 1$, entonces

(6.63) Probabilidad $[y_i = 1] = F(\beta' x_i)$

De esta forma, el logaritmo máximo verosímil L será

(6.63.1) $$L = \sum_{i=1}^{n}\left[y_i F(\beta' x_i) + (1 - y_i)(1 - F(\beta' x_i))\right] = \sum L_i$$

Sea $$G_{ij}(\beta) = \frac{\partial L_i}{\partial \beta_j} = \left[\frac{y_i}{F(\beta' x_i)} - \frac{1 - y_i}{1 - F(\beta' x_i)}\right] f_i\, x_{ij}$$

$$=\left[\frac{y_i - F(\beta' x_i)}{F(\beta' x_i)(1-F(\beta' x_i))}\right] f_i \, x_{ij}$$

Donde f_i es la derivada de F_i, es decir, la función de densidad. El j-ésimo elemento del vector Score d(β) es igual a $\sum_{i=1}^{n} G_{ij}(\beta)$. Si nosotros deseamos probar la hipótesis de que β=0 entonces se deberá obtener un contraste Score con nR^2 en una regresión artificial de i sobre G(β) donde i es un vector de unos. Davidson y Mackinnon(1984) denotan este contraste como LM_1 .

Considere ahora el estadístico basado en la matriz de información. Diferenciando L_i dos veces con respecto a β, obtenemos

(6.64) $$\frac{\partial^2 L_i}{\partial\beta\partial\beta'} = \frac{\partial}{\partial\beta'}\left[\frac{y_i f(\beta' x_i)}{F(\beta' x_i)} - \frac{1-y_i f(\beta' x_i)}{1-F(\beta' x_i)}\right] x_i$$

$$\frac{\partial^2 L_i}{\partial\beta\partial\beta'} = \left[\frac{y_i F(\beta' x_i) f'(\beta' x_i) - f^2(\beta' x_i)}{F^2(\beta' x_i)} - \frac{(1-y_i)1 - F(\beta' x_i) f'(\beta' x_i) + f^2(\beta' x_i)}{(1-F(\beta' x_i))^2}\right] x_i x'_i$$

Tomando $E(y_i) = F_i$, entonces

(6.64.1) $$I(\beta) = E\left[-\sum_{i=1}^{n}\frac{\partial^2 L}{\partial\beta\partial\beta'}\right] = \sum_{i=1}^{n}\frac{f^2(\beta' x_i)}{F(\beta' x_i)(1-F(\beta' x_i))} x_i x'_i$$

Davidson y Mackinnon sugieren el siguiente contraste Score: Defina la matriz R(β) de $n \times p$ cuyos elementos típicos son

(6.64.2) $$R_{ij}(\beta) = \left[F(\beta' x_i)(1-F(\beta' x_i))\right]^{-\frac{1}{2}} f(\beta' x_i) x_{ij}$$

El n-vector, tendrá como elemento típico

(6.64.3) $$i_r(\beta) = y_i\left[\frac{1-F(\beta' x_i)}{F(\beta' x_i)}\right]^{\frac{1}{2}} - (1-y_i)\left[\frac{F(\beta' x_i)}{1-F(\beta' x_i)}\right]^{\frac{1}{2}}$$

De esta forma, la matriz de información será I(β) = R´(β)R(β) y el vector Score d(β) = R´(β)r(β). De aquí, el contraste estadístico será

(6.64.4) $LM_2 = d'Id = r'R(R'R)^{-1}R'r$

El cual, no es más que nR^2 de la regresión artificial $r= RC+ \varepsilon_i$. Davidson y Mackinnon arguyen que LM_1 tiene propiedades débiles con relación a LM_2. Para comprobar Heterocedasticidad, se especifica $Var(\mu_i) = (1 + \gamma' Z_i)^2$ y se contrasta Ho: $\gamma=0$. Por otro lado, Davidson y Mackinnon especifican $Var(\mu_i) = \exp(2\gamma' Z_i)$ y Harvey(1976) $Var(\mu_i) = [\exp(\gamma' Z_i)]^2$. Reemplazando $\beta' x_i$ en el modelo de regresión por $\beta' x_i / (1 + \gamma' Z_i)$ y haciendo $\hat{\beta}$ el estimador máximo verosímil de β bajo Ho: $\gamma=0$ entonces se reemplaza β por $\hat{\beta}$ en LM_1 y LM_2. La matriz G será ahora de $n \times (p + m)$ donde m es la dimensión de γ, de esta forma

(6.65) $$\left.\frac{\partial L_i}{\partial \beta}\right|_{\beta=\hat{\beta},\gamma=0} = \left[\frac{y_i f\left(\frac{\hat{\beta}' x_i}{1+\gamma' Z_i}\right)}{F\left(\frac{\hat{\beta}' x_i}{1+\gamma' Z_i}\right)} - \frac{1-y_i f\left(\frac{\hat{\beta}' x_i}{1+\gamma' Z_i}\right)}{1-F\left(\frac{\hat{\beta}' x_i}{1+\gamma' Z_i}\right)}\right] x_i$$

(6.65.1) $$\left.\frac{\partial L_i}{\partial \gamma}\right|_{\beta=\hat{\beta},\gamma=0} = \left[\frac{y_i f\left(\frac{\hat{\beta}' x_i}{1+\gamma' Z_i}\right)}{F\left(\frac{\hat{\beta}' x_i}{1+\gamma' Z_i}\right)} - \frac{1-y_i f\left(\frac{\hat{\beta}' x_i}{1+\gamma' Z_i}\right)}{1-F\left(\frac{\hat{\beta}' x_i}{1+\gamma' Z_i}\right)}\right] \hat{\beta}' x_i(-Z_i)$$

La matriz R es ahora $n \times (p + m)$ y en la i- ésima fila tendremos

$$[R_i(\beta), R_i(\gamma)] = \left[F\left(\frac{\beta' X_i}{1+\gamma' Z_i}\right)\left(1-F\left(\frac{\beta' X_i}{1+\gamma' Z_i}\right)\right)\right]^{-1/2} f\left(\frac{\beta' X_i}{1+\gamma' Z_i}\right)\left[X_i, \beta' \chi_i(-Z_i)\right]$$

El Score vector $d(\hat{\beta},0)$ es $R'r$ el cual es un vector $(p + m)$

$$\left[f\left(\frac{\beta' x_i}{1+\gamma' Z_i}\right)\left(y_i - F\left(\frac{\beta' x_i}{1+\gamma' Z_i}\right)\right)\right]\left[F\left(\frac{\beta' x_i}{1+\gamma' Z_i}\right)\left(1-F\left(\frac{\beta' x_i}{1+\gamma' Z_i}\right)\right)\right]^{-1} x_i = 0$$

$$\left[f\left(\frac{\beta' x_i}{1+\gamma' Z_i}\right)\left(y_i - F\left(\frac{\beta' x_i}{1+\gamma' Z_i}\right)\right)\right]\left[F\left(\frac{\beta' x_i}{1+\gamma' Z_i}\right)\left(1-F\left(\frac{\beta' x_i}{1+\gamma' Z_i}\right)\right)\right]^{-1} (\beta' x_i) Z_i = 0$$

Dado que $\left[f\left(\frac{\beta' x_i}{1+\gamma' Z_i}\right)(y_i - F\left(\frac{\beta' x_i}{1+\gamma' Z_i}\right))\right]\left[F\left(\frac{\beta' x_i}{1+\gamma' Z_i}\right)(1 - F\left(\frac{\beta' x_i}{1+\gamma' Z_i}\right))\right]^{-1}$ es el residuo generalizado[Gauriéroux, Monfort y Trognon(1987)] y denotándolo por φ, entonces (6.65) puede escribirse como

(6.66) $\frac{1}{n}\sum \hat{\varphi}_i x_i = 0$ y $\frac{1}{n}\sum \hat{\varphi}(\beta' x_i)Z_i = 0$

De esta forma, el contraste de heterocedasticidad involucra solamente $\hat{\varphi}_i$ y no $\hat{\varphi}_i^{\,2}$. Veamos que sucede en el modelo Tobit.

Lee y Maddala(1985) usan la forma funcional Var(μ_i) = $\sigma^2{}_i$ = G(α + δ ´ Z_i) y sugieren probar δ = 0. Asumiendo que $\sigma_i = \sigma + \delta Z_i$ y definiendo un indicador I_i como $I_i = \begin{cases} 1 & Si \quad y^*_i > 0 \\ 0 & De \quad otra \quad forma \end{cases}$ el logaritmo de la función de verosimilitud L, viene dado por

(6.67) $L = \sum_{i=1}^{n} I_i\left[-\frac{1}{2}Ln\sigma^2{}_i - \frac{1}{2\sigma^2{}_i}(y_i - \beta' x_i)^2\right] + \sum_{i=1}^{n}(1-I_i)Ln\left[1-\Phi\left(\frac{\beta' x_i}{\sigma_i}\right)\right]$

Bajo Ho: δ = 0, entonces

(6.68) $\frac{\partial L}{\partial \delta} \approx \sum\left[I_i\left(\frac{(y_i - \beta' x_i)^2}{\sigma^2} - 1\right) + (1-I_i)\left[\lambda_i\left(\frac{\beta' x_i}{\sigma_i}\right)\right]Z_i\right]$

Donde λ_i es la razón de Mills, esto es

(6.68.1) $\lambda_i = \frac{\phi\left(\frac{\beta' x_i}{\sigma}\right)}{1-\Phi\left(\frac{\beta' x_i}{\sigma}\right)}$

El vector Score, será proporcional a

(6.68.2) $\sum\left(\hat{\mu}^2{}_i - \sigma^2\right)Z_i$

La condición de momentos para probar Heterocedasticidad en el contraste CM, es

(6.68.3) $\frac{1}{n}\sum\left(\hat{\mu}^2{}_i - \sigma^2\right)Z_i = 0$

De esta forma, el test para comprobar Heterocedasticidad se reduce al CM. Otro contraste comúnmente usado consiste en el contraste de las razones de verosimilitud o LR-test, el cual consiste en

(6.69) LR = -2[Ln $\hat{L}_r$ - Ln $\hat{L}$]

Donde $\hat{L}_r$ y $\hat{L}$ son los logaritmos de las funciones de verosimilitud del modelo con heterocedasticidad y el modelo sin heterocedasticidad. De esta forma, nosotros podemos probar Heterocedasticidad sobre una varianza del tipo Var(μ_i) = $(1 + \gamma' Z_i)^2$ ó Var(μ_i) = exp($2\gamma' Z_i$) ó
Var(μ_i) = $[\exp(\gamma' Z_i)]^2$ bajo Ho: γ=0. Por lo tanto, si no existe evidencia estadística para rechazar Ho existirá homocedasticidad y de forma contraria existirá. Además, LR ~ χ^2_{d-dr} donde d y dr son los grados de libertad asociados al modelo general y al modelo con heterocedasticidad.

6.9.6 Contrastes de normalidad

Cuando no existe normalidad, se existen sesgos en los estimadores [Goldberger(1983)]. Pagan y Vella (1989) sugieren un contraste CM para normalidad con base en el tercer y cuarto momento de los residuos. En términos generales, ya se trate de un Logit, Probit o Tobit, el problema consiste en evaluar el momento p E[μ^p_i | y_i = 0]. Lee y Maddala(1985) sugieren usar el método de recursividad para los momentos de μ_i = $\left(\frac{y_i - \beta' x_i}{\sigma}\right)$, el cual consiste en

(6.70) $E\left[\mu_i^{p+1} \middle| y_i = 0\right] = p\sigma^2\left[\mu_i^{p-1} \middle| y_i = 0\right] - \sigma\lambda_i(-\beta' x_i)^p$

Para p $\geq$ 1 t y E[μ_i | y_i = 0]=-$\sigma\lambda_i$ donde λ_i es la razón de Mills. Definiendo el indicador

$I_i = \begin{cases} 1 & Si \quad y_i > 0 \\ 0 & De \quad otra \quad forma \end{cases}$, entonces

(6.71) $\theta_{1i} = E\left(\mu^3{}_i \middle| y_i\right) = I_i \mu^3{}_i - (1 - I_i)\sigma\lambda_i(2\sigma^2 + (\beta' x_i)^2)$
(6.72) $\theta_{2i} = E\left(\mu^4{}_i - 3\sigma^4 \middle| y_i\right) = I_i(\mu^4{}_i - 3\sigma^4) + (1 - I_i)\sigma\lambda_i(\beta' x_i)(3\sigma^2 + (\beta' x_i)^2)$

De esta forma, el contraste de normalidad se basa en θ_{1i} y θ_{2i} [Pagan y Vella(1989) y Skeels y Vella(1993). Para implementar este contraste, se necesitan los residuos, pero éstos no se pueden obtener de (y_i - $\beta' x_i$) pues éstos no tienen media cero. Una forma de corregir este problema, consiste en obtener los residuos generalizados

(6.73) $\hat{\eta}_i = -\hat{\sigma}(1 - I_i)\hat{\lambda}_i + I_i \hat{\mu}_i$

Para el Probit, el valor esperado del error será

(6.74) $E(\mu_i | y_i) = \phi_i (y_i - \Phi_i)\Phi_i^{-1}(1 - \Phi_i)^{-1}$

El residuo generalizado $\hat{\eta}$ para un modelo de elección discreto, viene dado por

(6.75) $\hat{\eta}_i = \hat{f}_i (y_i - \hat{F}_i)\hat{F}_i^{-1}(1 - \hat{F}_i)^{-1}$

Powell (1986) desarrolla un estimador de cuadrados ordinarios censurado simétricamente(SCLS) para un Tobit de la forma siguiente: Se eliminan las observaciones para las cuales $\hat{\beta}' x_i < 0$. A continuación, se hace $y_i = 2\hat{\beta}' x_i$. De esta forma, los y_i son distribuidos simétricamente sobre (0, $2\hat{\beta}' x_i$) y los errores ($y_i - \hat{\beta}' x_i$) son distribuidos simétricamente sobre
$-\hat{\beta}' x_i$ y $\hat{\beta}' x_i$, por lo cual, tendrán media cero.

Newey(1985) parte de una distribución F_i, bajo la hipótesis alternativa de que $F_i = \Phi(\mu_i + \gamma_1 \mu_i^2 + \gamma_2 \mu_i^3)$ donde $\mu_i = \left(\frac{\beta' x_i}{\sigma}\right)$. Las condiciones de momentos resultantes serán

(6.76) $\frac{1}{n}\sum \hat{\mu}_i^2 \hat{\mu}_i = 0$ y $\frac{1}{n}\sum \hat{\mu}_i^3 \hat{\mu}_i = 0$

Bera, Jarque y Lee(1984) construyen el siguiente contraste: Suponga una función de densidad $g(\mu)$ que satisface la ecuación diferencial

(6.77) $\frac{\partial Lng(\mu)}{\partial \mu} = \frac{a + \mu}{b_0 + b_1 \mu + b_2 \mu^2}$

Si μ tiene media cero, esto implica que $a = -b_1$, después de reparametrizar

(6.78) $\frac{\partial Lng(\mu)}{\partial \mu} = \frac{c_1 + \mu}{c_0 + c_1 \mu + c_2 \mu^2}$

De esta forma, si se hace $c_1 = 0, c_2 = 0$ y $c_0 = \sigma^2$ obtendremos una densidad de $N(0, \sigma^2)$, por lo tanto, un contraste de normalidad consiste en contrastar $c_1 = 0, c_2 = 0$. Smith(1989) sugiere usar polinomios ortogonales y construir una contraste Score para los supuestos distribucionales. Bajo condiciones generales, se puede escribir una densidad $h(Z, \theta, \psi)$ de Z como el producto de otra densidad $f((Z, \theta)$ con momentos finitos de todos los ordenes y una serie de polinomios ortogonales $p_k(Z, \theta)$ de esta forma tendríamos

(6.79) $h(Z, \theta, \psi) = f(Z,\theta)\sum_{k=0}^{\infty} a_k(\theta,\psi)\, p_k(Z,\theta)$

Donde $a_0(\theta, \psi) = 1$ y $p_0(Z, \theta) = 1$. La hipótesis de Smith para normalidad consiste en hacer Ho: $a_k(\theta, \psi) = 0$ para $k > 1$.

Newey(1987) considera un contraste de normalidad tipo Hausman de la forma

(6.80) $y^*_i = Z_t\,\delta + \varepsilon_i$ con $y_i = \begin{cases} 1 & Si \quad y^*_i > 0 \\ 0 & De \quad otra \quad forma \end{cases}$

Donde Z_t incluye variables endógenas. El contraste se basa en la diferencia entre un estimador Tobit máximo verosímil (MLE), $\hat{\delta}$, y un estimador SCLS, $\hat{\delta}_s$. El contraste de normalidad de Hausman, será

(6.81) $Ho: n(\hat{\delta}_s - \hat{\delta})'\left[\nu(\hat{\delta}_s - \hat{\delta})\right]^{-1}(\hat{\delta}_s - \hat{\delta})$

Los supuestos distribucionales del contraste, podrían ser afectados por especificaciones erróneas en la ecuación [Blundell y Meghir(1986)]. Por otro lado, cuando existen distorsiones en el tamaño el contraste podría sobrerechazar la hipótesis nula [Newey(1987)]. Algunos autores como Chesher, Lancaster e Irish(1985) proponen usar métodos gráficos para detectar fallas en los supuestos distribucionales. Aunque el procedimiento es menos formal, este procedimiento podría informarnos preliminarmente de problemas de distribución. El procedimiento consiste en tomar los residuos $\hat{\mu}_i$ y computar la función de distribución usando el método de Kaplan-Meier(KPM). La comparación visual de esta función con respecto a la distribución F se hace a través de graficar $F^{-1}[KPM(\hat{\mu}_i)]$ contra $\hat{\mu}_i$ en el caso de una distribución normal $F = \Phi$. Si el modelo es correcto, la gráfica que resulta será continua [Horowitz y Neuman(1989)].

6.9.7 *Contraste de correlación contemporánea*

Considere, el siguiente Probit bivariado

(6.82) $$y_1^* = \beta'_1 x_{1i} + \mu_{1i}$$
$$y_2^* = \beta'_2 x_{2i} + \mu_{2i}$$

Donde y^*_1 y y^*_2 son variables dependientes dicotómicas y $\{\mu_1$ y $\mu_2\}$ son normales bivariadas con media cero varianza unitaria y correlación ρ. El vector Score para este modelo viene dado por

(6.83) $$\frac{\partial L}{\partial \beta_1} = \sum x_{1i} \hat{\mu}_{1i} \, , \; \frac{\partial L}{\partial \beta_2} = \sum x_{2i} \hat{\mu}_{2i} \, , \frac{\partial L}{\partial \rho} = \sum \hat{\mu}_{1i} \hat{\mu}_{2i}$$

Donde $\hat{\mu}_{1i}$ y $\hat{\mu}_{2i}$ son los residuos generalizados para las dos ecuaciones Probit. El contraste estadístico viene dado por nR^2 de una regresión artificial de un vector de unos sobre ($x_1 \hat{\mu}_1$, $x_2 \hat{\mu}_2$, $\hat{\mu}_1 \hat{\mu}_2$). Una alternativa simple, consiste en computar los residuos generalizados, su correlación al cuadrado r^2, y usar nr^2 como una χ^2_1 (chi-cuadrado con un grado).

Kiefer(1982) desarrolla un contraste Score para un Probit multivariado, él comienza con el supuesto general de que la matriz de correlación de los errores es R y desarrolla un contraste Score para la hipótesis R = I. Kiefer, también desarrolla un contraste para la hipótesis $\rho=0$ cuando $R = (1-\rho)I + \rho ee'$ donde e es un vector de unos, éste contraste es bastante conveniente en modelos de efectos con datos de panel.

6.9.8 Contraste de sesgos de selección

El contraste para sesgos de selección, fue el primer contraste de especificación en modelos con variables dependientes limitadas. Este contraste fue desarrollada por Gronau(1974) y Heckman(1979). En términos generales se le conoce como el contraste de Heckman. El problema planteado parte del modelo de autoselección tipo Heckman, de la forma

(6.84) $$y^*_{1i} = \beta'_1 x_{1i} + \mu_{1i}$$

Y, la ecuación de selección es

(6.85) $$y^*_{2i} = \beta'_2 x_{2i} + \mu_{2i} > 0$$

De esta forma, y^*_{1i} es observado e igual a y_{1i} sí y solo sí $y^*_{2i} > 0$. Por lo tanto, y^*_{1i} es censurado por la ecuación de selección. Es también de esperar que μ_{1i} y μ_{2i} tengan un grado de correlación ρ. Dado que y^*_{2i} es observado como una variable dicotómica, normalizando (6.85) a través de Var(μ_{2i}) = 1, deberemos asumir

$$(6.86)\ \begin{pmatrix}\mu_{1i}\\ \mu_{2i}\end{pmatrix} \sim N\left[\begin{pmatrix}0\\0\end{pmatrix}, \begin{pmatrix}\sigma^2 & \rho\sigma\\ \rho\sigma & 1\end{pmatrix}\right]$$

Una estimación por mínimos cuadrados ordinarios de (6.84) da como resultado estimadores sesgados de β_1 dado que E [μ_{1i} | y^*_{2i} > 0] es diferente de cero, esta expresión viene dada por $\rho\sigma\lambda_i$ donde λ_i es la razón de Mills. De esta forma (6.84) en términos de la variable observada y_{1i} puede escribirse como

(6.87) $y_{1i} = \beta'_1 x_{1i} + \rho\sigma\lambda_i + \varepsilon_{1i}$ donde $E[\ \varepsilon_{1i}\ ] = 0$

Una inspección de (6.87) nos muestra la naturaleza del error, esto es, la omisión de λ_i. Dado que λ_i no es observado, Heckman sugiere obtener el estimador preliminar de $\hat{\lambda}_i$ con base en $\hat{\beta}_2$, el Probit estimado de (6.85) y luego estimar (6.87) por mínimos cuadrados ordinarios. Adicionalmente si ρ es igual a cero no existirán sesgos. Melino (1982) muestra que el contraste de significancia de λ_i en (6.87) es un contraste de Score sobre ρ = 0. El modelo puede ser estimado por dos etapas, sin embargo, existen restricciones: Si x_{2i} contiene solamente una constante entonces λ_i es una constante y el coeficiente $\rho\sigma$ no es estimable; Si λ_i es una función lineal de los componentes de x_{1i} entonces se producirá multicolinealidad, esto ocurre cuando x_{2i} contiene solamente variables dummy (falsas) y x_{1i} incluye las mismas variables dummy y sus combinaciones. Usando ML el problema puede resolverse, sin embargo la estimación bajo ML tiene como problema que se podría entrar en un LOOP al buscar convergencia, y en muchas ocasiones podría encontrarse máximos locales y no globales. Recientemente se ha popularizado el estimador ML provisto por el programa LIMDEP para este tipo de Modelos, Nawata(1993a, 1993b) ha demostrado como los estimadores obtenidos por LIMDEP no son confiables. Olsen(1982) muestra que la función de verosimilitud para el modelo de selección tiene un único máximo condicionado sobre ρ. Olsen, sugiere que uno puede obtener estimadores ML condicionados sobre ρ y examinarlos en ρ. El procedimiento ha sido usado por Nawata, mostrando que este procedimiento da resultados más confiables que usar el programa LIMDEP.

6.9.9 Contraste de estabilidad

No es muy común contrastar estabilidad en modelos de variables dependientes limitadas, sin embargo, Anderson(1987) abre el camino en este tipo de contrastes. Anderson propone comparar el logaritmo de la verosimilitud cuando el modelo es regresado sobre un período, con respecto a un período posterior. El trabajo se inspira en el contraste de estabilidad de Chow, extendiéndose el uso de las variables dummy a los modelos Tobit y Probit. Hoffman y Pagan (1989) sugieren, siguiendo a Anderson, definir primero un periodo de 1 hasta s y un periodo de s+1 hasta s+S, y elaborar el estadístico

$$(6.88)\quad \hat{\tau} = \frac{1}{n}\sum_{t=s+1}^{s+S} \hat{d}_t$$

Donde $\hat{d}_t$ son los estimadores Score. La media del Score período posterior deberá ser cercana o igual a cero en el caso de estabilidad en los parámetros. La varianza asintótica de $\hat{\tau}$ será $S(1+k)I_{\theta\theta}$ [Hoffman y Pagan(1989)]donde K = s/S. De esta forma, el contraste estadístico será

$$(6.89)\quad \frac{1}{S(1+k)}\hat{\tau}'\,\hat{I}_{\theta\theta}^{-1}\,\hat{\tau}$$

El cual sigue una distribución chi-cuadrada con s grados de libertad. Este contraste puede ser aplicado a cualquier modelo de variables dependientes limitadas y se estima por máxima verosimilitud.

6.9.10 Contraste de exogeneidad

En modelos de ecuaciones simultáneas que involucran variables dependientes limitadas Groger(1990) considera un contraste de exogeneidad tipo Hausman a través de una estimación de mínimos cuadrados ordinarios no-linelaes. Smith y Blundell(1986) consideran el siguiente modelo

$$(6.90)\quad \begin{matrix} y^*_{1i} = y'_{2i}\gamma_1 + x'_{1i}\beta_1 + \varepsilon_{1i} \\ y'_{2i} = x'_i \pi_2 + \gamma'_{2i} + \varepsilon_{2i} \end{matrix} \;;\; y_{1i} = \begin{cases} y^*_{1i} & Si \quad y^*_{1i} > 0 \\ 0 & De \quad otra \quad forma \end{cases} \text{ y } \begin{pmatrix} \varepsilon_{1i} \\ \varepsilon_{2i} \end{pmatrix} \sim IN\left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sum_{22} \end{pmatrix}\right]$$

Donde y^*_{1i} no es observado. La hipótesis para probar exogeneidad consiste en Ho: $\sigma_{12} = 0$. Smith y Blundell sugieren

- Suponga que $\varepsilon_{1i} = \nu'_{2i}\lambda + e_{1i}$ y sustitúyalo en (6.90)
- Denote $Var(e_{1i}) = \sigma_{12}$
- Estime $y^*_{1i} = y'_{2i}\,\gamma_1 + x'_{1i}\beta_1 + \nu'_{2i}\,\alpha + E_{1i}$ como un Probit y pruebe la hipótesis Ho: $\alpha = 0$. Smith y Blundell a este estimador, el estimador condicional ML.

  Rivers y Voung(1988) consideran el mismo modelo que Blundell y Smith pero en un contexto

Probit y Vella y Verbeek(1993) lo consideran para el caso de un modelo de panel.

6.10 Variables latentes

Las variables latentes representan conceptos unidimensionales en su más pura forma, puede decirse, que se trata de variables abstractas como inteligencia, paisaje, etc. Como todas las variables latentes corresponden a conceptos, ellas son variables hipotéticas que varían en su grado de abstracción: inteligencia, clase social, poder y expectativas son variables latentes abstractas creadas en la teoría. Variables menos abstractas son la educación y el tamaño de la población.

Un ejemplo es la hipótesis de Emile Durkheim, sobre la relación inversa entre la cohesión social y el suicidio, la cohesión social se refiere a la solidaridad de grupo, la cual es una variable abstracta, el suicidio es directamente observable, pero la relación directa - indirecta es muy débil de acuerdo a la misma clasificación de los suicidios.

Un modelo latente se acompaña de un conjunto de ecuaciones estructurales que resumen las relaciones entre las variables latentes. Bollen(1989) usa las relaciones entre la democracia política y la industrialización, en países desarrollados, para introducir la noción de modelos de variables latentes. Dado que algunas sociedades han alternado entre dictaduras y regímenes electorales es difícil discernir si la asociación realmente existe. La democracia política se refiere a la extensión de los derechos políticos (imparcialidad de las elecciones) y libertades políticas (libertad de prensa) en un país. La industrialización es el grado en el cual la economía de una sociedad se caracteriza por el proceso de manufactura mecanizado, esto implica riqueza social, población educada, avances en el estándar de vida, y éstas son las oportunidades de una democracia.

Suponga que se tienen tres variables latentes aleatorias: democracia política en 1965 y 1960 e industrialización en 1960. Uno podría asumir que la democracia política en 1965 es una función de la democracia política e industrialización de 1960. No existe nada que nos diga que el nivel de industrialización es una variable latente exógena (independiente) y se simboliza como ξ_1, esta es exógena, en tanto sus causas están por fuera del modelo. La variable democracia política es una variable latente endógena, ella esta determinada por variables en el modelo, cada variable latente es representada por η_i. De esta forma, la democracia política en 1960 es representada por η_1 y la democracia política en 1965 por η_2, las variables latentes endógenas son parcialmente explicadas en el modelo y el componente no explicado γ_i es un término aleatorio, de esta forma, el modelo de variables latentes para el ejemplo será

$$(6.91)\quad \begin{aligned} \eta_1 &= \gamma_{11}\xi_1 + \gamma_1 \\ \eta_2 &= B_{21}\eta_1 + \gamma_{21}\xi_1 + \gamma_2 \end{aligned}$$

Escribiéndolo en notación matricial

$$(6.91.1)\begin{bmatrix}\eta_1\\ \eta_2\end{bmatrix} = \begin{bmatrix} 0 & 0 \\ B_{21} & 0\end{bmatrix}\begin{bmatrix}\eta_1\\ \eta_2\end{bmatrix} + \begin{bmatrix}\gamma_{11}\\ \gamma_{21}\end{bmatrix}\left[\xi_1\right] + \begin{bmatrix}\gamma_1\\ \gamma_2\end{bmatrix}$$

Compactamente, tenemos

$$\eta = B\eta + \gamma\xi + \gamma \tag{6.92}$$

Además, E(η) = 0, E(ξ) =0, E(γ) = 0 y γ no está correlacionado con ξ. Por otro lado, (I-B) es no singular. De esta forma (6.92) es una matriz general que representa las ecuaciones estructurales para un modelo de variable latentes.

6.10.1 Ecuaciones estructurales con variables observadas

La ecuación (6.92) es una representación general de ecuaciones estructurales con variables observadas de la forma

$$y = By + \Gamma x + \gamma \tag{6.93}$$

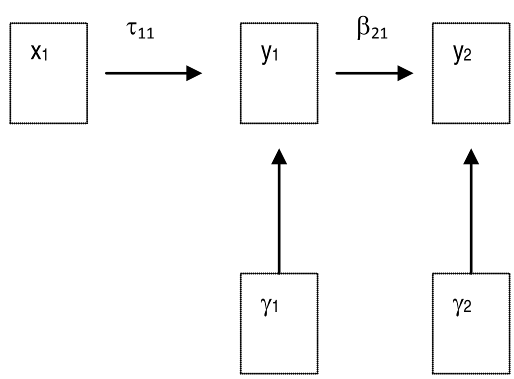
Donde, y es un vector de variables endógenas de $m\times1$, B es una matriz de coeficientes de $m\times m$, Γ es una matriz de coeficientes de $m\times n$, x es un vector de variables exógenas $n\times1$ y γ es un vector de errores en las ecuaciones de $m\times1$. Los γ representan los errores aleatorios en las relaciones entre los y's y los x's. El supuesto estándar consiste en que los errores γ no están no-correlacionados con x. Por otro lado, E(γ^2_{ik}) = Var γ_i $\forall$ k y Cov(γ_{ik} , γ_{ik}) = 0 $\forall$ k $\neq$ i.

El modelo implícito para las ecuaciones estructurales con variables observadas, será

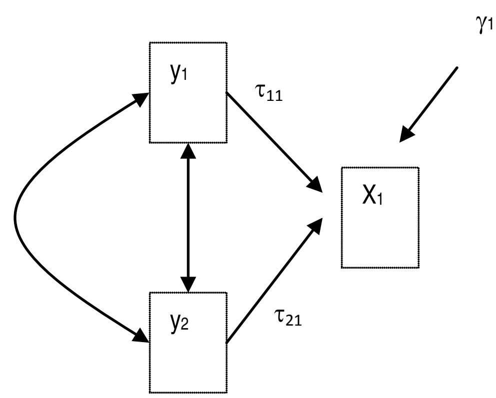
$$\begin{aligned} y &= \eta \quad donde\ y = m\times 1\ vector\ de\ \text{var}iables\ observadas \\ x &= \xi \quad donde\ x = n\times 1\ vector\ de\ \text{var}iables\ observadas \end{aligned} \tag{6.93.1}$$

Los modelos recursivos son sistemas de ecuaciones que no obtiene casualidad no recíproca como en 6.3.a o retroalimentación como en 6.3.b. Cuando esto es cierto es posible denotar a B como una matriz triangular inferior, en adición la matriz de covarianzas de los errores en las ecuaciones (Ψ) será diagonal, esto significa, que los errores para una ecuación no están correlacionados con los errores de las otras ecuaciones. Por ejemplo, si y_1 causa y_2, y_2 no tiene efecto sobre y_1 ni directamente, ni a través de alguna cadena de otras variables, veamos:

Gráfica 6.3. a. Causalidad unidireccional

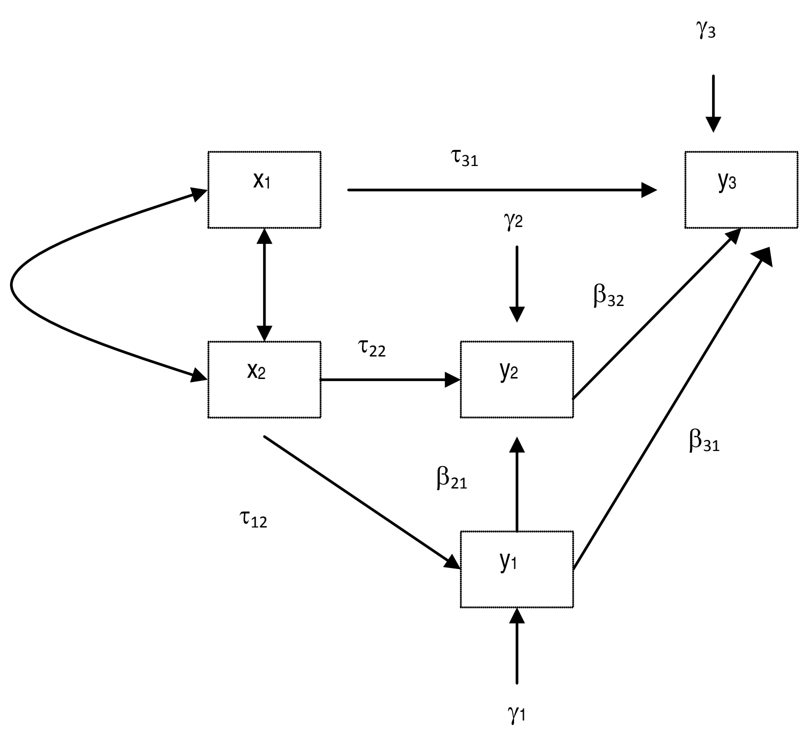


Gráfica 6.3. b. Retroalimentación



McDonald y Clelland (1984) proponen el siguiente modelo:

Gráfica 6.4.Retroalimentación en un modelo de sindicatos



Este modelo representa el deseo de unión de los trabajadores textiles en el suroeste de los Estados Unidos. Donde x_1 representa los años en la fábrica de textiles, x_2 representa la edad, y_1 la sumisión al empresario, y_2 el apoyo a las actividades de los trabajadores y, y_3 la propensión hacia las uniones.

El ordenamiento causal de McDonald y Clelland, consiste en que la sumisión influencia la actitud hacia el activismo y_2 y las uniones $y_{3,}$ y el activismo afecta el sentimiento de unión. Por otro lado, los γ's errores no están correlacionados a través de las ecuaciones. La ecuación en términos matriciales será

(6.94)
$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ \beta_{21} & 0 & 0 \\ \beta_{31} & \beta_{32} & 0 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} + \begin{bmatrix} 0 & \tau_{12} \\ 0 & \tau_{22} \\ \tau_{31} & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} \gamma_1 \\ \gamma_2 \\ \gamma_3 \end{bmatrix}$$

6.10.2 La Matriz de Covarianzas

La hipótesis del modelo de ecuación estructural general, consiste en

(6.95)
$$\sum = \sum(\theta)$$

Donde Σ es la matriz de covarianzas de la población de y e x, y $\Sigma(\theta)$ es la matriz de covarianzas como función de un modelo de parámetros libres en θ. La ecuación (6.92) implica que cada elemento de la matriz de covarianzas es una función de uno o más parámetros del modelo. La relación entre Σ y $\Sigma(\theta)$ es fundamental para entender la identificación y estimación del modelo ajustado.

La matriz de covarianzas $\Sigma(\theta)$ reúne los siguientes elementos: Primero, la matriz de covarianzas de y. Segundo, la matriz de covarianzas de x con y. Tercero, la matriz de covarianzas de x. Considérese primero $\Sigma_{yy}(\theta)$ esto es la matriz de covarianzas de y

$$
\begin{aligned}
\sum\nolimits_{yy}(\theta) &= E(yy') \\
&= E\left[(I-B)^{-1}(\Gamma_x+\gamma)\left((I-B)^{-1}(\Gamma_x+\gamma)'\right)\right] \\
&= E\left[(I-B)^{-1}(\Gamma_x+\gamma)(x'\Gamma'+\gamma')(I-B)^{-1'}\right] \\
&= (I-B)^{-1}\left[E(\Gamma_{xx'}\Gamma')+E(\Gamma_x\gamma')+E(\gamma x\Gamma')+E(\gamma\gamma')\right](I-B)^{-1'} \\
&= (I-B)^{-1}(\Gamma\Phi\Gamma'+\Psi)(I-B)^{-1'}
\end{aligned}
\tag{6.95.1}
$$

Donde Φ es la matriz de covarianzas de x, Ψ es la matriz de covarianzas de y. La matriz de covarianza de x, $\Sigma_{xx}(\theta)$, es igual a Φ, esto es

$$\sum\nolimits_{xx}(\theta) = E(xx') = \Phi \tag{6.95.2}$$

La parte final de la matriz de covarianzas es $\Sigma_{xy}(\theta)$, esto es, la covarianza de x con y

$$
\begin{aligned}
\sum\nolimits_{xy}(\theta) &= E(xy') \\
&= E\left[x\left((I-B)^{-1}(\Gamma x+\gamma)\right)'\right] \\
&= \Phi\Gamma'(I-B)^{-1'}
\end{aligned}
\tag{6.95.3}
$$

De esta forma, encontramos que $\Sigma(\theta)$ será

$$\sum(\theta) = \begin{bmatrix} (I-B)^{-1}(\Gamma\Phi\Gamma'+\Psi)(I-B)^{-1'} & (I-B)^{-1}\Gamma\Phi \\ \Phi\Gamma'(I-B)^{-1'} & \Phi \end{bmatrix} \tag{6.96}$$

6.10.3 Identificación

La identificación del modelo (6.96) con una o más ecuaciones, requiere una investigación de cuáles parámetros son conocidos y desconocidos. Por parámetros conocidos, entiéndase aquellos que pueden ser identificados, estos parámetros generalmente son características de la población y de la distribución de las variables observadas como las varianzas y covarianzas para los cuales los estimadores de la muestra son consistentes. Los parámetros desconocidos, son aquellos parámetros cuyo estatus de identificación no es conocido, estableciendo entonces el investigador cuando existen valores únicos para estos.

El modelo, se encuentra identificado, si se muestra que los parámetros desconocidos son funciones solamente de los parámetros identificados y, que estas funciones, llevan a soluciones únicas. Suponga que la Var(y) es el parámetro identificado, θ_1 y θ_2 son parámetros desconocidos y, la ecuación que los relaciona es Var(y) = $\theta_1 + \theta_2$. La identificación, deberá establecer cuando se alcanzan valores únicos de θ_1 y θ_2 en esta ecuación. Claramente con dos parámetros desconocidos θ_1, θ_2 en una sola ecuación la identificación no es posible. Para algún valor dado de Var(y) un conjunto finito de valores de θ_1 y θ_2 satisfacen dicha ecuación. Sin embargo adicionando una segunda ecuación $\theta_1 = \theta_2$ se puede observar la identificación para la cual cada parámetro será igual a $\frac{Var(y)}{2}$.

Este principio general, deberá mantenerse para ecuaciones estructurales más complicadas. Los parámetros cuya identificación es desconocida están en θ y θ contiene las (t) libres y los parámetros restringidos(no-redundantes) de B, Γ, Φ y Ψ. La ecuación que relaciona Σ a θ es la hipótesis de la estructura de la covarianza $\Sigma = \Sigma(\theta)$. Este principio general, lo enuncia Bollen de la siguiente forma: "Si un parámetro desconocido en θ se escribe como una función de uno o más elementos de Σ, este parámetro es identificado, si todos los parámetros desconocidos en θ son identificados, entonces el modelo es identificado" (pag 89).

Un conjunto más apropiado de identificación es conocido como la regla t, la regla del B nulo, la regla recursiva y las condiciones de rango y orden.

6.10.3.1 Regla t

Esta es la condición más sencilla, pero no es una condición suficiente. La regla t, parte de que el número de elementos no-redundantes en la matriz de covarianzas de las variables observadas deberá ser mayor o igual que el número de parámetros desconocidos en θ, esto es

(6.97) $$t \leq \left(\tfrac{1}{2}\right)(p+q)(p+q+1)$$

Donde, p+q , es el número de variables observadas y t es el número de parámetros libres en θ.

6.10.3.2 Regla del B nulo

Es un modelo multiecuacional donde las variables que no son endógenas afectan a alguna variable endógena, la matriz B es cero. Veamos

$$\begin{aligned} y_1 &= \tau_{11}x_1 + \tau_{12}x_2 + \gamma_1 \\ y_2 &= \tau_{21}x_1 + \tau_{22}x_3 + \gamma_2 \\ \operatorname{cov}\left(x_i, \gamma_j\right) &= 0 \ \forall\, i = 1,2,3\,;\ j = 1,2 \end{aligned} \tag{6.98}$$

La matriz B es cero dado que y_1 no afecta a y_2, ni y_2 afecta y_1. De esta forma, se establece que la identificación de algún modelo donde B es cero, los parámetros desconocidos en Γ, Φ y Ψ son funciones de los parámetros identificados de Σ, sustituyendo B=0 en (6.96) y particionando Σ obtenemos

$$\begin{bmatrix} \sum_{yy} & \sum_{yx} \\ \sum_{xy} & \sum_{xx} \end{bmatrix} = \begin{bmatrix} \Gamma\Phi\Gamma'\Psi & \Gamma\Phi \\ \Phi\Gamma' & \Phi \end{bmatrix} \tag{6.99}$$

Como puede observarse $\Phi = \Sigma_{xx}$, por lo cual Φ es identificado. Por otro lado

$$\Phi\Gamma' = \sum_{xy} = \sum_{xx}\Gamma' = \Gamma' = \sum_{xx}^{-1}\sum_{xy} \tag{6.100}$$

De esta forma, Γ es una función conocida, que se identifica a partir de las matrices de covarianzas y que en sí misma es identificada. También podemos observar que

$$\begin{aligned} \Psi &= \sum_{yy} - \Gamma\Phi\Gamma' \\ &= \sum_{yy} - \sum_{yx}\sum_{xx}^{-1}\sum_{xx}\sum_{xx}^{-1}\sum_{xy} \\ &= \sum_{yy} - \sum_{yx}\sum_{xx}^{-1}\sum_{xy} \end{aligned} \tag{6.101}$$

Entonces, cuando B=0, Γ, Φ y Ψ, cada una puede escribirse una como función identificada de las matrices de covarianzas de las variables observadas. Si los errores de una ecuación no están correlacionados con aquellos de las otras ecuaciones es un sistema (Ψ es diagonal). Por lo tanto, esas ecuaciones pueden tratarse como separadas o no relacionadas. Si Ψ no es diagonal y los errores de las últimas dos ecuaciones están correlacionadas entonces, tal modelo será llamado “Seemingly unrelated regresions” Kamenta (1980). La regla B nula es una condición suficiente para identificar un modelo.

6.10.3.3 Regla recursiva

A diferencia de la regla anterior, la regla recursiva no requiere que B=0, para aplicar la regla recursiva B deberá ser una matriz triangular, y Ψ deberá ser diagonal. Una condición más exacta para B, consiste en que ésta deberá ser una matriz triangular inferior. Si ambas condiciones se mantienen el modelo es identificado

$$
\begin{aligned}
y_1 &= \tau_{12}x_2 + \gamma_1 \\
y_2 &= \beta_{21}y_1 + \tau_{22}x_2 + \gamma_2 \\
y_3 &= \beta_{31}y_1 + \beta_{32}y_2 + \tau_{31}x_1 + \gamma_3
\end{aligned}
\tag{6.102}
$$

Una propiedad para todos los modelos recursivos, consiste, en que para una ecuación dada, el término de error γ no este correlacionado con las variables explicatorias. De esta forma,
cov (x_2, γ_2) = 0, cov (x_3, γ_1) = 0 y cov (x_2, γ_3) = 0 y, de igual forma, para cov (γ_2, y_1)= cov (γ_2,$\gamma_2 x_2+\gamma_1$) = 0. Así, γ_2 no esta correlacionado con y_1 y x_2 las dos variables explicatorias de la segunda ecuación y de igual forma cov (γ_3, y_1)=0 y cov (γ_3, y_2)=0. En general, para la i- ésima ecuación en algún modelo recursivo, γ_i no esta correlacionado con las variables endógenas las cuales son variables explicatorias en esa ecuación, esto se debe a que las variables endógenas están en función de las variables exógenas y de los errores de las otras ecuaciones, los cuales no están correlacionados con γ_i(Bollen,pag 96).

6.10.3.4 Condiciones de rango y orden

Si una condición de restricción en una ecuación se determina a partir de las variables excluidas, entonces “una condición necesaria para que una ecuación sea identificada, consiste en que el número de variables excluidas de la ecuación sea al menos p-1 ”. Considere el modelo

$$y_i = \left[B'_i \,|\, \gamma'_i\right] Z_i + \gamma_i \tag{6.103}$$

Multiplicando ambos lados por N y tomando el valor esperado

(6.104) $\sigma'_{y_i x} = \left[B'_i \,|\gamma'_i\right]\Sigma_{Z_i x}$

Sí B'_i y γ'_i son funciones solamente de los elementos de las covarianzas de σ'_{yix} y Σ_{Zix}, ellos son identificados. Una condición necesaria, consiste en que el número de ecuaciones en (6.104) sea al menos igual al número de parámetros libres desconocidos en $\left[B'_i \,|\gamma'_i\right]$. El número de ecuaciones es el número de elementos en σ'_{yix} , esto es, q. De esto se sigue, que q covarianzas con y_i resultan de q variables en x. El número de parámetros desconocidos en B'_i es (p-1) y en γ'_i es q. De esto se deduce, que con q ecuaciones en (p-1) + q desconocidos B'_i y γ'_i no pueden ser identificados, entonces los (p-1) + q desconocidos deberán ser reducidas a q.

La condición de orden se define propiamente como "si las variables excluidas son solamente el tipo de restricciones, entonces las (p-1) variables deberán ser excluidas de la i- ésima ecuación para hacer una posible identificación" Bollen (p 99). Sea

(6.105) $$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} 0 & \beta_{12} & 0 \\ \beta_{21} & 0 & 0 \\ \beta_{31} & \beta_{32} & 0 \end{bmatrix}\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} + \begin{bmatrix} \tau_{11} & 0 \\ 0 & \tau_{22} \\ 0 & 0 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} \gamma_1 \\ \gamma_2 \\ \gamma_3 \end{bmatrix}$$

Suponga que en la primera ecuación β_{13} y γ_{12} no sean restringida a cero, tendrían entonces

(6.105.1) $$y_1 = B_{12}y_2 + B_{13}y_3 + \tau_{11}x_1 + \tau_{12}x_2 + \gamma_1$$

El cual tiene la misma forma que (6.103). Multiplicando ambos lados por las variables exógenas y tomando valores esperados

(6.105.2) $$\begin{aligned} \operatorname{cov}(y_1, x_1) &= \beta_{12}\operatorname{cov}(y_2, x_1) + \beta_{13}\operatorname{cov}(y_3, x_1) + \gamma_{11}\operatorname{var}(x_1) + \gamma_{12}\operatorname{cov}(x_2, x_1) \\ \operatorname{cov}(y_1, x_1) &= \beta_{12}\operatorname{cov}(y_2, x_2) + \beta_{13}\operatorname{cov}(y_3, x_1) + \gamma_{12}\operatorname{var}(x_2) + \gamma_{11}\operatorname{cov}(x_1, x_2) \end{aligned}$$

El resultado son dos ecuaciones para cuatro desconocidas (B_{12}, B_{13}, γ_{11}, γ_{12}) y es claro entonces que una única solución no es posible, la condición de orden requiere (p-1) o 2 exclusiones de esta primera ecuación. La especificación de β_{13}=0 y γ_{12}=0 satisface este requerimiento, así que a partir de la primera ecuación, se encuentra una condición suficiente para la identificación.

Una forma para revisar la condición de orden para todas las ecuaciones en el modelo es formar una matriz, digamos C, la cual es [(1-B)| -Γ]. Así, para cada familia, se cuenta el número de cero elementos, si una familia tiene (p-1) o más ceros, esta es una condición de orden. Para el ejemplo anterior, tendremos

$$(6.105.3)\ C = \begin{bmatrix} 1 & -\beta_{12} & 0 & -\tau_{11} & 0 \\ -\beta_{21} & 1 & 0 & 0 & -\tau_{22} \\ -\beta_{31} & -\beta_{32} & 1 & 0 & 0 \end{bmatrix}$$

Cada fila en (6.105.3) tiene (p-1) o 2 exclusiones así que se satisface la condición de orden. De igual forma, la condición de orden nos brinda una regla para detectar sobreidentificación para modelos no recursivos. Con Ψ libres de sobreidentificación, esta podrá ocurrir si es posible producir una nueva ecuación con la misma forma pero con parámetros diferentes de una ecuación anterior, a través de usar una combinación lineal de las otras ecuaciones en un modelo, esto ocurre cuando dos o más ecuaciones tienen restricciones idénticas

$$(6.106)\ \begin{aligned} y_1 &= \beta_{12} y_2 + \tau_{11} x_1 + \tau_{12} x_2 + \gamma_1 \\ y_2 &= \beta_{21} y_1 + \tau_{22} x_2 + \gamma_2 \end{aligned}$$

Suponga que se excluye x_2 de ambas ecuaciones ($\tau_{12}=\tau_{22}=0$) entonces la matriz C será

$$(6.106.1)\ \begin{bmatrix} 1 & -\beta_{12} & -\tau_{11} & 0 \\ -\beta_{21} & 1 & 0 & 0 \end{bmatrix}$$

La condición de orden para ambas ecuaciones, consiste en hallar si cada ecuación tiene una exclusión (p=1). Multiplicando la segunda fila de (6.106.1) por a y al sumar este resultado a la primera fila se obtiene

$$(6.106.2)\ \begin{bmatrix} 1-\beta_{21}a & -\beta_{12}+a & -\tau_{11} & 0 \\ -\beta_{21} & 1 & 0 & 0 \end{bmatrix}$$

Dividiendo cada elemento de la primera fila por ($1-a\beta_{21}$) se obtiene

$$(6.106.3)\ C^* = \begin{bmatrix} 1 & -\beta_{12}^* & -\tau_{11}^* & 0 \\ -\beta_{21} & 1 & 0 & 0 \end{bmatrix}$$

Donde $\beta_{21}^* = \left(\dfrac{-\beta_{12}+a}{1-a\beta_{21}}\right)$ y $-\tau_{11}^* = \dfrac{-\tau_{11}}{\left(1-a\beta_{21}\right)}$. La primera ecuación representada por la primera fila de C* en (6.106.3) tiene la misma forma y las mismas variables excluidas que C en (6.105.3). De esta forma, existirá un infinito conjunto de valores para

β^*_{12} y de τ^*_{11} pero que no son iguales a B_{12} y τ_{11}, es por esta razón que B_{12} y τ_{11} no son identificados incluso cuando se satisface la condición de orden. Este procedimiento es sencillo con dos ecuaciones, pero en sistemas complejos multiecuacionales no es tan viable. Para determinar la regla del rango con C { [(I-B)| $-\Gamma$] } revise la identificación para la i- ésima ecuación, borre todas las columnas de C que no tengan ceros en la i- ésima fila de C y use las columnas que quedan para construir una nueva matriz C. Y de esta forma, una condición suficiente y necesaria para la identificación de la i- ésima ecuación consistirá en que el rango de C_i sea igual a p-1. Esta es la condición de rango de Bollen(pag.101).

A manera de ilustración considere (6.105) y la C matriz es (6.105.3) y examinemos la identificación de la primera ecuación, solamente existen ahora en la primera fila en la cuarta columna, de esta forma se borran las 3 columnas y C_i queda como

(6.106.3.1) $C_1 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$

Dado que el rango de una matriz o vector es el número de filas independientes o columnas, con ambos elementos de C_1 cero, su rango será cero, con un rango menor que uno, la primera ecuación no es identificable como se demostró anteriormente. Para la segunda ecuación de C en (6.106.3) C_2 será

(6.106.3.2) $C_2 = \begin{bmatrix} -\gamma_{11} & 0 \\ 0 & 0 \end{bmatrix}$

Excepto cuando γ_{11} es cero, el rango de C_2 será uno, lo cual satisface la condición de rango y entonces la segunda ecuación es identificada.

Haavelmo (1953) desarrolla un modelo de propensión marginal a consumir, esto es, la parte del ingreso disponible para comprar bienes de consumo. De esta forma, si este ingreso es igual a los gastos en inversión (x_1), a un gasto en consumo y_2 y los gastos en consumo están en función del ingreso disponible y_1 más un término de error γ, el sistema de dos ecuaciones será

(6.107) $$\begin{aligned} y_1 &= y_2 + x_1 \\ y_2 &= \beta_{21} y_1 + \gamma_2 \end{aligned}$$

En términos matriciales

(6.107.1) $$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ \beta_{21} & 0 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 & x_2 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} \begin{bmatrix} 0 & \gamma_2 \end{bmatrix}$$

La matriz C será

(6.107.2) $$C = \begin{bmatrix} 1 & -\beta_{12} & -\tau_{11} & 0 \\ -\beta_{21} & 1 & 0 & 0 \end{bmatrix}$$

La matriz, C_2 será

(6.107.3) $$C_2 = \begin{bmatrix} -1 \\ 0 \end{bmatrix}$$

El rango es uno, lo cual satisface la condición de rango, por lo tanto, es identificada.

6.10.4 6.10.3.5. Resumen de las reglas de identificación

Las reglas de identificación para ecuaciones estructurales con variables observadas asumiendo que no existen errores de medición, esto es $y = By + \Gamma_x + \gamma$ será

Reglas de identificación	**Evalúa**	**Requisitos**	**Condición Necesaria**	**Condición Suficiente**
t-regla	Modelo	$t \leq (1/2)(p+q)(p+q+1)$	Sí	No
Regla B nulo	Modelo	B = 0	No	Sí
Regla recursiva	Modelo	B triangular, Ψ diagonal	No	Sí
Condición de Orden	Ecuación	Restricción $\geq$ p-1, Ψ libre	Sí	No
Condición de rango	Ecuación	Rango C_i = p-1, Ψ libre	Sí	Sí

6.10.5 Estimación

El procedimiento de estimación, se deriva de la relación de la matriz de covarianzas de las variables observadas a los parámetros estructurales. De esta forma

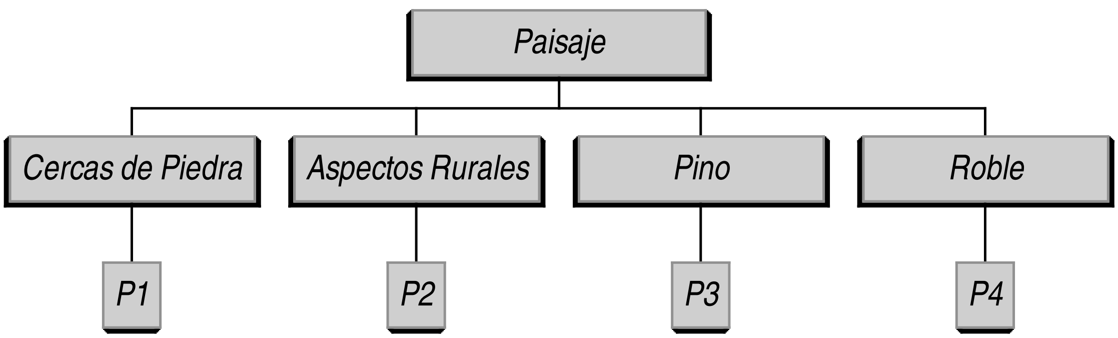
(6.108) $$\sum(\theta) = \begin{bmatrix} (I-B)^{-1}(\Gamma\Phi\Gamma'+\Psi)(I-B)^{-1'} & (I-B)^{-1}\Gamma\Phi \\ \Phi\Gamma'(I-B)^{-1'} & \Phi \end{bmatrix}$$

Si el modelo de ecuaciones estructurales es correcto y, los parámetros de la población son conocidos, entonces Σ podrá igualarse a $\Sigma(\theta)$. La forma de estimar el modelo de ecuaciones estructurales se realiza a través de una función de máximo verosimilitud, donde dicha función se ajusta maximizando

(6.109) $F_{ML} = Log|\Sigma(\theta)| + traza\left(S\,\Sigma^{-1}(\theta)\right) - LogS - p + q$

Donde S es la matriz de covarianzas para y_i y x_i. A través, de este procedimiento, se obtendrán estimaciones maximoverosimiles. Mora (1997), supone el siguiente modelo: Supóngase que el paisaje rural es una variable latente. Debido a que existen diferentes características que determinan un paisaje supondremos que este tiene cuatro indicadores principales, como se observa en la siguiente gráfica

Modelo de Variables Latentes



Con
$$\begin{aligned} P_1 &= \lambda_{11}\xi_1 + \delta_1 \\ P_2 &= \lambda_{21}\xi_1 + \delta_2 \\ P_3 &= \lambda_{31}\xi_1 + \delta_3 \\ P_4 &= \lambda_{41}\xi_1 + \delta_4 \end{aligned}$$

(6.110)
$$p_i^{\,*} = \Lambda_p \xi_p + \delta;$$
$$\varepsilon(\delta_i) = 0; \operatorname{cov}(\xi_i, \lambda_i) = 0 \forall : i..n; \delta_{iid}\,(0, \sigma_{jj}^{\;2})$$

Donde el paisaje $P^{*}{}_i$ es la variable latente y ξ_i es el verdadero paisaje. El P indicador del paisaje en $P^{*}{}_i$ sirve como indicador de la variable latente ξ_i, el verdadero paisaje. Sin pérdida de generalidad, si el indicador es centrado alrededor de cero, de tal forma que $P^{*}{}_i$ tiene un valor extremo, con los parámetros consistentes de los elementos de

λ, la varianza de P^*_i, Φ, y la varianza del error. Tomando segundos momentos a ambos lados de la ecuación (6.110)

(6.111) $\sum(\theta) = E(PP') = E(\Lambda_p P + \delta)(P'\Lambda_p + \delta_l) = \Lambda_p \Phi \Lambda'_p + \Theta_\delta$

(6.112) Cov(P^*_i) = $\Lambda_p \Phi \Lambda'_p + \Theta_\delta = \sum(\theta)$

Donde $\Theta_\delta = \begin{bmatrix} \delta_1 & 0 & 0 & 0 \\ 0 & \delta_2 & 0 & 0 \\ 0 & 0 & \delta_3 & 0 \\ 0 & 0 & 0 & \delta_4 \end{bmatrix}; \overline{\lambda} = \begin{bmatrix} \lambda_{11} \\ \lambda_{21} \\ \lambda_{31} \\ \lambda_{41} \end{bmatrix}; \Phi = (\Phi_{11})$

Siguiendo a Bollen (1989), el modelo anterior puede ser identificado siempre que t $\leq$ 1/2(p)(p+1), es decir, t <10 factores desconocidos. A través de (6.111) se obtiene

(6.113) $$\sum(\theta) = \begin{bmatrix} \lambda_{11}{}^2\Phi_{11} + Var\delta_1 & & & \\ \lambda_{21}\lambda_{11}\Phi_{11} & \lambda_{21}{}^2\Phi_{11} + Var\delta_2 & & \\ \lambda_{31}\lambda_{11}\Phi_{11} & \lambda_{31}\lambda_{21}\Phi_{11} & \lambda_{31}{}^2\Phi_{11} + Var\delta_3 & \\ \lambda_{41}\lambda_{11}\Phi_{11} & \lambda_{41}\lambda_{21}\Phi_{11} & \lambda_{41}\lambda_{31}\Phi_{11} & \lambda_{41}{}^2\Phi_{11} + Var\delta_4 \end{bmatrix}$$

En (6.113) existen 10 elementos desconocidos, escribiendo cada elemento en la matriz de covarianzas P1, P2, P3, P4 en términos de sus correspondientes parámetros estructurales y haciendo $\lambda_{11} = 1$ para escalar ξ_1 y poder identificar $\sum(\theta)$, se encuentran 9 ecuaciones con 9 elementos desconocidos lo que lleva a una solución para los parámetros desconocidos. Con este resultado, se calcula (6.113). De Saris (1978) se conoce además

(6.114) $P_i^* = \Phi\Lambda_p{}' \sum^{-1} P_i$

Se podrá observar además que $P_{ji(j=1..k)}$ es un indicador imperfecto de P^*_i si $\sigma_{jj} \rangle 0$ y si $\sigma_{jj} = 0$ $P_{ji(j=1..k)}$ es un indicador perfecto de P^*_i. Entonces, obtenemos (6.114) como una aproximación del paisaje y su función de verosimilitud vendrá dada por

(6.115) $likelh = \left[\left|\sum(\theta)\right|\right]^{-\frac{N}{2}} e^{-\frac{N}{2} trace\left[S_{WW}\sum(\theta)^{-1}\right]}$

Donde $S_{ww} = \frac{P^{'}P}{N}$. Para estimar (6.115) se requiere estimar (6.113) a través de máxima verosimilitud, los resultados encontrados usando una encuesta sobre paisaje que se describe en el capítulo 9 fueron

$$\lambda(i) = \begin{bmatrix} 1 \\ 22{,}8316 \\ 0{,}467748 \\ 1{,}596955 \end{bmatrix} \quad \Phi_{11} = [-0.00026] \quad \Theta_{\delta} = \begin{bmatrix} 0.01349 & & & \\ 0 & 0.152026 & & \\ 0 & 0 & 0.012205 & \\ 0 & 0 & 0 & 0.019107 \end{bmatrix}$$

$$\sum(\theta) = \begin{bmatrix} 0.013288 & & & \\ -0.00596 & 0.015987 & & \\ -0.00012 & -0.00279 & 0.012148 & \\ -0.00042 & -0.00952 & -0.00019 & 0.018442 \end{bmatrix}$$

Bollen (1989), provee el siguiente contraste de sobreidentificación

(6.116) FML $(s\hat{\sum}(\theta) = -Ln\left|\sum(\theta)\right| + Tr(s\sum^{-1}(\theta) - Log|s| - q$

Y, FML$\approx \chi^{2}_{q,gl}$ Donde Ho=$\sum = \sum(\theta)$ MAICE= Min AICH. El resultado encontrado fue FML =0.389563, dado este valor, bajo H_0 no existe sobreidentificación, esto es, no existe suficiente evidencia estadística para rechazar $H_{0.}$

6.11 Estimación en STATA

6.11.1 Modelos Binomiales

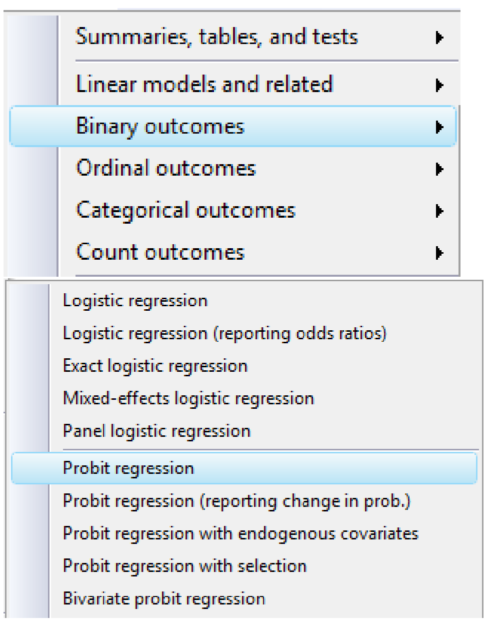
Los datos son tomados de Mora (2008) y explican la probabilidad de permanecer en un área de desempeño profesional. Perm es una variable dummy que toma valor de 1 si el individuo permanece en la misma área de desempeño y 0 en caso contrario. Sexo es una variable dicotómica que toma valor de 1 para las mujeres. La variable Exper es la experiencia en semanas reportada por el individuo. La variable SO1 indica la existencia de sobre-educación y toma el valor de 1 si el individuo tiene nivel profesional y acepta un trabajo que exige nivel ocupacional de tecnólogo o técnico. La variable SO2 es una variable dummy que toma valor de 1 si tiene un nivel profesional y acepta un trabajo que exige educación media. De esta forma, SO1 y SO2 muestran la magnitud del nivel de sobre-educación. La variable posg es una variable dummy que toma valor de 1 si el individuo tiene título de posgrado y 0 si no lo tiene. La variable No oficial es una variable dummy que toma valor de 1 si el título del individuo proviene de una institución no oficial. La estimación MPL se reporta a continuación:

```
. reg perman sexo exp SO1 SO2  posg nooficial

      Source |       SS       df       MS              Number of obs =    8229
-------------+------------------------------           F(  6,  8222) =  117.89
       Model |   141.00239     6  23.5003983           Prob > F      =  0.0000
    Residual |  1638.98303  8222  .199341161           R-squared     =  0.0792
-------------+------------------------------           Adj R-squared =  0.0785
       Total |  1779.98542  8228  .216332695           Root MSE      =  .44648

------------------------------------------------------------------------------
      perman |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        sexo |   .0869833   .0099006     8.79   0.000     .0675756     .106391
         exp |   .0086062   .0010564     8.15   0.000     .0065353     .010677
         SO1 |  -.1533382   .0158636    -9.67   0.000    -.1844349   -.1222415
         SO2 |  -.4245631   .0198123   -21.43   0.000    -.4634002    -.385726
        posg |  -.0313634   .0156514    -2.00   0.045    -.0620441   -.0006826
   nooficial |   .0471142   .0099284     4.75   0.000     .0276519    .0665764
       _cons |   .6417989   .0094671    67.79   0.000      .623241    .6603569
------------------------------------------------------------------------------
```

Como se puede observar de las predicciones del MPL (pMPL) no se encuentran acotadas al espacio [0,1]:

```
. gsort -pMPL

. list pMPL perman in 1/10

     +-------------------+
     |     pMPL   perman |
     |-------------------|
  1. | 1.898298        1 |
  2. | 1.349436        1 |
  3. | 1.288475        1 |
  4. | 1.285287        1 |
  5. | 1.185919        1 |
     |-------------------|
  6. | 1.104877        0 |
  7. | 1.099564        1 |
  8. |  1.09791        1 |
  9. | 1.067071        1 |
 10. | 1.060412        1 |
     +-------------------+
```

En el caso de modelos Binarios, Stata permite estimar tanto el modelo Logit como el Probit de la siguiente manera: en el menú de herramientas se selecciona la opción **Statistics** para después elegir **Binary outcomes**. Una vez hecho esto, se puede elegir bien sea el modelo Logit o el Probit, u otras opciones que proporciona el programa.



El siguiente paso consiste en ingresar la variable dependiente acompañada de las independientes y se da **Ok**. De igual forma, se puede escribir:

logit depvar var1 var2 var3 **[if] [in] [weight] [, options]**

El programa, con los datos anteriores arroja los siguientes resultados:

```
Iteration 0:   log likelihood = -5136.1985
Iteration 1:   log likelihood = -4820.1328
Iteration 2:   log likelihood = -4813.6324
Iteration 3:   log likelihood = -4813.5953
Iteration 4:   log likelihood = -4813.5953

Logistic regression                               Number of obs   =       8229
                                                  LR chi2(6)      =     645.21
                                                  Prob > chi2     =     0.0000
Log likelihood = -4813.5953                       Pseudo R2       =     0.0628
```

perman	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
sexo	.4410419	.0502824	8.77	0.000	.3424902	.5395936
exp	.0604604	.0072491	8.34	0.000	.0462524	.0746683
SO1	-.6752616	.0737608	-9.15	0.000	-.8198301	-.5306932
SO2	-1.836556	.0981449	-18.71	0.000	-2.028917	-1.644195
posg	-.169647	.0785897	-2.16	0.031	-.3236799	-.0156141
nooficial	.2358649	.0498234	4.73	0.000	.1382129	.3335169
_cons	.5274648	.0473659	11.14	0.000	.4346293	.6203004

Cabe recordar que los coeficientes β 's estimados solamente permiten contrastar los signos esperados a priori. Las interpretaciones económicas se deben realizar a través de los efectos marginales. El comando para el cálculo de estos es:

mfx

```
Marginal effects after logit
      y  = Pr(perman) (predict)
         =  .69408246
```

variable	dy/dx	Std. Err.	z	P>|z|	[95%	C.I.]	X
sexo*	.0928015	.01043	8.90	0.000	.072368	.113235	.459959
exp	.0128377	.00153	8.40	0.000	.009844	.015831	2.97893
SO1*	-.1552326	.01785	-8.69	0.000	-.190227	-.120239	.111192
SO2*	-.4293251	.0202	-21.25	0.000	-.468914	-.389736	.067809
posg*	-.0369017	.01748	-2.11	0.035	-.071163	-.002641	.113866
noofic~l*	.0502399	.01063	4.73	0.000	.029408	.071071	.541743

(*) dy/dx is for discrete change of dummy variable from 0 to 1

En el caso del modelo Probit:

```
Iteration 0:   log likelihood = -5136.1985
Iteration 1:   log likelihood = -4817.8281
Iteration 2:   log likelihood = -4816.2189
Iteration 3:   log likelihood = -4816.2188

Probit regression                                 Number of obs   =       8229
                                                  LR chi2(6)      =     639.96
                                                  Prob > chi2     =     0.0000
Log likelihood = -4816.2188                       Pseudo R2       =     0.0623
```

perman	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
sexo	.2620754	.0299219	8.76	0.000	.2034295	.3207212
exp	.0320677	.0037783	8.49	0.000	.0246624	.0394731
SO1	-.4188665	.0453618	-9.23	0.000	-.5077741	-.329959
SO2	-1.136434	.0591561	-19.21	0.000	-1.252378	-1.02049
posg	-.1018376	.0469733	-2.17	0.030	-.1939035	-.0097716
nooficial	.1417973	.0298244	4.75	0.000	.0833426	.200252
_cons	.3418753	.0283036	12.08	0.000	.2864012	.3973494

Y, como se puede observar aparecen los coeficientes estimados, los errores estándar y los valores respectivos de las z.

El contraste Lr (razón de verosimilitud) del modelo estimado versus el modelo ingenuo (solo la constante) es reportado como **LR chi2(6)** (el número 6 corresponde al número de variables independientes) y su respectivo p-valor. Esto permite concluir sobre la significancia global del modelo. Este test es una prueba de que todas las pendientes son cero, análogo a la prueba F, usualmente utilizada en los modelos de regresión lineal.

El comando para calcular los efectos marginales es el mismo del modelo anterior y se escribe después de la estimación:

mfx

```
Marginal effects after probit
      y  = Pr(perman) (predict)
         =  .69129815
```

variable	dy/dx	Std. Err.	z	P>|z|	[95%	C.I.]	X
sexo*	.0916076	.01035	8.85	0.000	.071321	.111894	.459959
exp	.0112926	.00133	8.51	0.000	.008691	.013894	2.97893
SO1*	-.1568788	.01768	-8.88	0.000	-.191523	-.122235	.111192
SO2*	-.4300989	.02016	-21.33	0.000	-.469616	-.390582	.067809
posg*	-.0365325	.01714	-2.13	0.033	-.07013	-.002935	.113866
noofic~l*	.0500488	.01054	4.75	0.000	.029389	.070709	.541743

(*) dy/dx is for discrete change of dummy variable from 0 to 1

Los efectos marginales se interpretan de la manera usual. Es decir, por ejemplo en el caso de sexo: Las mujeres tienen una probabilidad del 9% de permanecer en el área de desempeño con respecto a los hombres. Una semana más de experiencia incremente en un 1% la probabilidad de permanecer en el área de desempeño. También muestra los errores estándar, los valores z con su respectivo p-valor y el intervalo de confianza.

Si lo que se desea es calcular los efectos marginales sobre un valor puntual y no el efecto marginal promedio, se debe introducir:

```
. mfx, at(sexo=1 exp=54 posg=0 SO1=1 SO2=0 nooficial=1)

Marginal effects after probit
      y  = Pr(perman) (predict)
         =   .9802308
------------------------------------------------------------------------------
variable |      dy/dx    Std. Err.     z    P>|z|  [    95% C.I.   ]      X
---------+--------------------------------------------------------------------
   sexo* |   .0164412      .00656    2.51   0.012   .003593   .02929        1
     exp |   .0015374      .00047    3.26   0.001   .000614   .00246       54
   posg* |  -.0054221      .00334   -1.62   0.105  -.011976  .001132        0
    SO1* |  -.0131521      .00628   -2.09   0.036  -.025457 -.000848        1
    SO2* |  -.1584677      .04692   -3.38   0.001  -.250437 -.066498        0
noofic~1*|   .0078662      .00361    2.18   0.029   .000789  .014944        1
------------------------------------------------------------------------------
(*) dy/dx is for discrete change of dummy variable from 0 to 1
```

Como se puede observar, el comando anterior calcula los efectos marginales para las mujeres, con un año de experiencia, que no tienen posgrado, de instituciones no oficiales y que tienen un solo nivel de sobre educación.

6.11.2 Análisis de Especificidad, Sensibilidad y porcentaje de aciertos

Una forma adicional de analizar nuestro modelo consiste en calcular cuantas de las predicciones del modelo de los valores que fueron ceros o unos realmente son ceros o unos. De esta forma, se pueden calcular el porcentaje de predicciones que fueron correctas, el porcentaje de ceros correctos sobre el total de ceros (especificidad) y el porcentaje de unos que fueron correctos sobre el total de unos (sensibilidad). Siguiendo con el ejemplo anterior, a partir de la predicción de la variable dependiente se obtiene $\hat{Perman}$ y se construye la siguiente tabla:

		Valor real de Perman	
		Perman=0	Perman=1
Predicción $\hat{Perman}$	$\hat{Perman} \leq 0.5$	P_{11}	P_{12}
	$\hat{Perman} > 0.5$	P_{21}	P22

Donde 0.5 es el valor umbral el cual no es problemático en principio si la proporción de unos y ceros no es muy diferente. De esta forma, el porcentaje de aciertos, errores, la especificidad y sensibilidad se definen como:

Índices para medir la bondad del ajuste

Indice	Definición	Expresión
Tasa de aciertos	Cociente entre las predicciones correctas y el total de predicciones	$\frac{P_{11}+P_{22}}{P_{11}+P_{12}+P_{21}+P_{22}}$
Tasa de errores	Cociente entre las predicciones incorrectas y el total de predicciones	$\frac{P_{12}+P_{21}}{P_{11}+P_{12}+P_{21}+P_{22}}$
Especificidad	Proporción entre la frecuencia de valores 0 correctos y el total de valores 0 observados	$\frac{P_{11}}{P_{11}+P_{21}}$
Sensibilidad	Razón entre los valores 1 correctos y el total de valores 1 observados	$\frac{P_{22}}{P_{12}+P_{22}}$
Tasa de falsos ceros	Proporción entre la frecuencia de valores 0 incorrectos y el total de valores 0 observados	$\frac{P_{21}}{P_{11}+P_{21}}$
Tasa de falsos unos	Razón entre los valores 1 incorrectos y el total de valores 1 observados	$\frac{P_{12}}{P_{12}+P_{22}}$

Estos resultados se pueden obtener al correr el siguiente comando:

```
. estat classification

Probit model for perman

              -------- True --------
Classified |         D            ~D  |      Total
-----------+--------------------------+-----------
     +     |      5379          2126  |       7505
     -     |       246           478  |        724
-----------+--------------------------+-----------
   Total   |      5625          2604  |       8229

Classified + if predicted Pr(D) >= .5
True D defined as perman != 0
--------------------------------------------------
Sensitivity                     Pr( +| D)   95.63%
Specificity                     Pr( -|~D)   18.36%
Positive predictive value       Pr( D| +)   71.67%
Negative predictive value       Pr(~D| -)   66.02%
--------------------------------------------------
False + rate for true ~D        Pr( +|~D)   81.64%
False - rate for true D         Pr( -| D)    4.37%
False + rate for classified +   Pr(~D| +)   28.33%
False - rate for classified -   Pr( D| -)   33.98%
--------------------------------------------------
Correctly classified                        71.18%
--------------------------------------------------
```

6.11.3 Contrastes de Momentos Condicionales

Pagan y Vella (1989) proponen realizar un contraste sobre las restricciones de los momentos condicionales de la forma:

$$E\left[n^{-1}\sum_{i=N+1}^{N+n} Z_i(y_i - x_i\beta)\right] = 0$$

Donde N+1 hasta N+n son los individuos sobre los que se ha realizado la predicción y Z puede contener muchas variables pero en el caso de que s sea igual a x la expresión podría ser proporcional al vector de puntuación.

Debido a que los modelos Probit y Logit se estiman por máxima verosimilitud, en la cual cada observación es extraída de una distribución de Bernoulli, la probabilidad de suceso f(β'x) y con observaciones independientes lleva a una probabilidad conjunta o a una función de verosimilitud de la forma:

$$\Pr ob(Y_1 = y_1, Y_2 = y_2, Y_3 = y_3, \ldots, Y_n = y_n) = \prod_{y_i=0}\left[1 - F(\beta' x_i)\right]\prod_{y_i=1} F(\beta' x_i)$$

Reacomodando términos, la anterior expresión se puede escribirse como

$$L = \prod_i \left[F(\beta' x_i)\right]^{y_i}\left[1 - F(\beta' x_i)\right]^{1-y_i}$$

Que es la probabilidad para una muestra de n observaciones. Tomando logaritmos

$$Ln\ L = \sum_i y_i\, LnF(\beta' x_i) + (1 - y_i)\, Ln\left[1 - F(\beta' x_i)\right]$$

Al derivar con respecto al vector de parámetros se obtiene

$$\frac{\partial Ln\, L}{\partial \beta} = \sum_i \left[\frac{\left[y_i - (1 - F(\beta' x_i))\right] f(\beta' x_i)}{F(\beta' x_i)\left[1 - F(\beta' x_i)\right]}\right] x_i = 0$$

$$\frac{\partial Ln\, L}{\partial \beta} = \sum_i e_{g,i}\, x_i = 0$$

Donde $e_{g,i}$ son los residuos generalizados en la ecuación de Pagan y Vella (1989). De lo anterior, Pagan y Vella (1989) deducen los siguientes contrastes:

1.- Variables Omitidas

$$E\left[Z_i * e_{g,i}\right] = 0$$

2.- Contraste RESET

$$E\left[\hat{p}_i^2 * e_{g,i}\right] = 0$$

$$E\left[\hat{p}_i^3 * e_{g,i}\right] = 0$$

3.- Heteroscedasticidad

$$E\left[Z_i * \hat{p}_i * e_{g,i}\right] = 0$$

En el contraste de variables omitidas Z será la variable que se sospecha se omitió, mientras que en el contraste de normalidad (RESET) la variable $\hat{p}_i$ es la probabilidad estimada para el modelo logit o probit. Finalmente en el contraste de heteroscedasticidad Z_i es la variable que se sospecha tiene problemas de heteroscedasticidad. De igual forma, Mora (2002) provee una serie de contrastes adicionales y equivalentes para los modelos probit y logit. Por ejemplo, en el caso del contraste RESET se tendría.

Source	SS	df	MS	
Model	919.74064	6	153.290107	Number of obs = 8229
Residual	223.405105	8222	.027171626	F(6, 8222) = 5641.55
Total	1143.14575	8228	.13893361	Prob > F = 0.0000
				R-squared = 0.8046
				Adj R-squared = 0.8044
				Root MSE = .16484

$\hat{p}_i^2 * e_{g,i}$	Coef.	Std. Err.	t	P>|t|	[95% Conf. Interval]	
scoresexo	.3238262	.0043739	74.04	0.000	.3152521	.3324002
scoreexp	.0276042	.0004647	59.41	0.000	.0266933	.0285151
scoreSO1	-.0467561	.0072239	-6.47	0.000	-.0609167	-.0325954
scoreSO2	-.3343015	.0098024	-34.10	0.000	-.3535166	-.3150863
scoreposg	.0516048	.0076096	6.78	0.000	.036688	.0665216
scorenooficial	.2947448	.0042073	70.06	0.000	.2864974	.3029921
_cons	.0000626	.0018171	0.03	0.973	-.0034994	.0036246

Y,

Source	SS	df	MS		
Model	542.383343	6	90.3972238	Number of obs =	8229
Residual	92.2369676	8222	.011218313	F(6, 8222) =	8058.01
Total	634.62031	8228	.077129352	Prob > F =	0.0000
				R-squared =	0.8547
				Adj R-squared =	0.8546
				Root MSE =	.10592

$\hat{p}_i^3 * e_{g,i}$	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
scoresexo	.2553017	.0028105	90.84	0.000	.2497925	.260811
scoreexp	.0243145	.0002986	81.44	0.000	.0237292	.0248998
scoreSO1	-.0948046	.0046417	-20.42	0.000	-.1039035	-.0857057
scoreSO2	-.2998073	.0062985	-47.60	0.000	-.312154	-.2874607
scoreposg	.0138088	.0048895	2.82	0.005	.0042241	.0233936
scorenooficial	.2139213	.0027034	79.13	0.000	.2086219	.2192206
_cons	-.000241	.0011676	-0.21	0.836	-.0025298	.0020478

Dado el p-value de la constante en ambas regresiones no se puede rechazar el supuesto de normalidad.

Suponga a continuación que deseamos contrastar que la heteroscedasticidad es generada por la variable mujer. El contraste tomaría la forma:

Source	SS	df	MS		
Model	1065.44293	6	177.573821	Number of obs =	8229
Residual	16.5224777	8222	.002009545	F(6, 8222) =	88365.19
Total	1081.9654	8228	.131497983	Prob > F =	0.0000
				R-squared =	0.9847
				Adj R-squared =	0.9847
				Root MSE =	.04483

$Mujer_i * \hat{p}_i * e_{g,i}$	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
scoresexo	.7255415	.0011895	609.95	0.000	.7232098	.7278732
scoreexp	.0032798	.0001264	25.96	0.000	.0030321	.0035276
scoreSO1	-.0619329	.0019645	-31.53	0.000	-.0657839	-.0580819
scoreSO2	-.2437095	.0026658	-91.42	0.000	-.2489351	-.2384839
scoreposg	-.0104261	.0020694	-5.04	0.000	-.0144827	-.0063695
scorenooficial	.0235149	.0011442	20.55	0.000	.0212721	.0257578
_cons	-.0012903	.0004942	-2.61	0.009	-.002259	-.0003216

De acuerdo al p-value de la constante se rechaza la hipótesis de homocedasticidad y por lo tanto deberá estimarse un probit con heteroscedasticidad: Los resultados del probit y el probit en el que se corrigen los problemas de heteroscedasticidad se observan a continuación:

	Probit b/se	HetProb b/se
perman		
sexo (d)	0.0916***	0.0598***
	(0.0104)	(0.0101)
exp	0.0113***	0.0411**
	(0.0013)	(0.0150)
SO1 (d)	-0.1569***	-0.0376**
	(0.0177)	(0.0122)
SO2 (d)	-0.4301***	-0.4143***
	(0.0202)	(0.0412)
posg (d)	-0.0365*	-0.0297*
	(0.0171)	(0.0117)
nooficial (d)	0.0500***	0.0256***
	(0.0105)	(0.0069)
lnsigma2		
sexo (d)		0.0598***
		(0.0101)
exp		0.0411**
		(0.0150)
SO2 (d)		-0.4143***
		(0.0412)
Number	8.2e+03	8.2e+03
Log-Likelihood	-4.8e+03	-4.7e+03

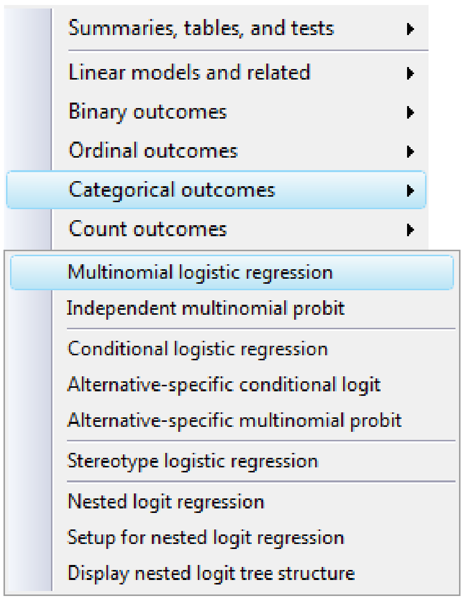
(d) for discrete change of dummy variable from 0 to 1
* p<0.05, ** p<0.01, *** p<0.001

La anterior tabla muestra los efectos marginales para el probit y el probit con heteroscedasticidad en las variables sexo, exp y SO2.

6.11.4 Modelos Multinomiales

Los datos para este tipo de estimación son tomados de Mora y Ulloa (2011) sobre calidad del empleo. En Mora y Ulloa (2011) se discute la relación entre la calidad del empleo y la educación. Se consideran tres niveles de calidad: baja calidad, cuando el índice de la calidad del trabajo es menor a 60 puntos; calidad media, si el índice de la calidad del trabajo se encuentra entre valores de 60 y 80 puntos; y buena calidad, cuando el índice de la calidad del trabajo es superior a 80 puntos. Las variables explicativas de la calidad del empleo se incluyen el sexo, el cual toma valor de uno si es hombre y cero si es mujer; el jefe de hogar y casado que toman valor de uno cuando la característica se presenta y cero en caso contrario. La variable años de educación se incluye como una variable continua la cual es igual a los años de educación. Con respecto a esta última la base de datos incluye la estimación de estas en lugar de la variable original [Ver Mora y Ulloa (2011)].

Con el fin de realizar la estimación de un modelo de variable dependiente con múltiples alternativas se deben seguir los siguientes pasos: en el menú se selecciona **Categorical outcomes** y posteriormente se elige **Multinomial logistic regression**, tal como se muestra aquí:



El comando que se utiliza es:

.mlogit vardep var1 var2 var3

Al utilizar este comando en Stata, la salida que muestra el programa es:

```
Iteration 0:   log likelihood = -33903.238
Iteration 1:   log likelihood = -32375.713
Iteration 2:   log likelihood = -32316.372
Iteration 3:   log likelihood = -32316.064
Iteration 4:   log likelihood = -32316.064

Multinomial logistic regression                   Number of obs   =      38679
                                                  LR chi2(34)     =    3174.35
                                                  Prob > chi2     =     0.0000
Log likelihood = -32316.064                       Pseudo R2       =     0.0468
```

DIC	RRR	Std. Err.	z	P>|z|	[95% Conf.	Interval]
Valores_menores_a_60_del_ICTtota	(base outcome)					
Valores_entre_60_y_80_del_ICEtot						
sexo	1.671325	.0594642	14.44	0.000	1.558748	1.792032
jefehogar	1.129309	.0324408	4.23	0.000	1.067483	1.194716
casado	1.089241	.0292182	3.19	0.001	1.033454	1.14804
edad	1.025899	.0025295	10.37	0.000	1.020953	1.030869
Med	1.140606	.0568583	2.64	0.008	1.034436	1.257672
Blla	.4224253	.0240687	-15.12	0.000	.3777902	.472334
Bga	.6140148	.0350801	-8.54	0.000	.5489689	.6867679
Cna	.5973351	.0344654	-8.93	0.000	.5334637	.6688538
Manz	.9850426	.0560297	-0.26	0.791	.8811264	1.101214
Mont	.3132066	.0206504	-17.61	0.000	.2752387	.3564121
Pas	.3889021	.0259693	-14.14	0.000	.3411933	.443282
Cuc	.3430424	.022685	-16.18	0.000	.3013415	.3905141
Per	.9836367	.0571411	-0.28	0.776	.8777827	1.102256
Vill	.533714	.0322714	-10.38	0.000	.4740674	.6008654
Ibg	.5736775	.0354664	-8.99	0.000	.5082111	.6475771
Cal	.6351048	.0354108	-8.14	0.000	.5693588	.7084428
shat	1.77975	.0515125	19.92	0.000	1.681598	1.883632
_cons	.00039	.0001537	-19.91	0.000	.0001801	.0008446
Valores_superiores_a_80_del_ICEt						
sexo	1.84008	.0898302	12.49	0.000	1.672177	2.024842
jefehogar	1.449714	.0557691	9.65	0.000	1.344427	1.563245
casado	1.2788	.0468161	6.72	0.000	1.190256	1.37393
edad	1.082429	.0041272	20.77	0.000	1.07437	1.090548
Med	1.164388	.0763598	2.32	0.020	1.023945	1.324094
Blla	.3580138	.028936	-12.71	0.000	.3055642	.4194662
Bga	.8318147	.0600925	-2.55	0.011	.721994	.95834
Cna	.4097119	.0351313	-10.41	0.000	.3463309	.4846921
Manz	.9500321	.0713977	-0.68	0.495	.8199133	1.1008
Mont	.5132297	.0405657	-8.44	0.000	.4395748	.5992263
Pas	.5762372	.0465309	-6.83	0.000	.491889	.6750493
Cuc	.4076554	.0356069	-10.27	0.000	.343514	.4837734
Per	.7028949	.0585579	-4.23	0.000	.5970038	.827568
Vill	.5820909	.0467069	-6.74	0.000	.4973825	.6812257
Ibg	.6097563	.049476	-6.10	0.000	.5201028	.7148639
Cal	.5665232	.0437272	-7.36	0.000	.4869872	.6590493
shat	2.178556	.0836016	20.29	0.000	2.02071	2.348732
_cons	2.11e-06	1.16e-06	-23.81	0.000	7.18e-07	6.18e-06

En primer lugar, observe que en el modelo anterior se estimó con base en la primera categoría, es decir, con base a valores menores a 60 puntos en el índice de calidad del empleo. La base se puede cambiar anexando simplemente la base de referencia se desee cambiar anexando simplemente ésta en el comando:

```
. mlogit DIC sexo jefehogar casado edad  Med Blla Bga Cna Manz Mont Pas Cuc Per Vill Ibg Cal shat,b(3)

Iteration 0:   log likelihood = -33903.238
Iteration 1:   log likelihood = -32375.713
Iteration 2:   log likelihood = -32316.372
Iteration 3:   log likelihood = -32316.064
Iteration 4:   log likelihood = -32316.064
```

De esta forma, la estimación del modelo se realizaría con respecto a la tercera categoría (valores superiores a 80 puntos en el ICE).

En segundo lugar, el comando rrr después de la coma permite estimar la razón de riesgo relativo o RRR este es un odds ratio, pero para una comparación especifica con la base elegida [b(.)]. El RRR muestra el incremento en el odd de Y1 dado x1=1 relativo al dos de

Y2 cuando x1=1 siendo Y1 y Y2 resultados mutualmente excluyentes. De esta forma, con respecto al género, se puede observar que la probabilidad de tener un empleo de calidad media sobre un empleo de baja calidad es 1.6 veces mayor para los trabajadores hombres.

Ahora bien, esto no significa que no se puedan calcular los efectos marginales. Estos se pueden calcular para cada categoría como se observa a continuación:

```
. mfx, predict(p outcome(3))

Marginal effects after mlogit
      y  = Pr(DIC==Valores_superiores_a_80_del_ICEt) (predict, p outcome(3))
         =  .10472125
```

variable	dy/dx	Std. Err.	z	P>|z|	[95%	C.I.]	X
sexo*	.0439924	.00431	10.21	0.000	.035544	.05244	.547429
jefeho~r*	.0321445	.00358	8.99	0.000	.025134	.039155	.4716
casado*	.0207596	.00328	6.33	0.000	.014334	.027185	.563975
edad	.0068025	.00033	20.40	0.000	.006149	.007456	38.3389
Med*	.0113356	.00633	1.79	0.073	-.00107	.023741	.110422
Blla*	-.0608693	.00419	-14.51	0.000	-.069089	-.05265	.090411
Bga*	-.0068594	.00626	-1.10	0.273	-.019135	.005416	.079268
Cna*	-.0562826	.00455	-12.36	0.000	-.065207	-.047358	.072365
Manz*	-.004369	.00657	-0.66	0.506	-.017248	.00851	.0689
Mont*	-.0364881	.00529	-6.90	0.000	-.046849	-.026127	.072261
Pas*	-.0304087	.00571	-5.33	0.000	-.041593	-.019224	.06267
Cuc*	-.0514713	.00488	-10.55	0.000	-.061032	-.04191	.065539
Per*	-.0289542	.00591	-4.90	0.000	-.040532	-.017377	.061816
Vill*	-.0332003	.00553	-6.01	0.000	-.04403	-.022371	.06797
Ibg*	-.030784	.0057	-5.40	0.000	-.041957	-.019611	.063342
Cal*	-.0374313	.00515	-7.27	0.000	-.047521	-.027341	.081982
shat	.0589441	.00341	17.30	0.000	.052265	.065623	10.2461

(*) dy/dx is for discrete change of dummy variable from 0 to 1

En tercer lugar, Es posible que sea deseable combinar algunas categorías de la variable dependiente, en parte porque hace el análisis simple, en parte porque el número de casos en algunas de las categorías resulte ser muy pequeño. Entre más categorías se tengan, mas parámetros a estimar y más difícil será concluir

```
**** Wald tests for combining outcome categories

 Ho: All coefficients except intercepts associated with given pair
     of outcomes are 0 (i.e., categories can be collapsed).

Categories tested |      chi2   df   P>chi2
------------------+------------------------
Valores_-Valores_ |  1677.071   17    0.000
Valores_-Valores_ |  1272.445   17    0.000
Valores_-Valores_ |   800.559   17    0.000
-------------------------------------------
```

Como se puede observar del anterior contraste, es posible rechazar la hipótesis de que todos los coeficientes son iguales en todas las categorías asociadas por lo tanto no se pueden agrupar categorías.

Finalmente, en los modelos multinomiales se debe cumplir el supuesto de que las razones de ocurrencia de cada par de valores de la variable independiente no deben estar afectado por el resto de las alternativas posibles (añadiendo o disminuyendo alguna de ellas). De esta forma, cada alternativa de este totalmente diferenciada y la probabilidad relativa de elegir la alternativa j relativa a k no dependerá del resto de resultados disponibles. Este supuesto es conocido como el supuesto de la independencia de las alternativas irrelevantes o IIA.

Una forma de contrastar si se cumple este supuesto es a través del contraste de Hausman. De esta forma se estima el modelo completo con las J alternativas y se estima un modelo restringido eliminando una o más J. A continuación se prueba la diferencia entre los dos modelos, la cual se distribuye asintóticamente como una chi cuadrado con tantos grados de libertad como ecuaciones a estimar existan en el modelo restringido y, valores significativos de la prueba indican que se viola el supuesto (la diferencia entre los dos modelos no es 0). La expresión del contraste es la siguiente:

Se estima el modelo con el conjunto de elección completo (*C*) y restringido (R). Bajo la hipótesis nula de que se cumple IIA, los coeficientes deberían ser similares. Se contrasta la significancia de la anterior diferencia:

$$\chi^2 = (\hat{\beta}_R - \hat{\beta}_C)' \left[\hat{V}_R - \hat{V}_C \right] (\hat{\beta}_R - \hat{\beta}_C)$$

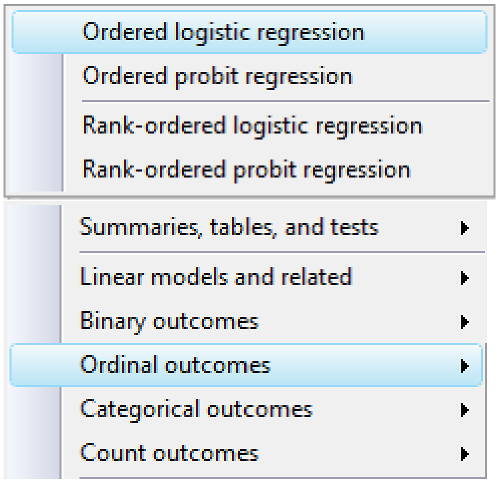
En la ecuación anterior $\hat{V}_C$ es la matriz de varianzas y covarianzas del modelo completo mientras que $\hat{V}_R$ es la matriz de varianzas del modelo restringido. Si rechazamos la hipótesis nula de validez de IIA, debemos emplear un modelo alternativo al logit multinomial.

```
. hausman partial all, alleqs constant

                 ---- Coefficients ----
             |      (b)          (B)            (b-B)     sqrt(diag(V_b-V_B))
             |    partial        all         Difference          S.E.
-------------+----------------------------------------------------------------
        sexo |    .5647727    -.2687546        .8335273               .
   jefehogar |    .2622747    -.4055201        .6677948               .
      casado |    .1846013    -.2837098        .4683111               .
        edad |    .0299979    -.0553233        .0853212               .
       yeduc |    .1894117    -.2414612        .4308729               .
         Med |    .1710914    -.1456461        .3167376               .
        Blla |   -1.016859     .2677948       -1.284653               .
         Bga |   -.4558985     -.433659       -.0222395               .
         Cna |   -.7033874     .5710211       -1.274409               .
        Manz |   -.0548348    -.0153654       -.0394695               .
        Mont |   -1.084292    -.5419053       -.5423868               .
         Pas |    -1.04352    -.1404773       -.9030424               .
         Cuc |   -.9005957    -.2228602       -.6777355               .
         Per |    .1424866     .0891806         .053306               .
        Vill |   -.5087335    -.2927301       -.2160034               .
         Ibg |    -.629747     .0313774       -.6611244               .
         Cal |   -.4151559      .054519        -.469675               .
        shat |    .4343524    -.2195432        .6538957               .
       _cons |   -8.643842     9.204713       -17.84856               .
------------------------------------------------------------------------------
                           b = consistent under Ho and Ha; obtained from mlogit
            B = inconsistent under Ha, efficient under Ho; obtained from mlogit

    Test:  Ho:  difference in coefficients not systematic

                 chi2(19) = (b-B)'[(V_b-V_B)^(-1)](b-B)
                          = -5202.13    chi2<0 ==> model fitted on these
                                        data fails to meet the asymptotic
                                        assumptions of the Hausman test;
                                        see suest for a generalized test
```

De acuerdo a los resultados no existe evidencia estadística suficiente para rechazar Ho. El valor negativo ha sido discutido ampliamente en los manuales de STATA. En particular Long and Freese (2006: 244-5) citan que es muy común que esto se presente y que es evidencia de que el supuesto de IIA no ha sido violado.

6.11.5 Modelo de respuesta ordenada

En Stata, se deben realizar los siguientes pasos para la estimación de este tipo de modelos. En la barra de herramientas se selecciona el menú **Statistics**, donde se desprende lo siguiente:



También se pueden utilizar los siguientes comandos:

ologit vardep var1 var2 var3
oprobit vardep var1 var2 var3

```
. ologit DIC sexo jefehogar casado Med Blla Bga Cna Manz Mont Pas Cuc Per Vill Ibg Cal shat

Iteration 0:   log likelihood = -33903.238
Iteration 1:   log likelihood = -32990.768
Iteration 2:   log likelihood = -32980.974
Iteration 3:   log likelihood = -32980.969
Iteration 4:   log likelihood = -32980.969

Ordered logistic regression                       Number of obs   =      38679
                                                  LR chi2(16)     =    1844.54
                                                  Prob > chi2     =     0.0000
Log likelihood = -32980.969                       Pseudo R2       =     0.0272

------------------------------------------------------------------------------
         DIC |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        sexo |   .1467956   .0235181     6.24   0.000     .1007011    .1928902
   jefehogar |   .3493868   .0236092    14.80   0.000     .3031137    .3956599
      casado |   .2516935   .0222128    11.33   0.000     .2081572    .2952298
         Med |   .1235174   .0421296     2.93   0.003     .0409449    .2060899
        Blla |   -.861604   .0490745   -17.56   0.000    -.9577883   -.7654197
         Bga |  -.3197155   .0479499    -6.67   0.000    -.4136955   -.2257355
         Cna |  -.5921552   .0502237   -11.79   0.000    -.6905918   -.4937186
        Manz |  -.0068455   .0482288    -0.14   0.887    -.1013723    .0876813
        Mont |  -.8993398   .0536451   -16.76   0.000    -1.004482   -.7941975
         Pas |  -.7081484   .0548046   -12.92   0.000    -.8155635   -.6007333
         Cuc |  -.9736096   .0556964   -17.48   0.000    -1.082773   -.8644466
         Per |  -.1401025   .0500421    -2.80   0.005    -.2381833   -.0420217
        Vill |  -.5725617   .0513804   -11.14   0.000    -.6732653    -.471858
         Ibg |  -.4944485   .0523701    -9.44   0.000     -.597092   -.3918049
         Cal |  -.4623591   .0478725    -9.66   0.000    -.5561874   -.3685308
        shat |   .1888666   .0098998    19.08   0.000     .1694633    .2082699
-------------+----------------------------------------------------------------
       /cut1 |   2.523334   .1148925                      2.298149    2.748519
       /cut2 |   4.025641   .1160232                       3.79824    4.253042
------------------------------------------------------------------------------
```

Observe que el comando solo estima un vector de parámetros para una de las categorías y a partir de esta se resécala con /cut1 y /cut2.

El contraste de Brant comprueba el supuesto de las regresiones paralelas o el supuesto de los odds proporcionales. El contraste compara los coeficientes de las pendientes de J-1 logits implícitos en la regresión ordenada. Stata reporta los resultados de un contraste ómnibus para todo el modelo y además contrasta el supuesto para cada una de las variables independientes con el contraste Brant:

Brant Test of Parallel Regression Assumption

Variable	chi2	p>chi2	df
All	1028.54	0.000	18
sexo	0.27	0.605	1
jefehogar	16.39	0.000	1
casado	10.18	0.001	1
edad	50.74	0.000	1
yeduc	398.34	0.000	1
Med	0.27	0.603	1
Blla	13.10	0.000	1
Bga	51.62	0.000	1
Cna	0.32	0.570	1
Manz	0.64	0.423	1
Mont	103.86	0.000	1
Pas	44.55	0.000	1
Cuc	35.49	0.000	1
Per	2.30	0.130	1
Vill	30.07	0.000	1
Ibg	11.58	0.001	1
Cal	4.44	0.035	1
shat	0.43	0.513	1

A significant test statistic provides evidence that the parallel regression assumption has been violated.

Una forma alternativa consiste en estimar un modelo completo con todas las categorías y un modelo restringido excluyendo alguna y realizar un contraste tipo LR de la forma siguiente:

lrtest constrained unconstrained
Likelihood-ratio test LR chi2(16) = 629.82
(Assumption: constrained nested in unconstrained) Prob > chi2 = 0.0000

Como se puede observar de ambos contrastes se rechaza la hipótesis de que el supuesto de las regresiones paralelas se mantiene. ¿Qué sucede si se viola el supuesto de las regresiones paralelas? Se puede utilizar un logit generalizado gologit2 o un logit multinomial. Y ¿que sucede cuando se presenta sesgo de selección?

7 Modelos de utilidad discreta

Generalmente las elecciones de los consumidores involucran elecciones discretas como usar gas o no, usar energía eléctrica o no, comprar un automóvil o no, etc.

En los capítulos anteriores hemos considerado a un individuo que elige una alternativa de un conjunto de elecciones finitas A. Si en A se cumplen los axiomas de *completitud, reflexividad* y *transitividad,* y dado que A es finita, entonces una alternativa óptima $a^* \in$ A para el individuo está definida como la alternativa $a^* \geq a \;\forall\; a \in$ A, por lo cual, es posible encontrar una función de utilidad que represente las preferencias individuales, esto es, existe una función U(·) que satisface la propiedad U(a*) $\geq$ U(a) si y solo si $a^* \geq a$ cuando la alternativa a* maximiza U sobre A .

La anterior aproximación, ha sido criticada por sicólogos como Thurstone (1927), Luce y Supes (1955), Tversky (1969) y por economistas como Georgescu-Roegen (1958), Quandt (1956) y Macfadden (1981, 1986) ya que implica fuertes postulados sobre el poder discriminatorio de los agentes, así como una capacidad ilimitada de procesar información.

Para Tversky, cuando se realiza una elección entre varias alternativas, las personas parten de experiencias inciertas e inconsistentes. Esto es, las personas no están seguras sobre cual alternativa deberían seleccionar, así como tampoco toman siempre la misma elección bajo condiciones parecidas. Este comportamiento, aparentemente irracional, lleva al autor a concluir que "el proceso de elección debe ser visto como un proceso probabilístico" (Tversky, 1972, p. 281).

Naturalmente, deberemos preguntarnos qué factores determinan dicha probabilidad. Esto es, el comportamiento de los agentes es intrínsecamente probabilístico o el modelador no puede representar el comportamiento del consumidor o ambos. Con respecto a lo primero Quandt (1956) arguye que una alternativa puede ser vista como un conjunto finito de características, donde las preferencias son definidas directamente sobre las características e indirectamente sobre las alternativas.

Para Quandt un individuo podrá en alguna ocasión considerar algunas características de una alternativa y/o cometer un error al evaluar la importancia de una característica asociada con una alternativa, así, las circunstancias bajo las cuales las elecciones son efectivamente realizadas pueden "perturbar" la percepción y/o la deseabilidad de una alternativa. Quandt cita el siguiente ejemplo: Un hombre que compra vino, podría comprar una botella sin tener en cuenta la cosecha; pero ante la presencia de un catador de vinos él podría comprar un vino de mejor cosecha; de esta forma, el comportamiento individual podría cambiar de acuerdo a factores externos, sin que las preferencias individuales sobre las características hayan cambiado. Desde este punto de vista, el proceso de elección es intrínsecamente probabilístico.

Para Manski (1977) la falta de información lleva al modelador a determinar reglas probabilísticas de elección en los individuos más que la falta de racionalidad.

Ambas interpretaciones llevan a un esquema probabilístico donde se pueden diferenciar dos familias de modelos: La primera familia parte de que la regla de decisión es estocástica mientras la utilidad es determinística (Luce, Tversky). La segunda familia parte de una regla de decisión determinística mientras la utilidad es estocástica (Macfadden, Thurstone). Ambas familias pueden distinguirse a través de la naturaleza del mecanismo aleatorio que gobierna la elección.

Otra aproximación es realizada por Machina (1985) para quien el individuo maximiza una utilidad determinística definida sobre loterías en el conjunto de elección. En la aproximación de Machina la utilidad se define sobre las loterías, y las probabilidades de estados alternativos de la naturaleza provienen exógenamente[15].

7.1 Reglas de decisión

7.1.1 Modelos con regla de decisión estocástica

La interpretación, proviene de Tversky (1972a), para quien la utilidad de diferentes alternativas es determinística, pero el proceso de elección en sí mismo es probabilístico. En este tipo de modelos el individuo no necesariamente elige la alternativa que da la mayor utilidad, en lugar de esto, existe una probabilidad de elegir cada una de las posibles alternativas, incorporando la idea de “racionalidad limitada” dado que los individuos no necesariamente seleccionan lo que es mejor para ellos [Macfadden (1981,pp.198)].

El primer modelo desarrollado bajo esta perspectiva es el de Luce (1959). Luce muestra que cuando las probabilidades de elección satisfacen los axiomas de elección, una escala puede ser definida sobre las alternativas, de tal forma, que las probabilidades de elección pueden ser derivadas de escalas de alternativas.

El modelo de Luce tiene como inconveniente que una nueva alternativa, que sea más que proporcional a las otras, reducirá las probabilidades de elección de alternativas existentes que son similares y causará reducciones menos que proporcionales en las probabilidades de elección en alternativas diferentes (Anderson, et. al, pp. 23-25).

Tversky (1972) propone que la elección de una alternativa puede verse como un proceso estocástico, en el cual, las alternativas son sucesivamente eliminadas hasta que quede solamente una, para esto supone que cada alternativa está compuesta por una lista de características las cuales son binarias en términos de que las alternativas poseen o no dichas características (por ejemplo, un automóvil puede o no tener aire acondicionado, sonido, etc.). En el caso de características que no sean estrictamente binarias (por ejemplo, el número de kilómetros recorridos por galón) Tversky sugiere usar niveles de umbrales, por ejemplo si el carro puede alcanzar más o menos de 30 Kilómetros por galón, para convertirlas en binarias.

A cada característica se asigna una escala positiva o valor de “utilidad” expresando la importancia de la característica para el individuo. El proceso de selección de una

[15] Los primeros trabajos originales al respecto, se deben a Von Newman y Morgensten(1944) posteriormente Pratt(1964) y Arrow(1971) desarrollan las famosas medidas de aversión al riesgo Arrow-Pratt, y, Machina(1985) trabaja la generalización de la utilidad esperada.

alternativa es el siguiente: Primero, una característica se selecciona y todas las alternativas que no posean esta característica son eliminadas del conjunto de elección. Segundo, se selecciona como el criterio para eliminar aquellas alternativas que quedan y así sucesivamente. Si una alternativa queda, ésta es la alternativa elegida por el individuo. Si varias alternativas quedan, ellas son elegidas con igual probabilidad. La probabilidad de seleccionar una característica como el criterio de elección de las alternativas que quedan depende de la escala de valores. Como podrán existir secuencias de eliminación diferentes, la probabilidad de elegir una alternativa particular es la suma de las probabilidades de todas las secuencias que finalizan con esta alternativa. Esta aproximación, es muy parecida a la ordenación de preferencias lexicográficas, sin embargo, es diferente debido a que el orden de selección de las características es aleatorio mientras éste viene dado *a priori* en el modelo lexicográfico.

Para ilustrar el proceso de eliminación de Tversky considere el siguiente ejemplo. Existen tres alternativas. A = $\{ a, b, c \}$ y siete características i = 1,2,...,7. Cada característica se asocia con una escala de utilidad μ_i. Las alternativas en A son discretas, por lo cual

$a = (U_1, 0 , 0 , U_4 , U_5 , 0 , U_7)$
$b = (0 , U_2 , 0 , U_4 , 0 , U_6 , U_7)$
$c = (0 , 0 , U_3 , 0 , U_5 , U_6 , U_7)$

Donde cero indica que la alternativa no posee la característica. Dado que la última característica, U_7, está presente en todas las alternativas, ésta no es considerada en el proceso de eliminación. Simplificando, a partir de K = $\sum_{i=1}^{S} U_i$ deberá mostrarse como $P_i (a)$ se determina. La alternativa a puede ser elegida si la primera característica es seleccionada dado que no está presente en b o en c. Este evento se debe asumir que ocurre con probabilidad $\frac{U_1}{\mathrm{K}}$. Si la característica cuarta o quinta es una de las seleccionadas, entonces, la alternativa a podrá ser seleccionada como la probabilidad de seleccionar la cuarta característica que tiene una probabilidad $\frac{U_4}{\mathrm{K}}$, c es eliminado pues no posee esta característica, la alternativa a se elegirá entonces con probabilidad P(a, b). La quinta característica es seleccionada con probabilidad $\frac{U_5}{\mathrm{K}}$ y si este evento ocurre b es eliminado y a es elegido entonces con probabilidad P(a , c). Dado que el proceso de selección de la primera característica es un evento mutuamente excluyente, tendremos

(7.1) $$P_A(a) = \frac{\left[U_1 + U_4 P(a,b) + U_5 P(a,c)\right]}{K}$$

Donde $P(a,b) = \dfrac{U_1 + U_5}{U_1 + U_2 + U_5 + U_6}$ y $P(a,c) = \dfrac{U_1 + U_4}{U_1 + U_3 + U_4 + U_6}$

Observe que en P(*a*, *b*) la característica cuatro es común en *a* y *b* por lo cual se elimina. De igual forma en P(*a*, *c*) la característica cinco es común en *b* y *c*. El procedimiento mostrado por Tversky se resume en los siguientes pasos

Paso 1: Elimine las características comunes a todas las alternativas.
Paso 2: Seleccione una de las características que permanecen.
Paso 3: Elimine las alternativas que no poseen esta característica.
Paso 4: Deténgase si las alternativas que quedan tienen la misma característica, de otra forma, regrese al paso 2.

Formalmente, suponga que existe una función U no negativa que especifica la utilidad para cada característica y denótese S como el número de características que están presentes después de haber eliminado las características comunes a las alternativas en el conjunto de elección $S \subseteq A$. Finalmente sea S_i el conjunto de las alternativas contenidas en S que contienen las características *i*, $i = 1, 2, ..., S$. En el modelo de Eliminación por Aspectos (EBA) propuesto por Tversky, la probabilidad de que la alternativa $a \in S$ sea elegida vendrá dada por

(7.2) $$P_S(a) = \sum_{i=1}^{S} \frac{U_i}{\sum_{j=1}^{S} U_j} P_{si}(a)$$

Cuando todas las características son comunes a todas las alternativas en S y $P_S(a) = \dfrac{1}{|S|}$ donde $|S|$ es el número de elementos en S. Como se puede observar (7.2) es recursivo, esto es, $P_S(a)$ es el peso de la suma de las probabilidades $P_{si}(a)$ donde *a* ha sido elegido del conjunto de S_i alternativas teniendo las características *i* en común, i =

1, 2, ..., S. Los pesos $\frac{U_i}{\sum_{j=1}^{S} U_j}$ k, representan las probabilidades de seleccionar las características $j = 1, 2, ..., S$ y las probabilidades $P_{si}(a)$ son definidas por (7.2).

7.1.2 Modelos con utilidad estocástica

Existen dos versiones tradicionales de los modelos de utilidad estocástica. El primero proviene de Thurstone a partir de la teoría sicológica de la elección individual y el segundo proviene de Macfadden en la versión económica de la elección discreta.

El modelo de Thurstone tiene su origen en una serie de experimentos donde se les preguntaba a los individuos acerca de comparar intensidades de estímulos físicos, por ejemplo, el rango de tonos en términos del ruido. Dada la variabilidad en las respuestas, Thurstone propone que un estímulo provoca una "sensación" o un estado sicológico que es la realización de una variable aleatoria. De esta forma, "las utilidades se asumen que varían de un momento a otro momento, y el proceso de decisión consiste en una regla fija de escoger la alternativa con la mayor utilidad momentánea" Edgell y Geisler (pp.266).

Considere un individuo compuesto por varios homo-económicos. Cada tipo obedece a la teoría neoclásica, y dependiendo del estado de la mente del individuo un homo-económico en particular es seleccionado, por lo cual, el individuo se comporta racionalmente de acuerdo a una utilidad determinística. De acuerdo a esta aproximación, los valores de las alternativas en A deberán ser considerados como variables aleatorias, $U_1 + \varepsilon_1$,..., $U_n + \varepsilon_n$, las variables U_1 ,..., U_n son escalas de valores asociados a alternativas constantes mientras que ε_1 ,..., ε_n son variables aleatorias. Suponga que la función de distribución acumulativa de $\varepsilon = (\varepsilon_1,.., \varepsilon_n)$ es continua con respecto a la medida de Lebesque [$\Pr(\varepsilon_i - \varepsilon_j = \alpha = 0)$ $\forall$ α constante e $i \neq j$]. Si ε_i tiene media cero (de lo contrario la media de ε_i puede ser adicionada al escalar μ_i), las probabilidades del conjunto de elección vienen determinadas por

$$(7.3) \quad P_A(i) = \Pr\left[\mu_i + \varepsilon_i = \max_{j=1,\ldots,n}(U_j + \varepsilon_j)\right] \quad , \; i = 1,\ldots,n$$

La versión de Macfadden es conceptualmente diferente: considere una población de individuos haciendo la misma elección sobre el conjunto A y determine la fracción de la población que elige una alternativa determinada. La población total puede ser dividida en subpoblaciones tales que cada subpoblación sea homogénea con respecto a ciertos factores socioeconómicos observables (ingreso, edad, profesión, etc.). Cada individuo se supone que tiene una función de utilidad determinista U definida sobre A. Sin embargo, el modelador podrá observar imperfectamente las características que influencian las decisiones individuales y entonces tendrá un conocimiento imperfecto de la función de utilidad U. La función U se descompone en dos partes, una parte, μ, que representa la parte conocida de la utilidad y definida sobre las características observables, y la otra parte, e, que representa la diferencia entre U y μ. Para cada $i = 1, 2, ..., n$ la utilidad deseada de la alternativa *i* puede escribirse como

$$(7.4) \quad U_i = \mu_i + e_i \,, \qquad e_i \sim (0, \sigma^2).$$

Pensando que el comportamiento es determinístico, para el modelador, es imposible predecir exactamente la elección del individuo dado que él no puede ser observado. Esto es posible, ya que cada miembro difiere de los otros en la subpoblación considerada con respecto a las características no observables y los factores que influencian al individuo en su elección. De esta forma, U_i puede ser modelado como una variable aleatoria

$$(7.5) \quad \bar{U}_i = U_i + \varepsilon_i$$

Aquí $\bar{U}_i$ es la utilidad observable y refleja las preferencias de la subpoblación para la i- ésima alternativa. ε_i toma en cuenta las diferencias entre los gustos en los individuos de la subpoblación. La probabilidad de que un individuo aleatoriamente seleccione la alternativa *i* viene dada por

$$(7.6) \quad P_A(i) = \Pr(\bar{U}_i = \max_{j=1,\ldots,n} \bar{U}_j), \; i = 1,\ldots,n$$

Existen diferentes fuentes de incertidumbre: características no observables, variaciones no observables en las utilidades individuales, desconocimiento de la cantidad de características observables y finalmente, especificación funcional errónea [Manski (1977)]. El modelador, podría estar satisfecho si pudiera agrupar una población y si en la función puede incluir los principales factores observables. Podemos observar entonces que aún cuando entre (7.3) y (7.6) existen diferencias epistemológicas, a un nivel agregado, como observa Macfadden las “variaciones intraindividuales e interindividuales en los gustos son indistinguibles en su efecto sobre la distribución observada en la demanda”(pp.205), lo cual indica que (7.3) y (7.6) tienen las mismas probabilidades de elección. Consideremos una población de N son estadísticamente idénticos e independientes, lo que significa que las elecciones están gobernadas por la misma distribución de probabilidades, aunque las elecciones actuales difieran entre los individuos. Esto implica que la probabilidad de un individuo de elegir una alternativa particular es independiente de las elecciones realizadas por otros individuos. Bajo este supuesto la distribución de elección es multinomial con media $\bar{X}_i$ dada por $\bar{X}_i$ = N P_A(i), i = 1,2,...,n, la cual es la demanda esperada para la alternativa i. Si N es suficientemente grande, $\bar{X}_i$ es una buena aproximación de la demanda agregada dado que la desviación estándar de la distribución de la función de demanda decrece a una razón $1/\sqrt{N}$.

7.2 Funciones de densidad para elecciones discretas

Para determinar las probabilidades de elección, se deberá especificar la distribución de las variables aleatorias ε_i. Estas probabilidades para una serie de distribuciones, como la del modelo de probabilidad lineal, del Logit, el Probit y el Multinomial, se conocen como modelos de elección binaria. Sea f(X), la densidad ε donde $\varepsilon \equiv \varepsilon_2 - \varepsilon_1$ tendremos

(7.7) $$P_A(1) = \int_{-\infty}^{U_1 - U_2} f(x)dx$$

Donde (7.7) es el valor de la función de densidad acumulativa de ε sobre $U_1 - U_2$. Supongamos que ε esté distribuido uniformemente en el intervalo [-L , L], entonces

(7.8) $$f(x) = \begin{cases} \dfrac{1}{2L}, & Si\ X \in [-L, L] \\ 0, & De\quad otra\quad forma \end{cases}$$

Por lo cual, la probabilidad de elección para la alternativa 1 será

(7.9) $$P_A(1) = \begin{cases} 0, & Si\ U_1 - U_2 < -L \\ \dfrac{U_1 - U_2}{2L} + \dfrac{1}{2}, & Si\quad -L \le U_1 - U_2 \le L \\ 1, & Si\quad L < U_1 - U_2 \end{cases}$$

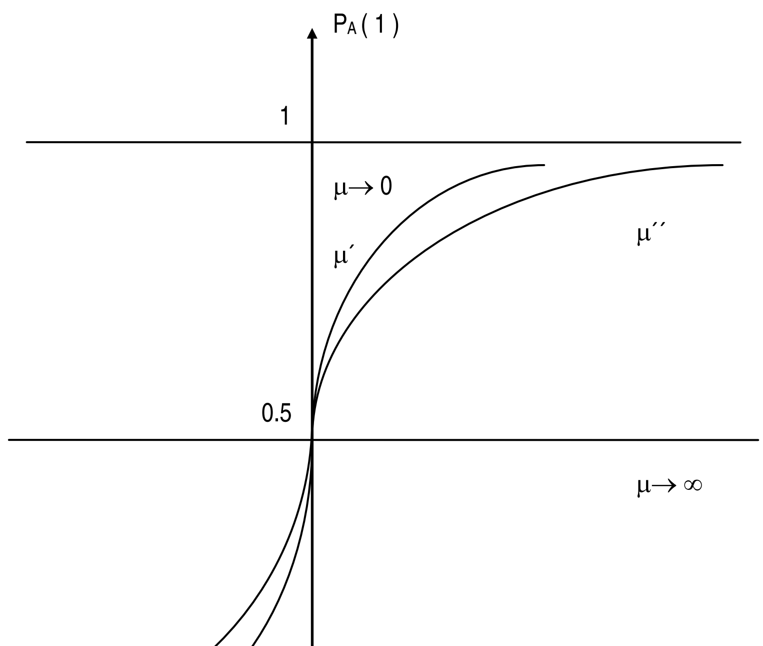
Este modelo es conocido como el *Modelo de Probabilidad Lineal*, ya que la probabilidad (7.9) es lineal sobre el intervalo [-L , L]. Consideremos ahora que ε se distribuya normalmente y que ε_1 y ε_2 se pueden ver como variables independientes no observables. Suponga ahora que $\varepsilon_1 \sim (0,\sigma_1^2)$ y $\varepsilon_2 \sim (0,\sigma_2^2)$ y que la covarianza entre ε_1 y ε_2 estará dada por σ_{12} , entonces $\varepsilon \sim (0 , \sigma_1^2 + \sigma_2^2 - 2\sigma_{12})$ y obtenemos

$$(7.10) \quad P_A(1) = \frac{1}{\sqrt{2\pi\sigma}} \int_{-\infty}^{U_1+U_2} e^{\left(\frac{-X^2}{2\sigma^2}\right)} dx$$

El cual es un *Probit.* Si asumimos que ε se distribuye logísticamente, la función de la distribución para ε viene dada por $\frac{1}{1+e^{(\frac{-X}{\mu})}}$, la media de ε es cero y su varianza es $\frac{\pi^2\mu^2}{3}$. La probabilidad de elegir 1 viene dada por el Logit definido como

$$(7.11) \quad P_A(1) = \frac{1}{1+e^{\left[\frac{(-U_1-U_2)}{\mu}\right]}} = \frac{e^{(\frac{U_1}{\mu})}}{e^{(\frac{U_1}{\mu})}+e^{(\frac{U_2}{\mu})}}$$

Gráfica 7.1. Modelo probit



La gráfica anterior se define para valores de $0 < \mu' < \mu'' < \infty$. La pendiente sigmoidal de la curva es más pronunciada para valores mayores de μ. La curva tiene un punto de inflexión en μ' - μ'' donde $P_A(1) = ½$. $P_A(1)$ es decreciente (creciente) en μ cuando U_1 es mayor, menor, que U_2 . Un modelo de elección determinístico inicialmente se puede definir entre los siguientes tres modelos: *Probabilidad Lineal*, *Probit* y *Logit*. Para $L \to 0$ en (7.9) con $U_1 \neq U_2$, $P_A(i) = 1$ si y solo si $U_i = \max\{U_1, U_2\}$. Este también es el límite para el modelo Probit cuando $\sigma \to 0$ y en el Logit cuando $\mu \to 0$. Cuando L, σ y μ tienden a infinito, el comportamiento individual es impredecible completamente y entonces $P_A(1) = P_A(2) = ½$.

7.3 Funciones de utilidad y funciones indirectas de utilidad

Un individuo consume de acuerdo a una función de utilidad definida sobre los bienes $X_1,......X_n$ y Z, siendo Z el numerario. La utilidad del consumidor podría también depender de una serie de atributos de los bienes X's denotadas por $b_1,.....b_n$, los cuales son tomados exógenamente, adicionalmente las preferencias podrían depender de características propias como la educación, la raza, la cultura, la edad,.., etc., representadas por el vector S [16]. De esta forma, la función de utilidad se escribirá compactamente como U(X, b, Z, S, ε) donde ε es una variable aleatoria con alguna función de densidad conjunta $f_\in (\varepsilon_1,..... \varepsilon_n)$ la cual induce una densidad sobre U. Un supuesto adicional, consiste en que los X's sean mutuamente excluyentes, esto es, un individuo no puede rentar su casa y vivir allí mismo por lo cual $X_i X_j = 0 \; \forall \; i \neq j$.

Suponga que el consumidor decide consumir solamente el bien j, condicionado sobre esta decisión, su función de utilidad será una función de X_j y Z, por lo cual la utilidad vendrá definida por

$$\bar{U} = U\left(0,\ldots,0,X_j,0,\ldots,0.b_1,\ldots,b_n,Z,S,\varepsilon\right) \quad (7.12)$$

[16] La introducción de S se debe originalmente a los trabajos de Pollack y Wales(1979).

Suponiendo que existe débil complementariedad. La utilidad directa condicional puede escribirse como

(7.13) $$\bar{U} = U_j\left(X_j, b_j, Z, S, \varepsilon\right)$$

El consumidor maximiza sobre la restricción tradicional de presupuesto y no negatividad para X_j y Z. Asumiendo estricta cuasiconcavidad en $\bar{U}_j$, con relación a X_j y Z la solución viene dada por $X_j > 0$. Las funciones de demandas Marshallianas ordinarias serán

(7.14) $$X_j = \bar{X}_j\left(p_j, b_j, Y, S, \varepsilon\right)$$

$$Z = \bar{Z}\left(p_j, b_j, Y, S, \varepsilon\right) = Y - p_j\,\bar{X}_j\left(p_j, b_j, Y, S, \varepsilon\right)$$

Y, la función indirecta de utilidad vendrá dada por

(7.15)

$$\bar{V}_j\left(p_j, b_j, Y, S, \varepsilon\right) \equiv \bar{U}_j\left(\bar{X}_j(p_j, \bar{U}_j, Y, S, \varepsilon), b_j, \bar{Z}_j\left(p_j, b_j, Y, S, \varepsilon\right), S, \varepsilon\right)$$

Siendo $\bar{V}_j$ cuasiconvexa, decreciente en p_j y creciente en Y. Usando la identidad de Roy para encontrar la demanda marshalliana obtenemos

(7.16) $$\bar{X}_j\left(p_j, b_j, Y, S, \varepsilon\right) = \frac{-\partial \bar{V}_j\left(p_j, b_j, Y, S, \varepsilon\right) / \partial p_j}{\partial \bar{V}_j\left(p_j, b_j, Y, S, \varepsilon\right) / \partial Y}$$

Las Cantidades $\bar{X}_j, \bar{Z}\, y\, \bar{V}_j$ son cantidades conocidas para el consumidor pero dado que las preferencias se observan incompletamente, estas serán variables aleatorias desde el punto de vista del investigador. Sea $f_{\bar{V}}(\bar{V}_1, .., \bar{V}_n)$ la densidad conjunta de

$\bar{V}_1,..,\bar{V}_n$ denotada por $f_\varepsilon(.)$ y $F_{\bar{V}}(.)$ la distribución acumulativa. Si la elección discreta, por medio de la cual, un bien puede ser seleccionado se representa por un conjunto binario de índices de valores de la forma

(7.17) $\delta_1,\ldots,\delta_N = \begin{cases} \delta_j = 1, & si \quad X_j > 0 \\ \delta_j = 0, & si \quad X_j = 0 \end{cases}$

Entonces, la elección puede ser expresada en términos de las funciones condicionadas de utilidad indirecta como

(7.18) $\delta_j\ (p,b,Y,S,\varepsilon) = \begin{cases} 1 \quad ; \quad si \quad : \bar{V}_j(p_j, b_j, Y, S, \varepsilon) \geq \bar{V}_i(p_i, b_i, Y, S, \varepsilon) \quad \forall i \\ 0 \qquad\qquad\qquad\qquad\qquad\qquad , \quad De \quad otra \quad forma \end{cases}$

Para el observador, los índices discretos de elección son variables de medición $E(\delta_j)=\pi_j$ que viene dados por

$$\Pr ob\left\{\bar{V}_j\left(p_j, b_j, Y, S, \varepsilon\right) \geq \bar{V}_i\left(p_i, b_i, Y, S, \varepsilon\right) \forall i\right\}$$

(7.19) $\pi_j(p,b,Y,S,\varepsilon) = \displaystyle\int_{-\infty}^{\infty} F_{\bar{V}}^{\,j}(u,\ldots,u)\partial u$

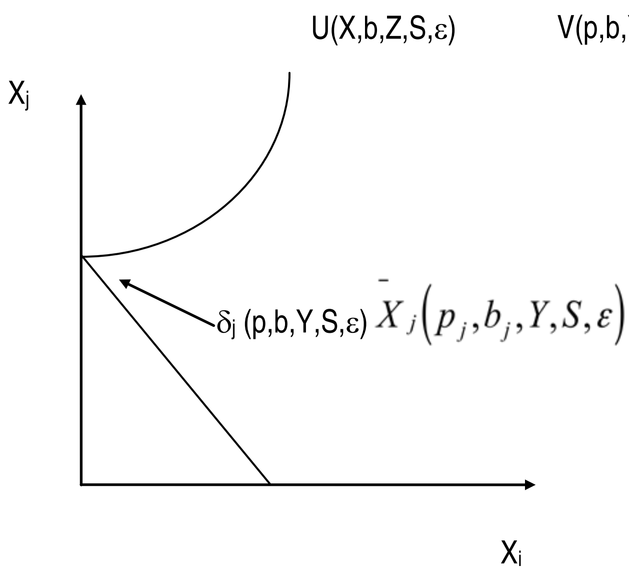
Donde $F_{\bar{V}}^{\,j}$ es la derivada de $F_{\bar{V}}$ ($\cdot$) con respecto a su i-ésimo argumento. De (7.18) nosotros sabemos que las demandas están condicionadas a la existencia de los índices, por lo cual

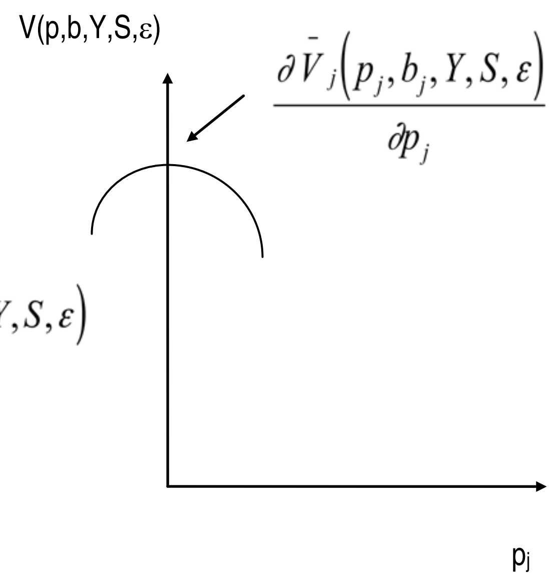
(7.20) $X_j\ (p,b,Y,S,\varepsilon) = \delta_j\ (p,b,Y,S,\varepsilon)\ \bar{X}_j\left(p_j, b_j, Y, S, \varepsilon\right)$ $\forall$ j=1,.....,N

$$V(p,b,Y,S,\varepsilon) = \max\left[\bar{V}_1\left(p_1, b_1, Y, S, \varepsilon\right), \ldots\ldots\ldots, \bar{V}_N\left(p_N, b_N, Y, S, \varepsilon\right)\right]$$

Dado que $V(\cdot)$ es cuasiconvexa, la solución de esquina puede representarse como

Gráfica 7.2 **Gráfica 7.3**





En el punto A, gráfica (7.3), la utilidad indirecta al precio j es máxima. A estos precios se cumple (7.17) y, dada la función de utilidad, la demanda de Xj cuando $\delta_j = 1$ se obtiene en el punto A, gráfica (7.2). Para encontrar las distribuciones de probabilidades de X_j y V, suponga que existe un conjunto $A_j \equiv$ $\{\varepsilon| \bar{V}_j\left(p_j, b_j, Y, S, \varepsilon\right) \geq \bar{V}_i\left(p_i, b_i, Y, S, \varepsilon\right) \forall i\}$; j=1…..N; de f_ε se puede construir como $f_{\varepsilon|\varepsilon \in A_j}$, esto es, la densidad conjunta de $\varepsilon_1, \ldots, \varepsilon_N$ dado que $\varepsilon \in A_j$ y cuando el bien j es seleccionado, la probabilidad de densidad de $\bar{X}_j$, $f_{X_j|\varepsilon \in A_j}(X) = \Pr\{X_j = X \mid \varepsilon \in A_j\}$ puede ser obtenida de $f_{\varepsilon|\varepsilon \in A_j}$. De esta forma, la probabilidad de densidad de X_j, $f_{Xj}(X) =$
Pr { X_j=X} tomará la forma

$$(7.21)\quad f_{x_j}(X) = \begin{cases} 1 - \Pi_j ; X = 0 \\ \Pi_j f_{X_j|\varepsilon \in A_j}(X) ; X > 0 \end{cases}$$

Así, una vez especificado el modelo, uno puede construir las densidades $F_{\bar{V}}$ y $f_{\varepsilon|\varepsilon\in A_j}$ las cuales son usadas en la función de probabilidades de elección discreta y las densidades condicionadas y no condicionadas de las X_j´s. Por ejemplo, si N fuese igual a dos se puede establecer que

$$(7.22)\quad X = \begin{cases} \dfrac{-\partial \bar{V}_1(p_1, b_1, Y, S, \varepsilon)/\partial p_1}{\partial \bar{V}_1(p_1, b_1, Y, S, \varepsilon)/\partial Y}; Si \quad \bar{V}_1(p_1, b_1, Y, S, \varepsilon) \geq \bar{V}_2(p_2, b_2, Y, S, \varepsilon) \\ \\ \dfrac{-\partial \bar{V}_2(p_2, b_2, Y, S, \varepsilon)/\partial p_2}{\partial \bar{V}_2(p_2, b_2, Y, S, \varepsilon)/\partial Y}; \quad De \quad otra \quad forma \end{cases}$$

Cuando hay T individuos y j^* es el índice seleccionado por el i- ésimo individuo y X_t^* su consumo observado sobre este índice para un individuo, la función de verosimilitud para la muestra vendrá dada por

$$(7.23)\quad L = \prod_{t=1}^{T} f_{X_{j^*t}}(X_t^*) = \prod_{t=1}^{T} \left\{ f_{X_{j^*t}}|_{\varepsilon\in A_{j^*t}}(X_t^*) \Pi_{j^*t} \right\}$$

En principio (7.23) puede ser obtenido por máxima verosimilitud, sin embargo Haneman(1984) sostiene que en la práctica las ecuaciones normales podrían tener múltiples raíces y, a menos que se comience con un estimador inicial consistente, no existe garantía de convergencia a un máximo global. Usualmente, se sugiere el procedimiento de dos etapas de Heckman, esto es, encontrar por máxima verosimilitud, usando un Logit, el conjunto de parámetros que serán consistentes pero no eficientes dado que ellos ignoran la información contenida en datos continuos; con estos parámetros uno puede entonces obtener estimadores consistentes de $(\lambda_i / \mu)_t$ y entonces realizar una regresión para elecciones continuas donde el modelo vendría determinado por

$$(7.24)\quad LnP_{j^*}X_{j^*} = Ln\theta_t + \eta Y_t + (\rho - 1)(\mu Ln \sum e^{\left(\frac{\lambda_i}{\mu}\right)_t} + 0.5722\mu) + V_i; t = 1,..,T$$

Con V_t´s iid, EV(μ,-0.57222μ) E(V_t)=0 Var V_t= $\dfrac{\Pi^2 \mu^2}{6}$

Finalmente, se puede usar la subrutina Maxlik del programa gauss y obtener estimadores eficientes, o usar el programa LIMDEP.

7.4 lecciones discretas con productos diferenciados

El consumidor representativo es un agente función, cuya utilidad nos muestra un conjunto de preferencias diversas. Ya que en la práctica, los consumidores tienden a comprar, solamente una, o en todo caso muy pocas de las variantes de un producto que se le ofrece, el consumidor representativo ha sido bastante criticado[17]. Como bien lo han señalado Archival, Eaton y Lipsey(1986) la cuestión sobre cuando el consumidor representativo puede constituir una descripción agregada válida de una población de consumidores caracterizados por elecciones discretas en el ámbito individual es un punto de discusión abierto.

En este sentido, el interés principal de esta sección, consistirá en mostrar como encontrar un consumidor representativo para una población de consumidores que realizan elecciones discretas dados unos supuestos sobre el proceso de elección o la elección de probabilidades y, cuales serían las propiedades de la función de utilidad correspondiente.

Suponga que existen m+1 bienes y N consumidores estadísticamente idénticos e independientes. El bien 0 es perfectamente divisible y se toma como numerario. Los bienes i= 1, 2, ..., m son los variantes de un producto diferenciado a los precios $p_1, ..., p_m$. Sea A el conjunto de variantes donde la variante i se asocia a un índice de calidad a_i [18]. Cada consumidor tiene un ingreso real Y con el que puede comprar una unidad de una variante singular. Asuma que $0 \leq p_i \leq Y$ lo cual asegura que cada variante es alcanzable por todos los consumidores. Suponga también que la función indirecta de utilidad derivada de la compra de la variante i viene dada por la siguiente forma aditiva

(7.25) $\bar{v}_i = Y - p_i + a_i + \varepsilon_i$; $i = 1, ..., m$

Donde ε_i son las n variables aleatorias con una densidad f(x) con $x = x_1, ..., x_n$. La probabilidad de que un consumidor seleccione la variante i, viene dada por

(7.26) $P_i = \text{Prob}[\ a_i - p_i + \varepsilon_i = \max_{j=1,..,m} (a_j - p_j + \varepsilon_j)]$

$= \text{Prob}[\ a_i - p_i + \varepsilon_i = \max_{j=1,..,m} (a_j - p_j + \varepsilon_j)]$

[17] Lo que produce soluciones de esquina o en el caso de que el consumidor adquiera más de una variante soluciones interiores.

[18] El índice de calidad a_i resume todas las características observables de la variante i. De esta forma, si la variante i tiene varias características observables y los consumidores tienen valoraciones idénticas e independientes entonces a_i es el producto de las características observables con las valoraciones de los consumidores.

Lo que se conoce como utilidad aleatoria(LRUM). La demanda esperada para la variante i, se define como

(7.27) $X_i = NP_i$

La estructura particular de LRUM depende del modelo en cuestión. Suponga un modelo multinomial, entonces la probabilidad de que un individuo elija una variante i viene dada por

$$(7.28)\ P_i\,(a-p) = \int_{-\infty}^{\infty} \int_{-\infty}^{u_i+x-u_1} \cdot\ \cdot\ \cdot \int_{-\infty}^{u_i+x-u_m} \times f(z_1 \ldots\ x \ldots\ z_m)\partial z_m \ldots\ \partial z_1 \partial x$$

La función (7.28) deberá satisfacer las siguientes propiedades

Primera Propiedad: $\int_{-\infty}^{p_2} \ldots \int_{-\infty}^{p_m} \frac{\partial^{m-1} P_1(a_1-p_1,\ldots,a_m-p_m)}{\partial p_2 \ldots \partial p_m} \partial p_m \ldots \partial p_2$. Esto deberá cumplirse también para P_m.

Segunda Propiedad: $\frac{\partial P_i}{\partial p_j} = \frac{\partial P_j}{\partial p_i}; i, j = 1....m, i \neq j.$

Tercera propiedad: Para algún $\theta \in \Re$ $P_i\,[a-(p+\theta)] = P_i\,(a-p)$, i = 1...m. Donde $p+\theta$ es un vector con componentes $p_j+\theta$, j= 1,...,m.

Cuarta propiedad: $\underset{(a_i-p_i)\to\infty}{Lim}$ $P_i = 1$; i=1...m, con $a_j - p_j$ finito y $j \neq i$.

Por la primera propiedad se tiene que las variantes de los productos diferenciados son débilmente sustitutas, esto significa que $\frac{\partial P_i}{\partial p_j} > 0; i, j = 1....m, i \neq j.$ Por la segunda propiedad se garantiza la igualdad de las derivadas cruzadas de los precios. De (7.27) nosotros conocemos que la demanda es independiente del ingreso. Si estas demandas son el resultado de la maximización individual, entonces las derivadas cruzadas representan los efectos sustitución en la matriz de Slutzky. La tercera propiedad significa que la probabilidad depende solamente de las diferencias en precios. La cuarta propiedad significa que todos lo N consumidores eligen una variante i con certeza cuando esta variante es infinitamente atractiva en términos de la medida de utilidad { $u_i = a_i - p_i$ } cuando las otras utilidades son finitas.

De la primera propiedad, se puede deducir la función de distribución acumulativa y la función de densidad, las cuales se definen como

(7.29) $F(z_1,\ldots, z_m) = \int_{-\infty}^{z_1} P_1(0, t - z_2, \ldots, t - z_m)\psi(t)\partial t \geq 0$

Donde $\psi(t)$ será una probabilidad de densidad univariada y derivando (7.29) encontramos la siguiente función de densidad

(7.30) $f(z_1,\ldots, z_m) = \dfrac{\partial^{m-1} P_1(0, z_1 - z_2, \ldots, z_1 - z_m)}{\partial z_2 \ldots \partial z_m}\psi(z_1) \geq 0$

Combinando (7.30) con (7.28) genera las probabilidades $P_1 \ldots P_m$.

7.4.1 La función de demanda para un continuo de consumidores

Considere un continuo de consumidores igual a N cada uno con gustos determinísticos. Un consumidor tiene un ingreso Y y compra una unidad de una variante de un producto diferenciado. La función de utilidad indirecta condicionada viene dada por

(7.31) $V_i = Y - p_i + a_i + e_i$; $i = 1, \ldots, m$

Donde $e_1 \ldots e_m$ describe las valoraciones de un consumidor para un conjunto de variantes. Cada conjunto de valoraciones define un tipo de consumidor. Aunque los índices de cualidades $a_1,\ldots,a_m$ son comunes a todos los consumidores las valoraciones son individuales y toman valores diferentes para consumidores diferentes. Las valoraciones se distribuyen sobre $\Re^m$ de acuerdo a la siguiente función de densidad

(7.32) $g(e) \equiv Nf(e)$

Donde $f(e)$ es la densidad de probabilidades de $\varepsilon_1 \ldots \varepsilon_m$. Por construcción $\int_{\Re^m} g(e)de = N$. El segmento de mercado para la variante i se puede definir como

(7.33) $S_i \equiv \{ e \in \Re^m ; V_i(e) \geq V_j(e), j = 1,\ldots,m. \}$

Mostrando el conjunto de tipos para los cuales la variante i es débilmente preferida sobre todos los otros. De esta forma, la demanda total para i viene definida por

(7.34) $X_i = \int_{S_i} g(e)de$; $i = 1, \ldots, m$

La cual es igual al total de consumidores en el segmento de mercado para la variante i. Es importante observar que la función de densidad del consumidor cumple el mismo papel que la función de densidad de probabilidades (7.26). Sin embargo, la diferencia entre (7.26) y (7.34) es sustancial, pues la función derivada de la integración de las utilidades máximas sobre la densidad de probabilidades (7.26) puede ser interpretada como la utilidad esperada del consumidor mientras que la integral de las utilidades máximas de los individuos sobre la densidad de tipos producirá la función de riqueza a partir de la utilidad. Esta última idea se desarrollará a continuación. Suponga que la función de riqueza derivada de la utilidad se define como

(7.35) $\mathbf{V} = W + \int_{\Re^m} max(a_i - p_i + e_i) g(e) \partial e \quad ; \quad i = 1,...,m$

Donde W = NY es el ingreso agregado. Note primero que **V** satisface las características propias de la función indirecta de utilidad (cap. 2, sec.2.4.1,pag 20)

Primera propiedad : **V** es continua en p_i y W.
Segunda propiedad : **V** no es creciente en p_i y es creciente en W.
Tercera propiedad : **V** es convexa en p_i.
Cuarta propiedad : **V** es homogénea de grado cero en todos los precios e ingreso.

La propiedad tercera se sigue del hecho de que la integral de funciones convexas es convexa [19], la cuarta propiedad se mantiene en tanto $p_1,..,p_m$ y W se encuentran en términos reales al dividir por el precio del bien 0, el cual es el bien numerario. Dado que las utilidades individuales son lineales en el ingreso, la utilidad marginal del ingreso es igual a uno para todos los individuos. Lo cual significa que un cambio en cualidades o precios que aumentan a **V** aumentará las transferencias de ingreso en todos los consumidores situándoles mejor que antes del cambio en precios o ingreso. El segundo término de (7.35) puede usarse para cuantificar los cambios en el excedente del consumidor atribuible a cambios en precios y cualidades, lo cual, se puede ver también como una medida del beneficio del consumidor de introducir una nueva variante.

7.4.2 El consumidor representativo multinomial

Suponga un consumidor con una función de utilidad aleatoria $\bar{v}_i$ = Y – p_i + a_i + ε_i . Las demandas esperadas vendrán dadas como

[19] Cuando se normaliza con el precio de la bien 0 **V** es cuasiconvexa en todos los precios (p_0, $p_1,...,p_n$)[Anderson, de Palma y Thisse(1992),(1995)].

(7.36) $X_i = N \frac{\exp[(a_i - p_i)/\mu]}{\sum_{j=1}^{m} \exp[(a_j - p_j)/\mu]}$; i = 1,...,m.

A través de (7.35) la función de utilidad indirecta para un consumidor representativo, se puede definir como

(7.37) $\mathbf{V} = W + N\mu \text{Ln}\left[\sum_{j=1}^{m} \exp\left(\frac{a_j - p_j}{\mu}\right)\right]$

Con el fin de ilustrar estas funciones, suponga el caso de que m=2 las demandas vendrán dadas por

(7.36.1) $X_i = N \frac{\exp[(a_i - p_i)/\mu]}{\exp[(a_1 - p_1)/\mu] + \exp[(a_2 - p_2)/\mu]}$; i = 1, 2.

Dado que las demandas para las variantes son independientes del ingreso y dependen multiplicativamente de N, debe esperarse una función de utilidad directa de la forma

(7.36.2) $\mathbf{U} = Nu(x_1, 1 - x_1) + X_0$

Donde $x_1 \equiv X_1 / N$ y X_0 es el consumo del bien numerario. Suponga que la restricción presupuestaria viene dada por

(7.36.3) $Y = \sum_{i=1}^{2} p_i X_i + X_0$

Y sustituyendo X_0 de (7.36.3) en (7.36.2) se obtiene

(7.36.4) $\mathbf{U} = Nu(x_1, 1 - x_1) + Y - p_1 N x_1 - p_2 N(1 - x_1)$

Tomando las condiciones de primer orden, se encuentra que

(7.36.5) $\frac{\partial u}{\partial x_1} = (p_1 - p_2)$

De (7.36.1) se observa la condición

(7.35.6) $$\frac{x_1}{1-x_1} = \exp\left[\frac{(a_1 - p_1) - (a_2 - p_2)}{\mu}\right]$$

Que también se puede ver como

(7.36.7) $p_1 - p_2 = a_1 - a_2 - \mu[\text{Ln } x_1 - \text{Ln } (1 - x_1)]$

Sustituyendo (7.36.7) en (7.36.5) e integrando para recobrar u

(7.36.8) $u(x_1, 1- x_1) = a_1x_1 + a_2(1 - x_1) - \mu[x_1\text{Ln } x_1 - (1 - x_1)\text{Ln } (1 - x_1)]$

Donde la constante de integración deberá ser igual a a_2. A través de (7.36.2) se tiene que la función de utilidad indirecta será

(7.36.9) $\mathbf{U} = N(a_1x_1 + a_2(1 - x_1) - \mu[x_1\text{Ln } x_1 - (1 - x_1)\text{Ln } (1 - x_1)] + X_0$

Dado que cada consumidor compra una unidad de una variante en el modelo de utilidad aleatorio, el consumidor representativo comprará N unidades. Esto bajo el supuesto de que la utilidad tiende a - ∞ para otros consumos.

Finalmente, una función de utilidad para un consumidor representativo con demandas tipo Logit multinomiales vendrá dada por

(7.38) $$U = \begin{cases} \sum_{i=1}^{m} a_i X_i - \mu \sum_{i=1}^{m} X_i (LnX_i - LnN) + X_0 & Sí \quad \sum_{i=1}^{m} X_i = N \\ -\infty & De \quad otra \quad forma \end{cases}$$

El Lagrangiano para el problema de la maximización del consumidor se define como

(7.39)
$$L = \sum_{i=1}^{m} a_i X_i - \mu \sum_{i=1}^{m} X_i (LnX_i - LnN) + X_0 + \lambda_1 (\sum_{i=1}^{m} X_i - N) + \lambda_2 (Y - X_0 - \sum_{i=1}^{m} X_i)$$

Donde λ_1 y λ_2 son los multiplicadores Lagranjianos de las restricciones. Dado que los precios e ingreso son tales que el consumo óptimo del bien numerario es positivo $\lambda_2 = 1$, el conjunto de primeras condiciones para X_i será

(7.40) $a_i - \mu(\text{Ln} x_i + 1) - p_i + \lambda_1 = 0, i = 1,...,m$

La condición (7.40) puede escribirse también como

(7.41) $x_i = \exp(-1 + \lambda_1/\mu) \exp(a_i - p_i) / \mu$, i = 1,...,m

Si hacemos $\sum_{i=1}^{m} X_i$ = 1 se encuentra que $\exp(1 - \lambda_1/\mu) = \sum_{i=1}^{m} \exp(a_i - p_i) / \mu$ de donde se sigue que las demandas multinomiales representadas por (7.36) son similares a (7.41).

7.5 Análisis de riqueza

En un análisis continuo se puede encontrar el excedente del consumidor a través de integrar la curva de demanda compensada entre dos precios. Sin embargo, en el análisis discreto existirán puntos de discontinuidad y no-diferenciación en la función indirecta de utilidad y en la función de gasto, por lo tanto, existirán problemas al integrar las funciones. La demanda se podría modelar, como observan Small y Rosen(1981) a través de tres aproximaciones

Primero, pensar que los bienes puedan ser disponibles en cantidades continuas, pero solamente en un pequeño número de variedades mutuamente excluyentes, un ejemplo sería una casa: usted podría alquilarla o vivir en ella, pero solamente la posesión de la misma, le daría una cantidad continua de usos para ser consumidas en ella, como clavar puntillas en ella para colgar cuadros, pintarla de todos los colores y las veces que usted quisiera, etc.

Segundo, los bienes pueden ser disponibles en unidades discretas entre más consumidores elijan una o dos unidades, como en el caso del transporte para trabajadores, los colegios, las antenas parabólicas, y en general muchos bienes durables.

Tercero, los bienes pueden ser comprados en unidades discretas pero debido a las no-concavidades en la función de utilidad, llevaría al consumidor a elegir entre soluciones alternativas de esquina. Por ejemplo, podríamos tener dos o más televisores con diferentes programas cada uno y observarlos al mismo tiempo; sin embargo, ver un sólo programa podría generar mayor utilidad que ver dos programas al mismo tiempo.

A continuación, se analizará primero el caso de un modelo con 3 bienes, dos de los cuales son mutuamente excluyentes. Considere un consumidor con una función de utilidad $U=U(X_n,X_1,X_2)$ donde X_n es el bien numerario, la utilidad es finita siempre que X_1 o X_2 sean cero. Como en el caso estándar, la utilidad se maximiza con la restricción

(7.42) $X_n+P_1X_1+ P_2X_2 = Y$; $X_j \geq 0$ (j=1,2,n)

Por último, se requiere que $X_1 X_2 =0$ y asumiendo soluciones interiores, se llega al siguiente teorema

7.5.1 El teorema de Small y Rosen

Suponga un consumidor que maximiza sujeto a la restricción (7.42) con una función de utilidad dos veces diferenciable y estrictamente cuasiconcava. Asuma que U es finito siempre que X_1 o X_2 sean cero y, que U es estrictamente creciente en X_n y no decreciente en X_1 y X_2. Sea $e(P_1,P_2,U)$ el mínimo gasto requerido para alcanzar el nivel de utilidad U y sea $U^\circ = V(p_1^\circ, p_2, Y)$ el valor de la función indirecta de utilidad a los precios iniciales y el ingreso, entonces la variación compensada para un cambio de precios en p_1 de p_1° a p_f° viene definido como

(7.43) $$e(p_1^f, p_2, U^\circ) - e(p_1^\circ, p_2, U^\circ) = \int_{p_1^\circ}^{p_1^f} X_1^c(p_1, p_2, U^\circ)\partial p_1$$

Y, la función de demanda compensada para el bien 1 vendrá definida por

(7.44) $X_1^c(p_1,p_2,U^o)=X_1(p_1,p_2,e(p_1,p_2,U^o))$

Adicionando un continuo de consumidores se obtiene

(7.45) $$\Delta E = \int_{p_1^\circ}^{p_1^f} X_1^c(p_1, p_2, \{U^i\})\partial p_i$$

Donde U^i es el nivel de utilidad inicial para el i- ésimo consumidor y (7.45) es el excedente del consumidor. Considere ahora el caso de un bien comprado solamente en unidades discretas cuando una gran parte de los consumidores eligen una o dos unidades. Sea la función de utilidad $U(X_n,X_1)$ donde $X_n>0$ y X_1 puede tomar valores binarios de 0 y 1. La utilidad se maximiza sujeta a la restricción presupuestaria $X_n + X_1 p_1 = Y$. Si $X_1 = 0$ la cantidad óptima de X_n es igual a Y, esta característica da como resultado una función de gasto $\tilde{e}_n$ (U) condicionada sobre $X_1 = 0$. Si X_1 es igual a 1 la función de gasto $\tilde{e}_1$ (p_1,U) se puede encontrar. El punto de máxima utilidad esta asociado con el mínimo de estas dos funciones de gasto

(7.46) $e(p_1, \bar{U}) = \min\{\tilde{e}_n(\bar{U}), \tilde{e}_1(p_1, \bar{U})\}$

Dado que $\tilde{e}_n$, $\tilde{e}_1$ son continuamente diferenciables en precios [$\frac{\partial \tilde{e}_n}{\partial p_1} = 0$] excepto a los precios a los cuales $\tilde{e}_1 = \tilde{e}_2$ y e es continuo y diferenciable, las funciones de demandas serán continuas, pero un conjunto de precios particulares y de ingreso llevaran a soluciones de esquina en el cual el bien no es consumido totalmente.

Se puede observar que el rango de precios estará dividido por $p_1^*(\bar{U})$, esto es, para $p_1 < p_1^*(\bar{U})$ una solución interior en el bien 1 existe. De esta forma, las demandas compensadas y la función de gasto se derivan como $\tilde{X}_n^c$, $\tilde{X}_1^c$, y $\tilde{e}$ cuando $p_1 \geq p_1^*(\bar{U})$, es por esta razón, que una solución de esquina se alcanza y el bien 1 no es consumido; esta solución, viene determinada por $\tilde{X}_n^c$ y $\tilde{e}$ por lo cual la restricción presupuestaria requiere solamente que satisfaga la condición

$$(7.47) \quad \tilde{X}_n^c(p_1, \bar{U}) + p_1 \tilde{X}_1^c(p_1, \bar{U}) = \tilde{e}(p_1, \bar{U}) \text{ sí } p_1 < p_1^*(\bar{U})$$

$$\approx{X}_n^c(\bar{U}) = \approx{e}(\bar{U}) \text{ sí } p_1 \geq p_1^*(\bar{U})$$

Donde $p_1^*(\bar{U})$ se define como el precio límite en el cual la demanda se hace cero (precio máximo). De esta forma, la demanda compensada viene definida como

$$X^{c}{}_{n}(p_1,\bar{U}) = \begin{cases} \tilde{X}^{c}{}_{n}(p_1,\bar{U}), si\colon p_1 < p_1{}^{*}(\bar{U}) \\ \approx X^{c}{}_{n}(\bar{U}), si\colon p_1 \geq p_1{}^{*}(\bar{U}) \end{cases} \quad (7.48)$$

Que es continua en p_1, de donde se infiere que el $\underset{p_1 \to p_1^*}{Lim}\tilde{e}$ (p_1, $\bar{U}$)= $\approx e$ ($\bar{U}$).

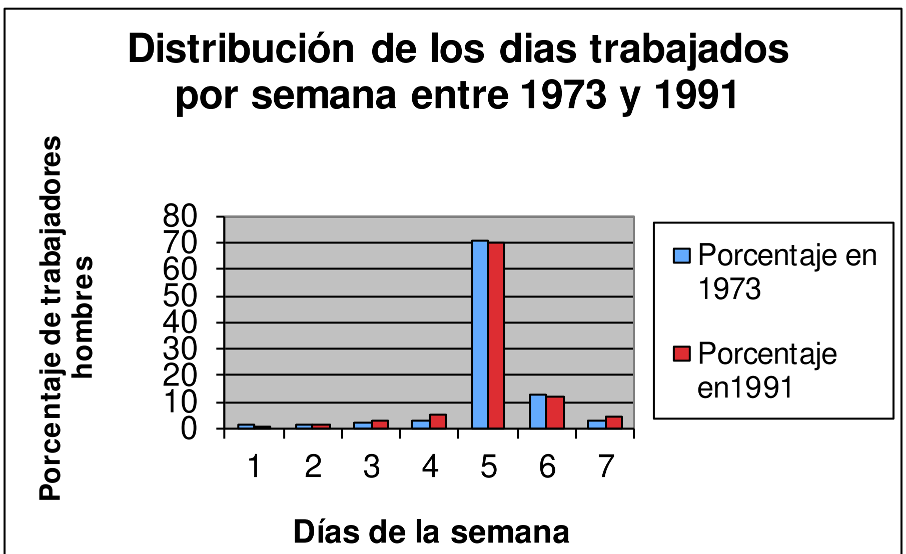
Cuando un bien es comprado en unidades discretas, pero existen no-concavidades en la función de utilidad el consumidor elige entre soluciones alternativas de esquina. Supongamos una canasta de 3 bienes donde las curvas de indiferencia entre X_1 , X_2 y el bien numerario son convexas, entonces a cada vector de precios el consumo tanto en X_1 como en X_2 podría ser cero. Los supuestos del teorema de Small y Rosen garantizan que los $U(X_n,0,X_2)$ y $U(X_n,X_1,0)$ sean funciones bien definidas manteniéndose el excedente del consumidor. Resumiendo, el excedente del consumidor se puede encontrar siempre que exista una función de gasto $\tilde{e}$ ($\bar{U}$) dado que dicha función es diferenciable en precios.

8 Aplicaciones de la teoría del consumidor a la elección de ocio

La teoría y el análisis empírico nos han mostrado, que los individuos producen y consumen teniendo como restricciones el ingreso y el tiempo. Ya hemos visto como el ingreso se puede considerar de forma exógena, sin exigir una mayor formalización. Sin embargo, debemos preguntarnos, por un lado, de dónde sale el ingreso y, por otro lado, cómo influye en el tiempo cuando los consumidores toman este tipo de decisiones.

Existen límites en las restricciones para asignar el tiempo en ocio y trabajo. Como anota Hamermesh (1998) existen restricciones biológicas, culturales e históricas que hacen que los consumidores asignen el tiempo entre ocio y trabajo. Aparte de estas restricciones no económicas, cambios en el precio del tiempo pueden inducir a cambios en las restricciones de ocio y trabajo. Por esta razón, un incremento en el valor del tiempo producirá restricciones en las interacciones entre la familia, los colegas etc. Como observa Hamermesh no solamente ha habido un cambio en las horas trabajadas en este siglo, también ha habido un cambio importante en franjas horarias no tradicionales, como de las 6:00 a 7:00 A.M y de las 5:00 a 6:00 P.M. La tendencia hoy día consiste en trabajar más horas en pocos días, lo que ha llevado a que se pase de semanas de trabajo de seis días a cinco como se puede observar

Gráfica 8.1. Distribución de los días trabajados



Fuente: Hammermesh(1998)

Como se puede observar, de la gráfica 8.1, que para 1991 el porcentaje de individuos que trabajaba 6 días disminuyó mientras el porcentaje que trabajaba 5 días aumento.

Supongamos un individuo o un hogar, compuesto por un trabajador, cuya elección consiste en la compra de diferentes canastas de bienes, a un vector de precios dados. Estas compras, pueden realizarse con un ingreso no laboral μ y un ingreso laboral wL. Donde w es la tasa de salario y L es la cantidad de tiempo que el individuo elige trabajar. Formalmente tendremos

(8.1) $\mu + wL = \sum p_i x_i$

Por el axioma de Insaciabilidad, se garantiza la igualdad. Adicionalmente $x_i \geq 0 \; \forall i$ y T≥ L ≥ 0, por lo cual no se pueden comprar cantidades negativas de los bienes y la oferta de trabajo no puede exceder el tiempo total del individuo.

Una mejor especificación, podría partir de que el tiempo dedicado a trabajar es de 24 horas menos el tiempo necesario para dormir y otras tareas mínimas de mantenimiento, como aseo, comida, etc. Adicionalmente, el tiempo podría ser de un año o más.

Aunque en el largo plazo, los individuos podrían elegir las horas que trabajen a través de la elección de diferentes trabajos, en el corto plazo muchas de las elecciones estarán condicionadas por las horas de trabajo requeridas de acuerdo a los regímenes laborales vigentes y, como en el caso de los mercados laborales en U.S.A y Europa la posibilidad de más de un empleo flexibiliza el número de horas trabajadas, cuando existen trabajos de medio tiempo.

Finalmente, la linealidad en la restricción presupuestaria es un punto de discusión, en tanto no necesariamente la tasa de salario varía en forma directa con el número de horas de trabajo sí las horas no tienen la misma tasa salarial. Por simplificar, se tomará linealmente. Sea la función de utilidad

(8.2) $\mu = V(x_0, x_1, \ldots, x_n)$

Siendo $x_0 = T - L = \text{ocio}$. El ocio es comparable con los otros bienes y separable de los mismos, el precio del ocio será la tasa de salario. Por cuestiones de simplificación, se puede asumir un bien agregado tipo Hicks [20] . Así, la función de utilidad podrá plantearse como

[20] Para Trabajar con un bien agregado tipo Hicks, se requiere que los n´s precios relativos permanezcan constantes.

(8.3) $\mu = V\left(x_0, x_1^*\right)$

Donde x_1* es el bien agregado de Hicks. El problema puede plantearse como

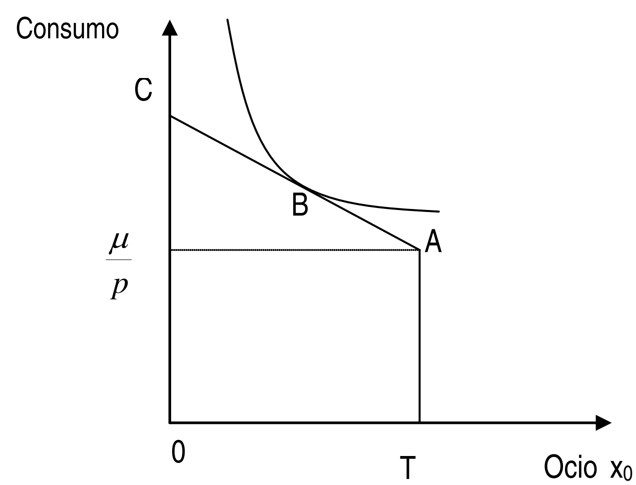
(8.4) Max $\mu = V\left(x_0, x_1^*\right)$

Sujeto a $V_i \geq 0$

$$\mu = V(x_0, x_1, \ldots, x_n)$$

Gráficamente se puede observar

Gráfica 8.2. Decisión de trabajar



Y, la restricción será

(8.5) $px + wx_0 = \mu + wT$

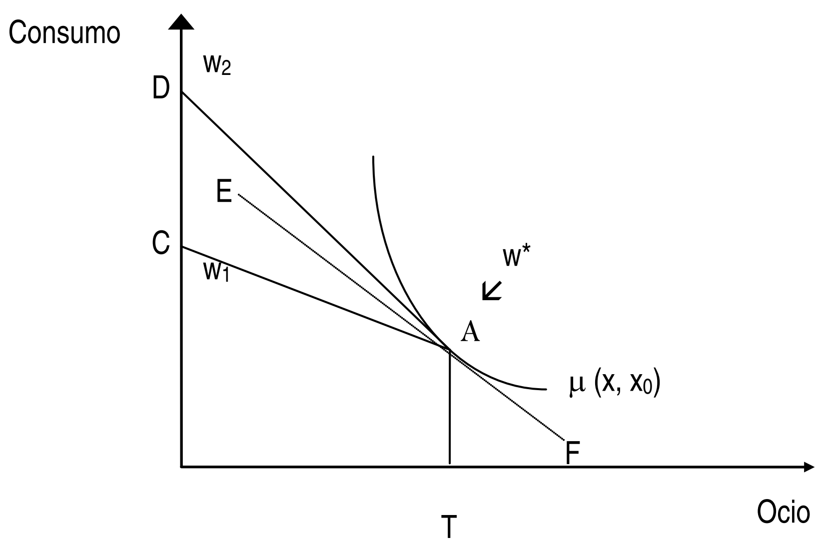
Si x = 0 y no existe ingreso adicional, entonces $wx_0 = wT$ ó $x_0 = T$ si $x_0 = T = 0$, por lo cual w = 0 y $x = \mu / p$. La cantidad de trabajo ofrecida se mide por la distancia 0 - T. Las curvas tendrán diferentes pendientes de acuerdo a la tasa de salario. La decisión de cuantas horas ofrecer a el individuo determinará la decisión de participar en el mercado laboral. Un cambio en w altera el análisis del ingreso y el efecto sustitución, la cantidad μ

+ wT representa el ingreso total disponible del consumidor que será gastado en ocio y bienes.

Ya que los trabajadores son libres de participar en el mercado y ellos varían las horas que deciden trabajar, en especial si trabajan tiempo parcial. Un grupo importante, podrá tomar dichas decisiones, por ejemplo, las mujeres casadas para quienes un segundo trabajo consiste en las actividades que realizan en el hogar.

La gráfica (8.3) muestra las diferentes restricciones presupuestarias de acuerdo a las tasas de salario w_1(CAT) y w_2 (DAT). La curva de indiferencia es tangente a la línea quebrada (E-F), mientras la pendiente de (E-F) es w*/p, esto es, la tasa marginal de sustitución la cual es un indicador de la tasa de salario de reserva w* el punto A en la gráfica (8.3)

Gráfica 8.3. Salario mínimo que induce a participar



Si el salario es menor que w*, el individuo no participará en el mercado, si el salario es w_2 se ofrecerán horas positivas. Si w = w* el trabajador es indiferente. Donde w* es el salario de reserva, aquel valor de w que hace que x_0 = T

$$T = x_0\left(\mu + w^{*}T, w^{*}, p\right) \tag{8.6}$$

Supongamos, la siguiente función de utilidad

$$\mu = \alpha_0 \mathrm{Log}\left(x_0 - \gamma_0\right) + \alpha \mathrm{Log}\left(x - \gamma\right) \tag{8.7}$$

Donde $\alpha_0 + \alpha = 1$ y, γ_0 y γ_1 son ocio y consumo respectivamente. Partiendo de (8.1) y (8.5) se obtiene T – L = x_0 y $\sum_{i=1}^{n} px_i$ = wL + μ, entonces px = w(T- x_0) + μ . De donde

(8.8) $px + wx_0 = \mu + wT$

Haciendo $Y = \mu + wT$, el problema se puede plantear como

(8.9) Max $\alpha_0 \mathrm{Log}(x_0 - \gamma_0) + \alpha \mathrm{Log}(x - \gamma)$

St. : $px + wx_0 = Y$

El Lagrangiano resultante será

$$\ell = \alpha_0 Log(x_0 - \gamma_0) + \alpha Log(x - \gamma) - \lambda(wx_0 + px - Y)$$

$$(1)\frac{\partial \ell}{\partial x_0} = \frac{\alpha_0}{x_0 - \gamma_0} - \lambda w = 0 (2)\frac{\partial \ell}{\partial x} = \frac{\alpha}{x - \gamma} - p\lambda = 0$$

$$(3)\frac{\partial \ell}{\partial \lambda} = wx_0 + px - Y = 0$$

Igualando (1) y (2)

$$\frac{\alpha_0}{(x_0 - \gamma_0)w} = \frac{\alpha}{(x - \gamma)p} \Rightarrow$$

$$\alpha_0 (x - \gamma)p = \alpha w(x_0 - \gamma_0)$$

$$\alpha_0 xp - \alpha_0 \gamma p = \alpha w(x_0 - \gamma_0)$$

$$x = \frac{\alpha w(x - \gamma_0)}{\alpha_0 p} + \frac{\alpha_0 \gamma p}{\alpha_0 p}$$

$$x = \frac{\alpha w(x - \gamma_0)}{\alpha_0 p} + \gamma$$

Sustituyendo en (3) se obtiene

$$p\left(\frac{\alpha w(x-\gamma_0)}{\alpha_0 p}+\gamma\right)+wx_0=Y$$

$$\frac{p\alpha x_0 w}{\alpha_0 p}-\frac{p\alpha w\gamma_0}{\alpha_0 p}+p\gamma+wx_0=Y$$

$$x_0 w\left(\frac{p\alpha}{\alpha_0 p}+1\right)=Y-p\gamma+\frac{p\alpha w\gamma_0}{\alpha_0 p}$$

$$x_0 w\left(\frac{p\alpha+\alpha_0 p}{\alpha_0 p}\right)=Y-p\gamma+\frac{p\alpha w\gamma_0}{\alpha_0 p}; hacemos \alpha+\alpha_0=1$$

$$x_0 w\frac{1}{\alpha_0}=Y-p\gamma+\frac{p\alpha w\gamma_0}{\alpha_0 p}$$

$$x_0=Y\frac{\alpha_0}{w}-p\gamma\frac{\alpha_0}{w}+\frac{p(1-\alpha_0)w\gamma_0}{\alpha_0 p}\frac{\alpha_0}{w}$$

$$x_0=Y\frac{\alpha_0}{w}-p\gamma\frac{\alpha_0}{w}+\frac{p\alpha_0 w\gamma_0}{\alpha_0 pw}-\frac{\alpha_0}{w}\frac{p\alpha_0 w\gamma_0}{\alpha_0 p}$$

$$x_0=\gamma_0+\frac{\alpha_0}{w}(Y-p\gamma-\gamma_0 w)$$

$$x=\gamma+\frac{\alpha}{p}(Y-p\gamma-\gamma_0 w)$$

Nosotros sabemos que $T - L = x_0$ y que $px + wx_0 = \mu + wT$ de donde

(8.10)
$$T-L=\gamma_0+\frac{\alpha_0}{w}\left(\mu+wT-\gamma_0 w-p\gamma\right)$$
$$L=\left(T-\gamma_0\right)-\frac{\alpha_0}{w}\left(\mu+w\left(T-\gamma_0\right)-p\gamma\right)$$

Si el tiempo total es un parámetro a ser estimado esta ecuación nos indicaría que $(T - \gamma_0)$ el tiempo que se excede sobre el ocio ya comprometido, se puede identificar.

Los γ parámetros, nos dan una forma natural de incorporar diferencias en los gustos y las características de los hogares, en función de la oferta de corte transversal. Estos, son usualmente incorporados a través de variables como el hogar, el sexo, la educación, el tamaño del hogar, la localización, etc. De igual forma, se puede observar que

(8.11)
$$L=T-\gamma_0-\frac{\alpha_0}{w}\left(\mu-p\gamma\right)-\frac{\alpha_0 w}{w}\left(T-\gamma_0\right)$$
$$L^*=\left(T-\gamma_0\right)\left(1-\alpha_0\right)-\frac{\alpha_0}{w}\left(\mu-p\gamma\right)$$

Donde L* indica la restricción de no-negatividad sobre las horas trabajadas. El salario de reserva se define cuando L*=0, esto implica que

(8.12) $$\left(T-\gamma_0\right)\left(1-\alpha_0\right)=\frac{\alpha_0}{w}\left(\mu-p\gamma\right)$$

Por lo cual

(8.13) $$w^{*}=\frac{\alpha_0\left(\mu-p\gamma\right)}{\left(T-\gamma_0\right)\left(1-\alpha_0\right)}$$

De esta forma, la oferta laboral L viene dada por

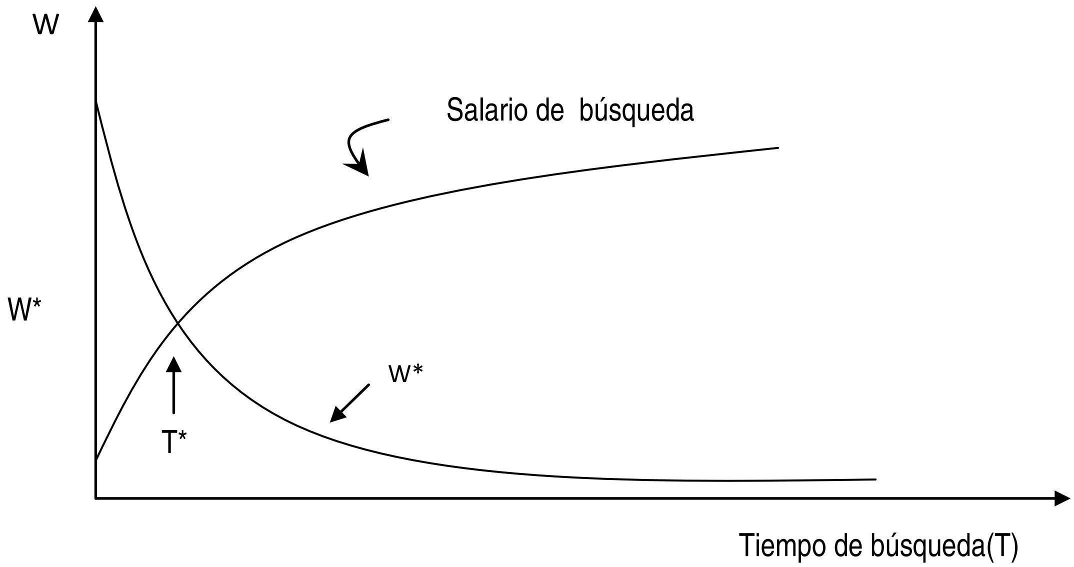
(8.14) $$L=L^{*}=\left(T-\gamma_0\right)\left(1-\alpha_0\right)-\frac{\alpha_0}{w}\left(\mu-p\gamma\right)\quad \text{si } w\geq w^{*}$$

(8.15) $$L=0\qquad\qquad \text{si } w\leq w^{*}$$

Entre los determinantes fundamentales del salario, se considera que la educación juega un papel importante y, por lo tanto, a mayor educación mayor salario de reserva [Kettunen, J(1994), Alba,A(1992) en Uribe (1998)]. Entre otros factores que inciden en el salario de reserva, se encuentran la probabilidad de recibir una oferta, la distribución de los salarios, los factores que determinan el bienestar de estar desempleado [Andres, Garcia y Jimenez (1989)]. El salario de reserva entonces podría sería una función que decrece con el tiempo de búsqueda.

Para Uribe (1998) W* depende de la posición en el hogar, de la edad, del sexo, del tiempo de búsqueda y de la educación. De igual forma Uribe (1998) determina el salario esperado por los trabajadores como una función de la dispersión salarial, del número de vacantes, del tiempo de búsqueda, del ingreso obtenido en el sector informal y de la tasa de desempleo. Al igualar el salario de reserva y el salario esperado se encuentran los determinantes del tiempo de búsqueda

Gráfica 8.4. Determinación del tiempo de búsqueda



Fuente: Uribe(1998)

Uribe encontró para la ciudad de Santiago de Cali, usando la Encuesta Nacional de Hogares de 1992, que la elasticidad educación del tiempo de búsqueda es del 0.7487, la elasticidad experiencia del tiempo de búsqueda es de 0.48. Que los jefes de hogar buscan un 44.6% menos de tiempo en comparación con los que no lo son y tienen características similares y que un incremento en el 100% de las vacantes disminuye el tiempo de búsqueda en el 0.6%.

En trabajos empíricos, la desigualdad determina cuando (8.14) ó (8.15) es observado. Supongamos que (8.13) se cumple. Ahora consideremos una familia en la cual existe una cantidad de horas institucionales que deberán trabajarse y nosotros estamos interesados en la decisión de participar de un segundo trabajador, en el mercado laboral, con lo cual, estaríamos considerando a dos integrantes de la familia que trabajarían.

Las ganancias del primer trabajador, podrían ser un ingreso exógeno μ. Inversamente, el salario de reserva podría disminuir ante la existencia de deudas o intereses. El parámetro γ_0 puede interpretarse como el ocio ya comprometido y, variará entre los hogares de acuerdo a las edades de los hijos y gustos. A mayor γ_0, cuando el grupo familiar tiene niños incrementa w* y reduce la probabilidad de trabajar ya que entre menos edad tenga el niño más tiempo demandará para su cuidado y reducirá la probabilidad de participar.

De esta forma, el conjunto de ecuaciones (8.14) y (8.15), sirve para estimar las horas trabajadas, como una función de la tasa salarial en estudios de corte transversal y series de tiempo. La tasa de participación y el número de horas trabajadas, han sido usados como medidas de la oferta de trabajo, donde ésta es explicada por el salario y la educación [Mincer (1970)] influyendo sobre la participación de la mujer, en el mercado laboral [Owen (1969, 1971) Abbot y Ashelfeltar (1976) y Phlips (1978)].

Esta aproximación, nos muestra porque muchos estudios son erróneos: cambios en la oferta agregada de trabajo como respuesta a cambios en w y μ son ocasionados no solamente a cambios en la oferta de aquellos quienes actualmente están trabajando, y que se puede ver en la ecuación (8.13), sino también, a través de los cambios ocasionados por la unión de los individuos en el hogar y, además del tiempo que queda para laborar en otros trabajos [21]. En últimas, la oferta depende del cambio de w* a través de w.

La evidencia agregada a través de series de tiempo es consistente con mostrar una caída en el largo plazo en ambas horas trabajadas y en un incremento en la participación de la mujer.

Para aquellos hombres que tienen un salario de reserva bajo y quienes trabajan relativamente largas horas el efecto ingreso es dominante. En los estudios observados, esto significa que aumenta el salario, el ingreso y L.

Para las mujeres, sin embargo, el tiempo que se pasa en el hogar tiene un mayor salario de reserva y, por lo tanto, la participación y las horas trabajadas son menos que la de los hombres. Por lo tanto, el analizar la oferta de trabajo deberá realizarse necesariamente en un contexto familiar.

En estudios de corte transversal de los hogares, los trabajadores decidirán cuándo participar o no, siendo la variable de participación dicotómica. En algunos estudios, esta variable dicotómica simplemente se regresa contra los determinantes de participación: número de hijos, número de personas que trabajan en la familia, sexo, edad, nivel de educación (Grenhalg (1980)).

Gronau (1973) ha propuesto el siguiente modelo: Sea γ_0 el ocio para el individuo h que viene dado por

$$(8.16)\quad \gamma_0^h = a_0 + b_0 Z^h + \varepsilon^h$$

Con Z^h una variable de composición y, ε^h un término aleatorio de error. Si w^h es el salario ofrecido por participar, h podría participar si $w^h > w^{*h}$, esto es, sí

$$(8.17)\quad w^h > \alpha_0\left(\mu^h - p\gamma\right)\frac{1}{\left(1-\alpha_0\right)\left(T - Y_0^h\right)} \Rightarrow w^h > \frac{\alpha_0\left(\mu - p\gamma\right)}{\left(1-\alpha_0\right)\left(T - \alpha_0 + b_0 Z^h + \varepsilon^h\right)}$$

Por conveniencia α_0, T, Y y p no varían entre los hogares. Reescribiendo en términos de ε^h tendremos

[21] Como menciona Pencavel(1998) existe una amplia literatura sobre las elecciones de trabajo de esposas y esposos, pero muy poca literatura sobre las elecciones "maritales" que explican conjuntamente a esposos y esposas. El trabajo de Pencavel apunta a las elecciones por educación en los matrimonios encontrando que mientras en 1940 el 12% de esposas y el 15% de maridos tenían más de 12 años de escolaridad en 1991 estos porcentajes eran del 55% para esposas y esposos.

$$w^h > \alpha_0\left(\mu^h - p\gamma\right)\frac{1}{(1-\alpha_0)\left(T-\alpha_0+b_0Z^h+\varepsilon^h\right)} \Rightarrow$$

$$(1-\alpha_0)\left(T-\alpha_0+b_0Z^h+\varepsilon^h\right)-(1-\alpha_0)\varepsilon^h > \frac{\alpha_0}{w^h}\left(\mu^h - p\gamma\right)$$

(8.18)

$$(1-\alpha_0)\left(T-\alpha_0+b_0Z^h+\varepsilon^h\right)-\frac{\alpha_0}{w^h}\left(\mu^h - p\gamma\right) > (1-\alpha_0)\varepsilon^h$$

$$ó \quad (1-\alpha_0)\varepsilon^h < (1-\alpha_0)\left(T-\alpha_0+b_0Z^h+\varepsilon^h\right)-\frac{\alpha_0}{w^h}\left(\mu^h - p\gamma\right)$$

Que es la decisión de participar. Nosotros sabemos, que si ε sigue alguna función de densidad, por ejemplo la normal, la desigualdad anterior nos daría la probabilidad de participar en términos de la función de distribución, de las variables Z^h, w^h , γ y los μ^h parámetros.

Si y^*_1 representa el salario ofrecido menos el salario de reserva, el más bajo salario que se está dispuesto a aceptar, y y^*_2 representa el salario ofrecido solamente cuando el salario ofrecido exceda el salario de reserva. Entonces nosotros observaremos el salario actual, el cual es igual al salario ofrecido, cuando

(8.19) $y_{2i} = y^*_{2i}$ Sí $y^*_{1i} > 0$

$y_{2i} = 0$ Sí $y^*_{1i} \leq 0$ $i = 1,2,3...$

Por otro lado, también sabemos que

(8.20) $y^*_{1i} = x'_{1i}\beta_1 + \mu_{1i}$

$y^*_{2i} = x'_{2i}\beta_2 + \mu_{2i}$

Donde μ_{1i} y μ_{2i} son distribuciones normales bivariadas con media cero, varianzas σ^2_1 y σ^2_2 y covarianza σ_{12}. La función de verosimilitud vendrá dada para este modelo como

(8.21) $$L = \prod_0 P\left(y^*_{1i} \leq 0\right)\prod_1 F\left(y_{21}\middle|y^*_{1i} > 0\right)P\left(y^*_{1i} > 0\right)$$

En últimas, la función de verosimilitud se puede escribir también como

(8.22) $$L = \prod_1 \Phi\left(x_i^{'}\beta/\sigma\right)^{-1}\sigma^{-1}\phi\left[\left(y_i - x_i^{'}\beta\right)/\sigma\right]$$

Con Φ la distribución normal estándar y ϕ la función de densidad. Siendo un estimador máximo verosímil, que es consistente incluso cuando los errores están serialmente correlacionados. De esta forma para un Tobit normal truncado tendremos

(8.23) $F\left(y_{1i}\right)=\begin{cases}\dfrac{1/\sigma\,\phi\left(\left(y_{1i}-\beta_j x_{ij}\right)/\sigma\right)}{1-\Phi\left(-\beta_j x_{ji}/\sigma\right)} & si\ y_{1i}>0\\ 0 & De\ otra\ forma\end{cases}$

Esta ecuación es directamente estimable. Ver por ejemplo Mora (2013) quien modela el efecto de las remesas sobre la participación laboral como ingreso no laboral. Las remesas se consideran como una variable endógena y truncada.

8.1 El efecto de las herencias sobre la oferta laboral

El efecto de la herencia sobre la oferta laboral se conoce también como la hipótesis de Carnegie [Holtz-Eakin, Joulfaian, Rosen (1993)]. Según esta hipótesis una mayor herencia disminuye la oferta laboral individual. En particular, los autores muestran cómo una persona que recibe una herencia de U$ 150.000 está cuatro veces más dispuesta a no trabajar que una persona que recibe una herencia de U$ 25.000. La hipótesis es importante en el sentido de que la corroboración de la misma es consistente con la hipótesis de que el ocio puede ser tratado como un bien normal. Sin embargo, su importancia no solo radica en términos de la verificación de la normalidad del ocio, también puede observarse cómo influye en el ciclo de vida a través de la decisión de trabajar. Holtz-Eakin, Joulfaian y Rosen (1993) encuentran los siguientes resultados

Tabla 8.1. Modelo Logit sobre herencias (Errores estándar en paréntesis)

Modelo	1	2	3
Constante	- 0,1485	0,0002771	1,272
	(0,1507)	(0,8588)	(0,8819)
Herencia (Millones de U$)	- 4,775	- 3,931	- 3,681
	(0,7769)	(0,7877)	(0,8017)
[Herencia (Millones de U$)]^2	2,427	1,944	2,069
	(0,7976)	(0,7781)	(0,7450)
LF´82 = 1si trabajo en 1982			
0 de otra forma	2,915	2,918	2,203
	(0,1679)	(0,1757)	(0,2056)
Edad en 1982		0,02375	- 0,03802
		(0,04990)	(0,05166)
[Edad en 1982]^2		- 0,0007358	- 0,00006577
		(0,0006405)	(0,0006626)
Ingresos(Millones de U$) en 1982.			0,05749

			(0,01091)
Dividendos + intereses en 1982.			- 0,01075
			(0,004479)
Dependientes en 1982			0,3361
			(0,1664)
LogLikelihood	- 552,0	- 537,9	- 513,8
N	1632	1632	1632

Variable dependiente: 1 si trabajo en 1985
0 de otra forma

Como se puede observar del modelo (1) un incremento en la herencia recibida reduce la probabilidad de trabajar en 1985, siendo este efecto significativamente importante. Del modelo (2) y (3) se puede deducir que la probabilidad de trabajar depende de la edad negativamente, lo cual significa que a mayor edad mayor probabilidad de salir de la fuerza de trabajo. En el modelo (3) se introducen los ingresos, los dividendos e intereses y el número de personas dependientes del individuo en 1982. Los ingresos se interpretan como una medida del costo de oportunidad de la fuerza de trabajo y debe esperarse que individuos con mayores ingresos estén dispuestos a permanecer trabajando. También se puede observar que la probabilidad de permanecer trabajando aumenta con el número de dependientes y decrece con la edad y los ingresos no dependientes del ingreso (Herencias). Finalmente Holtz-Eakin, Joulfaian y Rosen estiman el mismo modelo cuando los individuos reciben 2 o 3 ingresos adicionales y los resultados se mantienen como se puede observar

Tabla 8.2. Modelo Logit Multinomial sobre herencias (Errores estándar entre paréntesis)

Modelo	1	2	3
Constante	- 0,7466	- 4,420	- 3,979
	(0,2829)	(0,8325)	(0,8854)
Herencia(Millones de U$)	- 2,6832	- 2.353	- 2,365
	(0,4324)	(0,4326)	(0,4352)
[Herencia(Millones de U$)]^2	1,259	1,0794	1,087
	(0,3866)	(0,3724)	(0,3726)
LF´82 = 1 si la familia tiene un ingreso por trabajos en 1982 0 de otra forma	3,995	3,853	3,764
	(0,3093)	(0,3303)	(0,3337)
LF 2´82 = 1 si la familia tiene dos ingreso por trabajos en 1982 0 de otra forma	6,196	6,071	5,985
	(0,3135)	(0,3346)	(0,3398)

Edad en 1982			0,2315 (0,03963)	0,2078 (0,04372)
[Edad en 1982]^2			- 0,003139 (0,0004779)	-0,002854 (0,0005296)
Ingresos(Millones de U$) en 1982.				0,001964 (0,001613)
Dividendos + intereses en 1982.				- 0,0006123 (0,001051)
Dependientes en 1982				0,03711 (0,04144)
LogLikelihood	- 1768,0	-1727,1		- 1725,4
N	2700	2700		2700

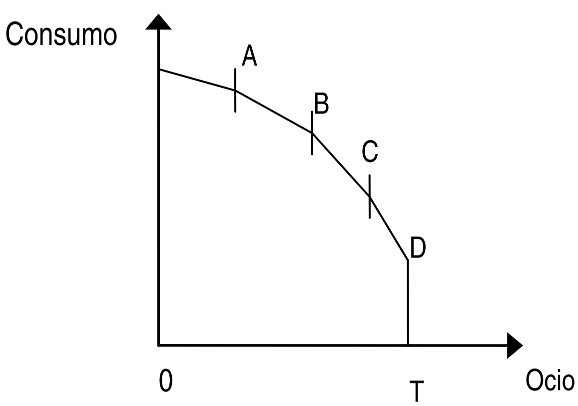
Variable dependiente: 0 si la familia tiene 0 ingresos en 1985
1 si la familia tiene 1 ingreso en 1985
2 si la familia tiene 2 ingresos en 1985

Los resultados confirman las conclusiones de que una mayor herencia recibida por la familia reduce la probabilidad de que ambos cónyuges participen en el mercado de trabajo e incrementa la probabilidad de que ninguno de ellos participe en el mercado de trabajo.

8.2 Restricciones no lineales y restricciones sobre las horas

Cuando en un país existe un sistema de impuestos complejos y seguridad social, esto origina restricciones no lineales, donde las tasas marginales aumentarán con el ingreso

Gráfica 8.5. No linealidades generadas por tasas impositivas diferentes



Dado que los sistemas impositivos no son lineales, los incrementos marginales aumentan con el ingreso produciendo una restricción no lineal en la participación de la fuerza de trabajo como se observa en la gráfica (8.5). La pendiente dependerá de la relación entre los ingresos, la existencia de patrimonios superiores a un monto determinado por el gobierno, loterías, la existencia de herencias y del número de dependientes. Por ejemplo, para el año 2000 el gobierno Colombiano estableció la siguiente tabla de retención en la fuente, declaración de renta y complementarios

Tabla 8.3. Impuesto de retención en la fuente: Año gravable 2000

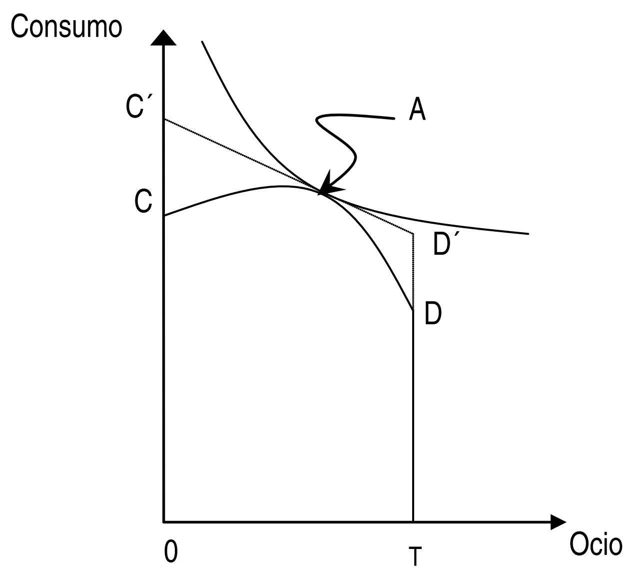
Intervalos	% retención	Valor a retener	Intervalos	% retención	Valor a retener
1 a 1.100.000	0.00%	0	2.000.001 a 2.050.000	9.80%	198.500
1.100.001 a 1.110.000	0.09%	1.000	2.250.001 a 2.300.000	11.91%	271.000
1.110.001 a 1.120.000	0.27%	3.000	2.500.001 a 2.550.000	13.60%	343.500
1.120.001 a 1.130.000	0.44%	5.000	2.750.001 a 2.800.000	14.99%	416.000
1.130.001 a 1.140.000	0.62%	7.000	3.000.001 a 3.050.000	16.15%	488.500
1.140.001 a 1.150.000	0.79%	9.000	3.250.001 a 3.300.000	15.13%	561.000
1.150.001 a 1.200.000	1.28%	15.000	3.500.001 a 3.550.000	17.97%	633.500
1.200.001 a 1.250.000	2.04%	25.000	4.000.001 a 4.050.000	19.34%	778.500
1.500.001 a 1.550.000	5.57%	85.000	4.500.001 a 4.550.000	20.51%	915.000
1.550.001 a 1.600.000	6.03%	95.000	5.000.001 a 5.050.000	22.04%	1.107.500
1.600.001 a 1.650.000	6.46%	105.000	5.250.001 a 5.300.000	22.65%	1.195.000
1.650.001 a 1.700.000	6.87%	115.000	5.500.001 en adelante	35%	-

Tabla 8.4. Tarifas del impuesto sobre la renta y complementarios: Año gravable 2000.

Intervalos de renta Gravable o ganancia ocasional	Tarifa del promedio del intervalo	Impuesto	Intervalos de renta gravable o ganancia ocasional	Tarifa del promedio del intervalo	Impuesto
1 a 12.300.000	0.00%	0	35.500.001 a 35.700.000	17.44%	6.208.000
12.300.001 a 12.500.000	0.16%	20.000	40.100.001 a 40.300.000	18.76%	7.542.000
20.100.001 a 20.300.000	8.62%	1.742.000	50.100.001 a 50.300.000	20.92%	10.502.000
30.100.001 a 30.300.000	15.37%	4.642.000	53.700.001 a 53.900.00	21.86%	11.762.000

Ya que existe una gran cantidad de variables que podrían causar no convexidades en la participación de la fuerza de trabajo, Hall (1973) ha sugerido que la restricción no lineal para un individuo, puede ser reemplazada por la restricción lineal C´D´ tangente al punto [A], gráfica(8.6). En el punto [A] el comportamiento correspondiente a la restricción TDC es idéntico a TD’C’. Sin embargo, al estimar linealmente C’D’, esta depende de la oferta de trabajo observada y la causalidad podrá presentarse no solamente del salario al número de horas sino a la inversa, lo cual significa que los individuos con bajas preferencias por ocio podrán tener un salario bajo y el efecto estimado de los salarios sobre las horas podría estar sesgado hacia abajo. En principio, una estimación bajo un sistema de ecuaciones simultáneas que reconozca que la oferta y la tasa de salario son endógenamente dependientes, solucionará el problema

Gráfica 8.6. Causalidad bidireccional del salario a las horas trabajadas



8.3 Restricciones sobre las horas trabajadas

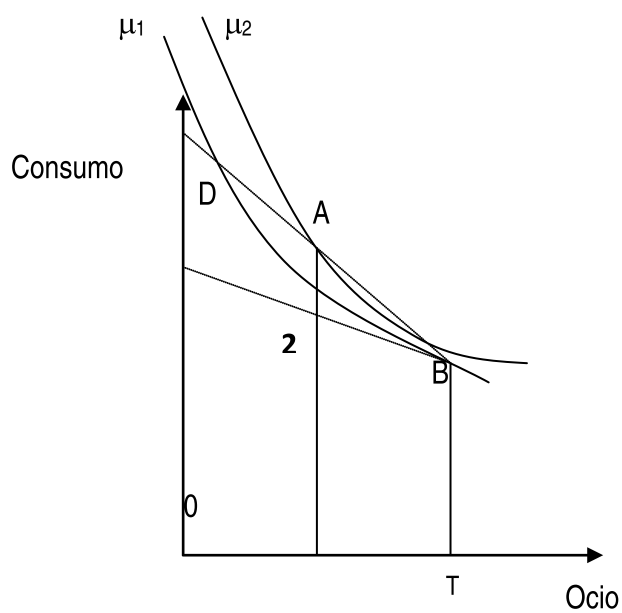
La duración del día de trabajo ha cambiado sustancialmente como lo observa Costa(1998) de 10 horas en 1880 pasó a 8 horas en 1940 y menos de 8 horas en 1991 en U.S.A. Estos cambios son explicados por cambios tecnológicos como la electrificación, lo cual se traduce en cambios en la demanda por días de trabajo de las firmas; y en gran medida estos cambios se deben también a cambios en la legislación. Costa (1998) también muestra que para la población masculina entre 25 y 64 años, que mientras la elasticidad del salario con respecto a las horas trabajadas en 1890 era de – 0.536 en 1991 es de 0.104. Cabe anotar también, que las desigualdades en ingresos entre los deciles 9 y 10, entre 1973 y 1991, son atribuibles a diferencias en las horas trabajadas.

Debido a factores tecnológicos y legislativos, muchos trabajadores, no tienen una completa flexibilidad para elegir las horas que desean trabajar. Esto afecta, al menos en el corto plazo, la elección efectiva entre trabajar un día, una semana o no trabajar. Para examinar cómo influye esta restricción, supongamos que el número de horas por semana es fijo a un nivel $\overline{L}$ (independiente si trabaja un día, una semana o un año). Si el individuo decide trabajar, la restricción presupuestaria será

(8.24) $px = w\overline{L} + \mu$,

De esta forma, la maximización de $U(x, T - \overline{L})$ con $T - \overline{L}$ = x0(ocio) sujeto a la restricción y usando la función indirecta de utilidad condicionada sobre el número de horas, tendremos que $\nu(w\overline{L} + \mu, p, T - \overline{L})$. Como la decisión de trabajar es una elección binaria, una comparación entre la utilidad de trabajar y la utilidad de no trabajar $W(\mu, p, T)$ determinará cuándo es preferible trabajar o no:

Gráfica 8.7. Elección de trabajar



El número de trabajadores que estarían dispuestos a trabajar, dependerá en el agregado, a través de la comparación binaria de la distribución conjunta de las tasas de salario, del ingreso no laboral, de las características observables y, aquellas no observables que entraran en la función de utilidad.

Como puede observarse, la curva de indiferencia $\mu 2$ no es tangente sobre [AB] en [A], en este caso el individuo trabaja más horas en [A], de lo que podría elegir si todo [AB] fuese disponible. Sin embargo, [B] no es la mejor elección para el individuo pues implica una menor curva de indiferencia, $\mu 1$. La situación está dada para que el empleador tenga una considerable flexibilidad en poder variar el número de horas requerido. En el caso anterior, el trabajador solamente puede elegir no trabajar al salario dado, si las horas trabajadas están por encima de [B] aunque él podría aceptar algún número de horas entre [A] y [B]. Por otro lado, estar desempleado podría ser resultado de información imperfecta acerca del rango de oportunidades disponibles [Spence(1973) y Stigliz(1975)].

Comprobar, cuando los individuos tienen algún tipo de restricción o no, puede ser un procedimiento simple, en tanto, se pueda obtener una estimación correcta del grado en el cual ellos estén subempleados o sobreempleados, entonces las tasas pueden ser corregidas dada la oferta actual de trabajo.

Si esto no es así, es decir, si la información no es disponible, una solución consiste en partir del grado de subempleo teórico y encontrar las funciones de probabilidades subyacentes, veamos. Suponga que L_h horas son reportadas por un individuo, cuya oferta de trabajo es L^s, entonces, hay tres posibilidades: Primero, él puede reportar que esta subempleado, en cuyo caso $L^s{}_h \geq L_{h,}$. Segundo, él puede reportar que esta sobreempleado, en cuyo caso $L^s{}_h \leq L_h$ y Tercero, él puede reportar que no está racionado, en cuyo caso $L^s{}_h = L_h$. Si especificamos la función de oferta de trabajo como $L^s{}_h = L^s\,(x_h) + \varepsilon_h$ y ε_h tiene una función de distribución F(•) y una función de densidad f(•). Estos tres eventos pueden ser descritos como

Evento	**Condiciones**	**Probabilidad ó densidad**
Subempleado	$L^s{}_h \geq L_h = \varepsilon_h \geq L_h - L^s(x_h)$	$1 - F[L_h - L_s(x_h)]$
Sobreempleado	$L^s{}_h \leq L_h = \varepsilon_h \leq L_h - L^s(x_h)$	$F[L_h - L_s(x_h)]$
Sin racionamiento	$L^s{}_h = L_h = \varepsilon_h = L_h - L^s(x_h)$	$f[L_h - L_s(x_h)]$

Suponga que las observaciones de 1 hasta R_1 se refieren a los subempleados, de $R_1 + 1$ a R_2 se refieren a los sobreempleados y de $R_2 + 1$ hasta H se refiere a los no racionados, entonces la densidad de probabilidad conjunta de la muestra, asumiendo que ε_h es independiente de x_h a través de los hogares será

$$(8.25)\quad L = \prod_{h=L}^{R_1}\left[1 - F\left(L_h - L^s\left(x_h\right)\right)\right] - \prod_{h=R+1}^{R_2}\left[F\left(L_h - L^s\left(x_h\right)\right)\right]\prod_{h=R_2+1}^{H}\left[F\left(L_h - L^s\left(x_h\right)\right)\right]$$

En un estudio, realizado por Ham (1977) usando datos de la universidad de Michigan sobre ingresos, encontró a partir de la muestra, que de las mujeres con edad entre 25 y 50 años en 1967 solamente el 28% no experimentó ninguna forma de racionamiento en 1967 - 74. El porcentaje anual de desempleados varió del 3.5% al 7%. Ham censuró la muestra (ver capítulo 6) y uso máxima verosimilitud para estimar los datos.

8.4 Asignación del tiempo para dormir

Como hemos visto hasta ahora, los consumidores asignan su tiempo entre las más diversas actividades incluyendo el tiempo para dormir. Parece existir un consenso de que las necesidades de dormir vienen determinadas biológicamente, razón por la cual, su análisis no debe ir más allá. En la mayoría de estudios de oferta de trabajo individual se asume implícitamente que es fija la cantidad de tiempo asignado entre trabajo-ocio-dormir [Michael(1973),Heckman y MaCurdy(1980), Deaton y Muellbauer(1980)]. Sin embargo, la cantidad de trabajo que un individuo ofrece podría ser variable si el tiempo usado para dormir cambia de semana a semana y de año en año. Esto podría ser así, si la variación en el tiempo usado para dormir cambia como respuesta a cambios en los incentivos económicos. Webb(1985) encuentra que la presencia de niños en el hogar reduce la duración del sueño y que las personas duermen menos en los días de trabajo que en los fines de semana. De igual forma Biddle y Hamermesh(1990) encuentran que una hora adicional de trabajo reduce el tiempo para dormir en aproximadamente 10 minutos.

8.4.1 Demanda de tiempo para dormir

Cuando las personas no derivan utilidad de dormir y éste no tiene impacto sobre la productividad del trabajo, entonces la elección de un consumidor es simple: La duración del sueño es igual al mínimo biológico necesario T^*_{NB} el cual variará dependiendo de las elecciones que las personas realicen en torno a la asignación de su tiempo.

Para algunos individuos se puede asumir que el sueño es una bien intensivo en tiempo y que su consumo produce utilidad como lo hacen otras bienes. Particularmente, será un bien que toma solamente tiempo y no bienes aunque reduce la cantidad de tiempo disponible para producir ingresos salariales[22]. Suponga que el salario de mercado W_m sea

(8.26) $W_m = W_1 + W_2 T_s$

Donde W_1 y W_2 son positivos y T_s es el tiempo asignado para dormir. Asuma que la función de utilidad esta definida sobre T_s y un bien Z de la forma

(8.27) $u = U(Z, T_s)$

Asuma también que el tiempo de trabajo T* se divide en $T_Z + T_s + T_{w.}$
Por otro lado, $T_Z = bZ$ es el tiempo usado en producir Z . La producción de Z requiere una serie de insumos X tales que $X = \alpha Z$ [23]. El precio de X es p y la restricción individual vendrá dada por

[22] No solamente es indispensable por estas características, como observa Becker(1965), el tiempo para dormir es necesario para la productividad individual y la eficiencia.

[23] Deberá adoptarse coeficientes fijos en la tecnología de la función de producción de hogares(Capitulo 5).

(8.28) $pX = W_m T_W + \mu$

Donde μ es un ingreso no laboral como herencias, rifas o ingresos ocasionales. De (8.26) y (8.28) se deduce que la restricción para un individuo viene dada por

(8.29) $(W_1 + W_2 T_s)(T^* - T_Z - T_s) + \mu = p\alpha Z$

La persona entonces maximizará (8.27) sujeto a (8.29). De esta forma, el problema se plantea como

(8.30) Max $u = U_1(Z) + U_2(T_s)$
Sujeto a $(W_1 + W_2 T_s)(T^* - T_Z - T_s) + \mu = p\alpha Z$

El Lagrangiano será

$$\ell = U_1(Z) + U_2(T_s) - \lambda\left[(W_1 + W_2 T_s)(T^* - bZ - T_s) + \mu - p\alpha Z\right]$$

(8.31)
$$\frac{\partial \ell}{\partial Z} = U_1 - \lambda\left[-b(W_1 + W_2 T_s) - p\alpha\right]$$
$$\frac{\partial \ell}{\partial T_s} = U_2(T_s) - \lambda\left[+W_2(T^* - bZ - T_s) - (W_1 + W_2 T_s)\right]$$
$$\frac{\partial \ell}{\partial \lambda} = (W_1 + W_2 T_s)(T^* - bZ - T_s) + \mu - p\alpha Z$$

De donde se deduce

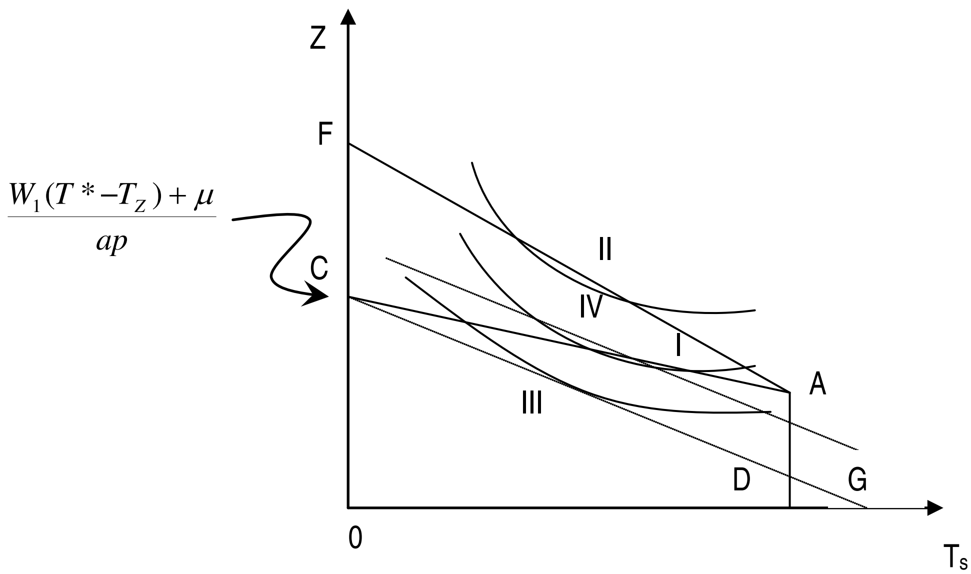
(8.32) $$\frac{U_1}{U_2} = \frac{bW_m + p\alpha}{W_1 + W_2(T_s - T_W)}$$

La ecuación (8.32) muestra que la razón de las utilidades marginales entre consumo y sueño deberán ser iguales a la razón entre precios. El precio de una unidad de Z refleja el costo de los bienes requeridos para producir éste y, el precio sombra del tiempo necesario para su producción. El precio de una unidad de sueño será la tasa de salario menos alguna adición al ingreso laboral proveniente del efecto extra del sueño sobre la productividad.

8.4.2 Efecto sustitución y efecto ingreso en la demanda de tiempo para dormir

A continuación analizaremos qué sucede con el tiempo para dormir ante un cambio en los incentivos económicos. Suponga que I sea la situación inicial sobre la línea AC. Un incremento en W produce una rotación hacia afuera partiendo desde A hacia AF, gráfica (8.8). El efecto es contrario a lo que sucede en el caso tradicional donde un incremento en el precio del bien sobre el eje horizontal produce que la línea de presupuesto gire desde CA hacia CG. La diferencia entre CG y FA consiste en el efecto ingreso extra cuando se reasigna el tiempo total. De esta forma, el efecto sustitución ante un cambio de salarios será el movimiento de I a IV mientras el efecto ingreso será de IV a II y no de IV a III como se observa en la gráfica

Gráfica 8.8. Efecto sustitución y efecto ingreso



La gráfica (8.8) muestra cómo el efecto ingreso es positivo ante un incremento en W para la demanda de Z y T_s, dado que son bienes normales. Este resultado gráfico se puede corroborar de la solución al problema(8.30), veamos

$$(8.33)\quad T_s = \frac{\mu + W_1T^* - Z\left(W_1b + \frac{U_1W_1}{U_2} - \frac{U_1W_2}{U_2} + bW_m\right)}{W_1 + W_2\left(\frac{U_1T_Z}{bU_2} - T_w\right)}$$

De (8.33) se puede observar que la demanda por sueño se ve afectada por cambios en aquellos factores exógenos que cambian la demanda por bienes o a través del bien compuesto Z. La demanda por sueño dependerá también de la cantidad de ingresos laborales y del tiempo total, e inversamente del salario derivado por trabajar y de la producción de Z. Los efectos sustitución e ingreso se pueden observar a través de un cambio en W_m sobre T_s en la forma usual

$$(8.34)\quad \left.\frac{\partial T_s}{\partial W_m}\right|_{\mu} = \frac{\partial T_s}{\partial W_m} - \frac{\partial T_s}{\partial (T_W W_m + \mu)}T_s + \frac{\partial T_s}{\partial (T_W W_m + \mu)}T^*$$

En (8.34) el efecto ingreso será positivo mientras el efecto sustitución será negativo. También se puede observar que cuando cambian los ingresos no laborales, siendo productivo el sueño, una caída en T_W aumenta el precio del sueño

$$(8.35)\quad \frac{\partial T_s}{\partial \mu} = \frac{1}{W_1 + W_2\left(\frac{U_1 T_Z}{bU_2} - T_w\right)}$$

Biddle y Hamermesh estiman una ecuación de demanda usando una muestra de consumidores sobre usos de tiempo entre 1975-1976. La ecuación estimada por ellos, es

$$(8.36)\quad T_j = \gamma_{1j} + \gamma_{2j} W_m + \gamma_{3j} I + \beta_j X + \varepsilon_j$$

Donde T es el logaritmo del tiempo cuando se trata de la demanda por sueño j = s y caso de la demanda por el bien Z , j = Z. W_m es el logaritmo de la tasa de salario, I es el logaritmo de otros ingresos, X es un vector de variables demográficas y ε_j es el termino aleatorio de error. Los resultados encontrados fueron

Tabla 8.5. Demanda de tiempo para dormir.

Variable dependiente	Salarios	Ingreso	$\bar{R}^2$
		Todos los individuos	
8.4.2.1.1 Sueño y Sueño ligero	- 141.44	- 1.78	
	(77.35)	(4.80)	0.24
Tiempo para producir Z	132.18	- 1.71	
	(129.37)	(8.09)	0.162
		Hombres	
Sueño y Sueño ligero	- 181.68	- 2.88	
	(120.88)	(5.77)	0.40
Tiempo para producir Z	233.34	- 6.69	
	(193.67)	(9.30)	0.50
		Mujeres	
Sueño y Sueño ligero	- 64.30	1.55	
	(93.44)	(8.43)	0.18
Tiempo para producir Z	- 262.42	14.44	
	(166.99)	(4.80)	0.053

En la tabla (8.5) se muestran los resultados para γ_{2j} y γ_{3j}. Los resultados son presentados para la muestra entera y en forma separada para hombres y mujeres. Tal vez, la conclusión más importante que se pueda obtener de un estudio como éste, consiste en que al menos una parte del tiempo para dormir es una reserva de la cual las personas extraen tiempo cuando las circunstancias hacen atractivos otros usos. Esto significa, que la elección del consumidor con relación a la elección por tiempo para dormir, se ve afectada por variables económicas al igual que otras elecciones que usan tiempo.

9 Aplicaciones de la teoría del consumidor al medio ambiente

En los últimos años, el desarrollo de la legislación medio ambiental en países como Colombia [24], ha despertado un creciente interés en estimar los cambios en el bienestar de los individuos ante cambios en las provisiones de bienes naturales como el medio ambiente, esto es, el efecto de una modificación por ejemplo en la calidad del aire, o en la calidad de zonas naturales como parques, lagos, paisajes, etc.

En este capítulo, se presentarán los siguientes métodos de valoración ambiental: el método de coste de viaje, el método de los precios hedónicos y el método de la valoración contingente. El objetivo fundamental de este capítulo es ampliar el horizonte de pensamiento de aquel lector que ha seguido los capítulos anteriores. Es en este sentido, una revisión de los capítulos 1-3 de este libro y luego una lectura al libro de Azqueta (1994), Pearce y Turner(1990) y al libro de Freeman III(1993) podrán brindar una idea más general de los métodos de valoración ambiental.

9.1 El método de coste de viaje

Para ilustrar el método de coste de viaje, usaré dos aproximaciones: La primera parte consiste en el modelo tradicional de coste de viaje adicionando el uso de variables latentes [Mora(1997)]. La segunda parte consiste en el modelo de utilidad aleatorio para el número de visitas.

9.1.1 El uso de variables latentes

Considere una serie de consumidores que deciden visitar un paisaje específico, el cual es considerado como un bien. Los agentes económicos toman la decisión de visitar dicho paisaje, de acuerdo a los "precios" del Paisaje, y aunque no existe un precio explícito por el bien Paisaje, esto no quiere decir que este precio no exista, ya que el consumidor, realiza una serie de gastos cuando visita un lugar determinado y, a través, de estos gastos se puede estimar una función de demanda por Paisaje. Los gastos dependen del coste del viaje en cualquier tipo de transporte (Pt), del gasto derivado de estar en un lugar determinado(incluyendo alimentación, etc.) (PA), y del coste de oportunidad del salario (PW).
Dado que cada visita tiene un gasto, el consumidor buscará minimizar el gasto de cada visita manteniendo su utilidad. Así el problema se plantea como

(9.1) $MinC(u,p): Y = (P_t + P_A + P_W)Z; St: v(z) = u$

[24] Especialmente con la expedición de la ley 99 de 1993 se reglamenta el cobro por uso de los recursos ambientales. A partir de la expedición de esta ley, diferentes modificaciones se han realizado, una de las ultimas, el decreto 901 de abril de 1997, reglamenta las tasas retributivas por uso directo o indirecto del agua como receptor de vertimientos puntuales.

Donde c(u,p) es la función de gasto e Y el ingreso. De esta forma, un consumidor planea una serie de actividades derivadas de contemplar un paisaje, pasear por un lugar, etc. y elige un bien Z, la cantidad de viajes a ese lugar. El problema planteado de la anterior forma, es el simple modelo de coste de viaje utilizado por Bockstael et-al(1987) Smith y Kouru(1990). Siguiendo a Kealy y Bishop(1986) y asumiendo una función de utilidad cuadrática(cuasilineal) de la forma

(9.2) $$\nu = A_0 Z + \frac{A_1}{2} Z^2 + A_2 Z$$

Donde A_0, A_1 y A_2 son los parámetros de la función de utilidad. A_0 depende de las características individuales S(Sexo, Edad, Ingreso), Pollak y Walles (1969), Pollak(1969,1970) [25] y A_2 depende de las características paisajísticas P_j. Siendo A_0 y A_2 de forma lineal, el problema se puede también plantear como

(9.3) $$Max(Z): Y - (P_t + P_a + P_W)Z - (\alpha_0 + \sum_{k=1}^{K} \alpha_k S_k)Z - \frac{A_1}{2}Z^2 - (\alpha_2 + \sum_{j=1}^{J} \alpha_j P_j)Z$$

Solucionando la anterior ecuación, para el problema de maximización en Z, bajo una solución interior obtenemos

(9.4) $$Z = -\frac{1}{A_1}(\alpha_0 + \sum_{k=1}^{K} \alpha_k S_k) - \frac{1}{A_1} P - \frac{1}{A_1}(\alpha_2 + \sum_{j=1}^{J} \alpha_j P_j) + \varepsilon$$

Siendo $P = P_t + P_A + P_W$ y $\varepsilon \approx (0, \sigma^2{}_\varepsilon)$. La función de demanda anterior requiere una solución interior para el mercado de trabajo, dependiendo la misma del tiempo cuando este es exógeno o endógeno McConell(1992), Bockstael, Hanemann y Strand(1987). Si se asume como en Bockstael et-al que la tasa de salario refleja el valor del tiempo individual, dado que trabajo y ocio se intercambia marginalmente, obtenemos un valor real para el parámetro asociado al coste de oportunidad. En caso contrario el valor marginal del tiempo individual en otros usos, no es igual a la tasa de salario y el coste de oportunidad no es igual al valor del parámetro obtenido.

Balkan y Kahn(1988), Willis y Garrod(1991), Kealy y Bishop(1986), Bokstael, Strand y Hanemman(1987) entre otros, muestran lo inapropiado de utilizar mínimos cuadrados ordinarios para estimar (9.1.4) cuando se usan encuestas debido al sesgo obtenido cuando no se tiene en cuenta a toda la población. Los resultados muestran que bajo mínimos cuadrados ordinarios se sobrestima la verdadera magnitud del excedente del consumidor. Esta es una consecuencia del sesgo de truncamiento asociado con la colección de datos cuando se estima sólo una parte de la población real o cuando existen sesgos de información en la misma encuesta. De esta forma, asumiendo que la demanda por paisaje, derivada de

[25] En Deaton y Muellbauer(1980).

una encuesta, provee la información sobre la parte de la población que elige un determinado sitio por visitar, pero no toma en cuenta la información sobre otros grupos que demandan paisaje como serían los ganaderos, los pastores, etc. [26] o sobre los que no viajan aún cuando pudieran demandar paisaje, cualquier estimación bajo mínimos cuadrados ordinarios mostraría sesgos de truncación. Siguiendo la especificación de Englin y Schonkwiler(1995) la función de demanda tiene la siguiente forma semilogarítmica

(9.5) $E(Z_i, X_i)= u_i= e^{X^*iB}$ donde $X^*_i=[P^*_i|X_i] = \left[\phi' \Lambda'_P \sum^{-1} P_i | X_i \right]$

Esto significa que una de las variables independientes se construye a partir de un modelo Latente P^*_i, y de las variables X_i definidas en la ecuación de demanda (9.4). Teniendo en cuenta que un estimador es consistente, si los términos de los errores son normales, [Grogger y Carson(1989)] y, obteniendo una función de verosimilitud conjunta para un modelo de variables latentes y truncado, obtenemos un estimador máximo verosímil basado en la función de densidad de Z_{ji} la cual es truncada a una normal [27]

$$(9.6)\ LLikelih = -\frac{N}{2} Ln\left|\sum(\theta)\right| - \frac{N}{2} Traza\left[S_{WW} \sum^{-1}(\theta)\right] + \sum_{i=1}^{N}\left[\frac{1}{\sigma}\psi\left[V_{ji} - \beta_j X_{ji}^{\ *}\right]\sigma^{-1}\left[1-\beta_j X_{ji}^{\ *}\right]\sigma^{-1}\right]$$

La ecuación (9.6) representa la función de demanda por paisaje, en donde la variación compensatoria y la equivalente son iguales al excedente del consumidor, que a su vez viene definido como

$$(9.7)\ CS \approx \int_{Pmin}^{Pmax} Z(P)dp = -\frac{Z^2}{2(\frac{1}{A_1})}$$

Donde A1 es el valor estimado del parámetro en P, y Z el número de viajes. De Kealy y Bishop (1986), Bockstael, Hanemman y Strand (1987) se conoce que Cs es sesgado y de la forma $\approx \sigma^2 \left(\frac{1}{A_1}\right) * \left(\frac{1}{A_1}\right)^2$ ó (1/ (t-ratio)^2), y el excedente agregado del consumidor será

[26] Willis y Garrod(1991) justifican su uso, ya que se sobrestima la verdadera magnitud del excedente del consumidor de la siguiente forma: “Esta una consecuencia de los sesgos de truncar asociados con la colección de datos, los cuales surgen porque en el modelo de demanda estimado sobre encuestas existen sesgos conducidos en cada lugar de recreación, si bien, tales encuestas son usadas comúnmente en estudios de demanda por recursos, ellos no proveen información sobre los individuos que no eligen usar un lugar” Pág. 39.

[27] Englin y Schonkwiler(1995), utilizan una binomial negativa truncada con variables latentes, el hecho de que el parámetro de sobredispersión diera mayor que cinco, muestra la existencia de sesgos al 95% según Grogger y Carson(1989,pag32), siendo de esta forma los resultados inconsistentes. Por esta razón, y con base en los resultados de Grogger y Carson(1989) se propuso una función truncada a una normal con variables latentes.

(9.8) $\sum_{j=1}^{n} Z_j Cs$

Donde n es la población total de visitantes, Z el número de viajes por persona y Cs el excedente del consumidor por viaje definido en (9.7).

Para estimar (9.7) y (9.8) se realizó una encuesta aleatoria en la zona de Alameda del Valle(España), situada a 92 Kilómetros al norte de Madrid. El lugar tiene 2.520,8 Ha de zonas especialmente protegidas y 20,2 Ha de suelo urbano. La configuración del paisaje depende de una serie de características particulares como la presencia de pinos y robles sembrados en Montes de propiedad pública, la utilización de cercas de piedra para encerrar al ganado y que hoy día tiene poca función, pues la actividad ganadera viene decayendo a raíz de las medidas tomadas por el gobierno español para la inserción de España a la Comunidad Económica Europea. Esta serie de características muestran un paisaje rural, donde la actividad del hombre ha sido decisiva en la configuración de dicho paisaje. El tamaño de la encuesta es de 70 personas a quienes se les preguntó acerca de características como la vegetación anteriormente descrita, las cercas de piedra, el aspecto rural, y la relación entre el hombre y la naturaleza. Se preguntó además características individuales como sexo, edad, ingresos, profesión, años de educación, el tamaño del grupo con que se desplazaba y, el tipo de transporte utilizado. De acuerdo con estas preguntas se obtuvo que el 68.58 % de las personas que visitaban Alameda del Valle eran hombres, el 77.15 % le gustan los paisajes con pino, el 70% las cercas de piedra, el 67.72% los paisajes con roble, el 75.72% le gustan los paisajes con aspectos rurales, el 64.25% le gusta la relación entre el hombre y la naturaleza, el 54.29% le gusta el paisaje de la zona por su valor histórico, el 78.58% los espacios libres, el 87.15% de las personas había visitado antes la zona, el 70% de las personas tiene como único motivo visitar la zona. En relación con preferencias por zonas aledañas, el 34.28% prefiere además de Alameda del Valle, la zona de Presa Pinilla y el 31.42% prefiere además el Paular. También se encontró que el 62.85% de las personas no pertenecen a grupos relacionados con el medioambiente. Con relación a las variables de respuesta cuantitativa se obtuvieron los siguientes datos

Tabla 9.1. Resumen estadístico de las variables

Variable	Observaciones	Media	Mínimo	Máximo
Educación	70	15.17143	0	19
Años	70	34.97143	18	70
Ingreso	70	161.4286	20	500
Número de visitas	70	9.357143	1	45
Costo implícito	70	7.133857	4.98	10.98
Costo total	70	11.3035	6.3217	25.605

De la encuesta realizada, se procedió a estimar (9.6) por máxima verosimilitud usando los costos implícitos, esto es, los costos de desplazamiento que se construyen como 2*92 km*0,026 mil pesetas más los gastos en el día reportados en la encuesta. Por otro lado, 0.026 son los costos de desplazarse en automóvil por Km. (incluye Gasolina, gastos de depreciación, aceite, etc.) y 92 es la distancia desde Madrid a Alameda del Valle, a este valor se le suman los costos de estar en el lugar(alimentación, etc. que fueron reportados en la encuesta). Cuando se le suman a los costos implícitos, el costo de oportunidad del tiempo se obtiene el costo total. Los resultados se pueden observar en la tabla (9.2).

Tabla 9.2. Demanda estimada del modelo de coste de viaje

Variable dependiente [28]: Número de viajes
Número de observaciones: 70

Variables	Modelo 1	Modelo 2
Sexo [29]/	-1.02807	0.69317
	(-0.916)	(0.596)
Educación	-0.840202	-0.984842
	(-1.805)	(-1.960)
Años	0.240119	-0.001238
	(0.462)	(-0.023)
Ingreso	0.0000132	0.154519
	(1.803)	(1.674)
Latpaisa [30]/	0.1765821	0.1757131
	(1.681)	(1.484)
Costo implícito	-0.4963451	
	(-1.849)	
Costo total		-0.1943029
		(-4.226)
Constante	17.4997	16.9545
	(2.186)	(1.973)
σ	2.6444	2.800848
	(0.592606)*	(0.6312835)*
χ^2 (6)=16.37 Prob>χ^2 =0.119		χ^2 (6)=12.80 Prob>χ^2 =0.0464
Pseudo R^2=0.1403		Pseudo R^2=0.1097
Log Likelihood=-50.153017		Log Likelihood=-51.93669

Likelihood Ratio Test [31]

[28] t entre paréntesis a excepción de (*) errores estandar.

[29] El resultado poco significativo en las variables de sexo y educación contrasta con los resultados obtenidos cuando se usaron los indicadores del paisaje, pues usando estos indicadores dichas variables fueron significativas.

[30] Esta variable se construyo a partir del modelo de variables latentes descrito en él capitulo 6.

χ^2 (5) = 13.43	χ^2 (5) = 12.49
Prob>χ^2 = 0.0197	Prob>χ^2 = 0.028
Information Matrix	
F(5,64)=1.67	F(5,64)=1.30
Prob>F=0.1545	Prob>F=0.2737

Para el primer modelo, se tuvieron en cuenta los costos implícitos. El coeficiente del precio es negativo y diferente de cero al 6%, igualmente resultaron significativos las variables educación, ingreso, la variable latente paisaje y la constante. Sin embargo, las variables sexo y edad no resultaron significativas [32]. En cuanto a la variable latente del paisaje (Latpaisa) su valor se aproxima al verdadero valor del paisaje [33]. Con relación con el excedente del consumidor, éste fue de 1007.36 pesetas con un sesgo de $\pm$ 29.25%. En el segundo modelo, se usó el costo total que es igual al costo implícito más el coste de oportunidad del tiempo [34], de aquí se observa, que el coeficiente del precio es negativo y significativamente diferente de cero, y las otras variables conservan sus signos y significancia. El excedente del consumidor es de 5302.06 pesetas, con un sesgo del 5,5999%, de donde se observa que los resultados no son significativamente diferentes con relación a los costos en su conjunto, usando variables latentes

[31] Maddala(1995) conjetura que Score Test es más eficiente que Lr-test, sin embargo Maddala no presenta resultados que invaliden Lr-test. Dado que $\chi^2_{(5)}$ al 99% = 15,09 y al 99.5%= 16.75 no se puede rechazar la hipotesis de homocedasticidad.

[32] Este resultado se debe a la especificación del paisaje como variable latente. En el modelo tradicional(usando los indicadores del paisaje) esta variables resultaron significativas.

[33] Cuando se usaron los indicadores del paisaje (pino, Roble, Cercas de Piedra, Aspecto Histórico) los resultados para estos valores eran inconsistentes. Con la nueva especificación se obtiene (1/0.49645)*(0.17499)=0.3557647 es decir 35,6% de cambio en el número de visitas ante una visión unidimensional del paisaje.

[34] El coste de oportunidad del tiempo se calculó como el tiempo de Viaje por 2 por (1/4) del salario del momento más el salario por(3/4) si el Bienestar es mayor del 60%. El tiempo de Viaje por 2 por (1/4) del salario del momento más el salario por (1/2) sí el Bienestar es igual al 50% y, el tiempo de Viaje por 2 por (1/4) del salario del momento más el salario por (1/4) si el Bienestar es inferior al 60%. Donde la variable Bienestar proviene de la encuesta y, capta que tanto consideran los agentes que le reporta beneficios el viaje a la zona de Alameda del Valle.

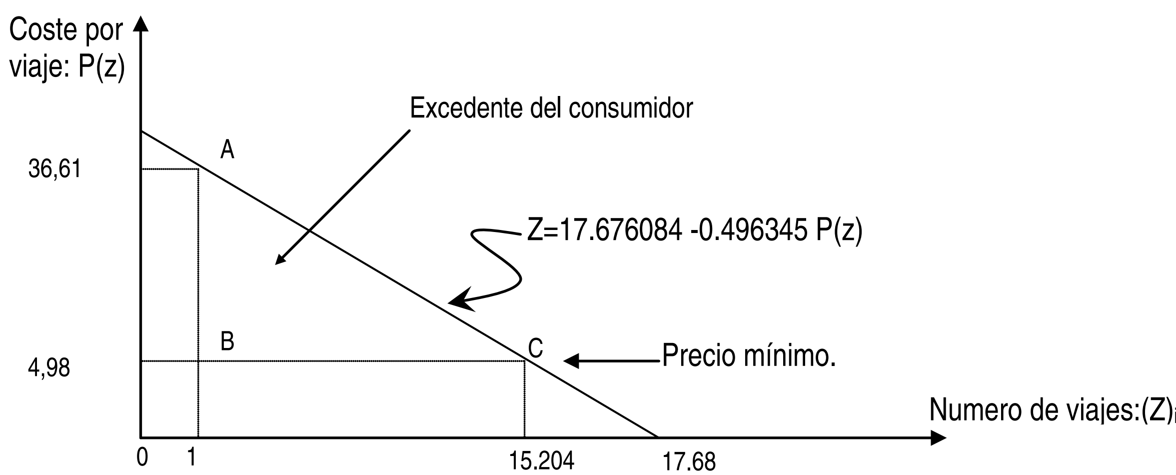
Tabla 9.3. Excedente del consumidor por viaje

Diferencia entre modelos	Costo implícito	Costo total
Paisaje como variable latente.	1007.36	5302.06
Paisaje como variable normal [35].	1398,81	5318,18
Variación del excedente del consumidor.	38.82%	3.5%

Si se toma el costo implícito, los resultados de modelar el paisaje como una variable latente, o usar los indicadores del mismo tienen grandes efectos sobre el excedente del consumidor. Si se toman los costos totales, el efecto sobre el excedente del consumidor no es muy significativo.

El excedente del consumidor se grafica a través de la curva de demanda, de la siguiente forma: Siguiendo a Balkan y kahn(1988) [36] y, usando el coste mínimo, se despeja el número óptimo de visitas (z) de la ecuación de demanda. Por otro lado, la ecuación de demanda en la gráfica (9.1) se calculó teniendo en cuenta que si la función de utilidad es lineal en el ingreso, la utilidad marginal del ingreso es constante y por lo tanto igual a 1. Con respecto a la variable Latpaisa su valor es de 1 mientras para las otras variables su valor será de cero.

Gráfica 9.1. Demanda estimada y excedente del consumidor



Si tenemos en cuenta el número de viajes óptimo derivado de la primera regresión, de acuerdo a la gráfica (9.1) el excedente del consumidor estimado por (9.7) sería de (15.20485)^2/2*(0.496345) el cual es 232.889 pesetas que es aproximadamente el área ABC. El excedente agregado sería de 16'302.292(232.889*70); si incorporamos el

[35] Esta regresión se calculó usando los indicadores del paisaje, esto es cercas de piedra, pino, roble y aspectos históricos junto a las variables de costo.

[36] Balkan y kahn(1988) usan el coste promedio.

sesgo(29,5%), su valor sería de 11'574.627 el cual es aproximadamente igual al valor obtenido a través de (9.7) esto es de 11'982.590. Por otro lado, la valoración social del paisaje a través del excedente agregado del consumidor es de 63'068.049 pesetas usando el costo total en el modelo de variables latentes y la ecuación (9.8).

9.1.2 El modelo de utilidad aleatorio

Parsons y Kealy(1992) consideran que un individuo toma el número total de viajes a un lago como predeterminado y decide cual lago visitar en cada viaje. El o ella, tiene una utilidad cuando viaja al lago (a_i) de la siguiente forma

(9.9) $\mu_{ai} = V_a + V_{ai} - \varepsilon_{ai} - \varepsilon_a$

V_a es un componente sistemático de utilidad común a todos los lagos en el área de Wisconsin
(a = 1 si el lago se localiza en el norte y a = 0 si se localiza en el sur). V_{ai} es un componente sistemático para el lago i en el área a (i=1,...,N sí está en el norte e i=1,...,S si está en el sur). El término $\varepsilon_{ai} + \varepsilon_a$ es un elemento aleatorio que captura las características excluídas del lago. La parte ε_a incorpora las características excluídas comunes a todos los lagos en el área a. Definiendo $V_a = V(X_{ai}, p_{ai})$ donde X_{ai} es un vector de las características del lago como el tamaño, facilidades comerciales, calidad del agua del lago y p_{ai} es el precio de visitar el lago incluyendo el costo de oportunidad del tiempo y los costos de viaje. Parsons y Kealy(1992) usan una función de utilidad lineal de la forma

(9.10) $V_{ai} = \beta Z_{ai}$ donde $Z_{ai} = (X_{ai}, p_{ai})$

Dado que los lagos del noroeste de Wisconsin tienen substanciales diferencias con respecto a los del sur, Parsons y Kealy definen $V_a = \alpha' d_a$ donde $d_a = 1$ cuando el lago se encuentra en el norte y $d_a=0$ cuando se encuentra en el sur. V_a captura una contribución "promedio" a la utilidad para un viaje tomado en el norte con relación a un viaje en el sur. De esta forma, la utilidad aleatoria para una visita a un lago (a_i) es

(9.11) $\mu_{ai} = \alpha' d_a + \beta Z_{ai} + \varepsilon_{ai} + \varepsilon_a$

Dado que un individuo decide cuando visitar un lago en el norte o en el sur se asume que ε_{ai} es una variable aleatoria idéntica e independientemente distribuida con un parámetro de escala $\delta'=1$. De esto se sigue, que la probabilidad individual de visitar el lago i´ dado que él o ella realizan un viaje al norte o al sur viene definida por el Logit

$$\Pr(i'|a=1)=\frac{e^{\beta Z_{1i}}}{\sum_{1i\in N} e^{\beta Z_{1i}}}$$

(9.12) $$\Pr(i'|a=0)=\frac{e^{\beta Z_{0i}}}{\sum_{0i\in S} e^{\beta Z_{0i}}}$$

Donde N es el conjunto de lagos que entra en el conjunto de elección cuando el individuo se desplaza al norte y S es el conjunto de lagos que entra en el conjunto de elección cuando el individuo se desplaza al sur. De aquí se sigue que

(9.13) $\text{Max}(\beta Z_{11}+\varepsilon_{11},\ldots\ldots, \beta Z_{1N}+\varepsilon_{1N})$
(9.14) $\text{Max}(\beta Z_{01}+\varepsilon_{01},\ldots\ldots, \beta Z_{0N}+\varepsilon_{0N})$

Son variables aleatorias con

(9.15) $E(\text{Max}(\beta Z_{11}+\varepsilon_{11},\ldots\ldots, \beta Z_{1N}+\varepsilon_{1N}))=I_1=\text{Ln}\left[\sum_{1i\in N} e^{\beta Z_{1i}}\right]+0.577$

(9.16) $E(\text{Max}(\beta Z_{01}+\varepsilon_{01},\ldots\ldots, \beta Z_{0N}+\varepsilon_{0N}))=I_0=\text{Ln}\left[\sum_{0i\in S} e^{\beta Z_{0i}}\right]+0.577$

Definiendo $\bar{\varepsilon}_1 = \text{Max}(\beta Z_{11}+\varepsilon_{11},\ldots\ldots, \beta Z_{1N}+\varepsilon_{1N})-I_1$ y $\bar{\varepsilon}_0 = \text{Max}(\beta Z_{01}+\varepsilon_{01},\ldots\ldots, \beta Z_{0N}+\varepsilon_{0N})- I_0$ y, asumiendo que $\bar{\varepsilon}_1 + \varepsilon_1$ y $\bar{\varepsilon}_0 +\varepsilon_0$ son i.i.d variables aleatorias con un parámetro de escala δ, entonces la probabilidad de elegir un lago en el norte o en el sur, será

$$\Pr(a=1)=\frac{e^{\alpha-\delta I_1}}{e^{\alpha+\delta I_1}-e^{\alpha-\delta I_0}}$$

(9.17) $$\Pr(a=0)=\frac{e^{\delta I_0}}{e^{\alpha-\delta I_1}-e^{\delta I_0}}$$

Dado que $\delta'=1$ en el lugar de elección, entonces $\delta/\delta'=\delta$ en el área de elección. También se asume que $\alpha=\alpha'\ \delta$ en el modelo. Si un individuo visita el lugar (ai) un total de T_{ai} veces durante el año, la probabilidad es $[\Pr(ai)]^{T_{ai}}$. Si el individuo visita más de un lugar, la probabilidad de las visitas será
$\Pi_a\ \Pi_i\ [\Pr(ai)]^{T_{ai}}$. La función de verosimilitud para el Logit vendrá dada por

(9.18) $L = \Pi_n \Pi_a\ \Pi_i\ [\Pr_n(i|a) \times \Pr_n(a)]^{T_{ain}}$

Donde $Pr_n(i|a)$ proviene de las ecuación (9.12) y Pr(a) de la ecuación (9.17) y n denota el n-ésimo individuo.

La variación equivalente y compensatoria para (9.18) ante un cambio en la calidad del agua, viene determinada por

$$(9.19)\ \nabla w = \left[\frac{1}{\delta\beta_y}\right][\,Ln[e^{\alpha+\delta\bar{I}_1} + e^{\delta\bar{I}_0}] - Ln\left(e^{\alpha+\delta I_1} + e^{\delta I_0}\right)\,]$$

Donde $\bar{I}_0$ y $\bar{I}_1$ se refiere a los valores de I_0 e I_1 con mejoramientos en la calidad del agua y, β_y es la utilidad marginal del ingreso, esto es, el coeficiente sobre P_{ai} en el modelo de utilidad aleatoria.

Dado que Wisconsin tiene una gran variedad de lagos, los autores proponen estimar el modelo de la siguiente forma: todos los sitios entran en el conjunto de oportunidades de la persona, pero el modelo se estima usando un subconjunto aleatorio extraído del conjunto total. De esta forma, cuando un individuo visita un lago en el norte, 23 lagos son extraídos aleatoriamente del conjunto de lagos (esto significa incluir todos aquellos en un radio de 180 millas desde el hogar) y se le adiciona al subconjunto el lago que visita actualmente. Este método también se usó para el sur. Los autores usan conjuntos de oportunidades aleatorias de 3,6,12 y 24 lagos.

MacFadden(1978) muestra que estimar el modelo de esta forma, da estimadores insesgados del modelo cuando se usa el conjunto de alternativas total. El resultado encontrado por Parsons y Kealy(1992) es el siguiente

Tabla 9.4. Demanda estimada del modelo de coste de viaje usando Máxima Verosimilitud para un Logit:

Datos para pesca

Número de lagos extraídos del conjunto de oportunidades

Variables	3	6	12	24	24(visitados)
PRICE	-.20(24.8)	-.27(33.8)	-.25(45.5)	-.23(52.4)	-.24(55.1)
LNACRES	.69(15.3)	.61(21.1)	.55(24.0)	.56(30.0)	.38(22.6)
CF	-.22(1.7)	.19(1.9)	.18(2.4)	.31(4.5)	.20(2.9)
REMOTE	.34(1.6)	-.47(3.4)	-.33(2.7)	-.48(4.3)	-.11(1.1)
LNMXD	.45(5.6)	.54(8.8)	.44(9.9)	.38(10.5)	.26(7.0)
BR	-.43(3.0)	-.61(6.9)	-.46(6.1)	-.28(4.3)	-.22(3.5)
INLET	.86(5.2)	.90(6.3)	93(8.7)	.63(7.0)	.64(7.0)
DONO	-.85(4.7)	-.88(6.7)	-.84(8.2)	-.79(8.8)	-.82(10.0)
DOYES	-.15(0.8)	-1.02(6.5)	.30(2.5)	.11(1.0)	.33(4.0)
CLEAR	----	----	----	----	----
INC VALUE (δ)	.20(9.9)	.16(9.8)	.17(9.9)	.18(9.9)	.18(9.6)
NORTH	.54(3.5)	.53(3.6)	.52(3.5)	.53(3.5)	.59(3.8)
Número de Visitas	3.598	3.598	3.598	3.598	3.598
Número de Individuos	239	239	239	239	239
Primer estado:					
Log Likelihood	-790	-1,599	-2,670	-4,154	-5,162
Pendiente = 0 Log-L	-3.954	-6,448	-8,943	-11,437	-11,437
Segundo estado:					
Log Likelihood	-338	-338	-338	-338	-337
Pendiente = 0 Log-L	-438	-438	-438	-438	-401

Donde la variable PRICE es el costo de oportunidad del viaje, esto es, 1/3*[Ingreso anual/2080]*Tiempo de Viaje más [0.10*2*distancia al lago]. LNACRES es el logaritmo de los acres que tiene el lago. CF que es igual a 1 si el lago tiene facilidades comerciales y cero si no. REMOTE que es igual a 1 si el lago es navegable y cero si no. NORTH si el lago está en el norte y cero si no. LNMXD que es el logaritmo de la máxima profundidad del lago. BR que es igual a 1 si existen rampas para botes y cero si no. INLET que es igual a 1si el lago tiene ensenadas y cero si no. DONO que es igual a 1 si el hypolimnion está vacío de oxígeno y cero si no. DOYES que es igual a 1 si el oxígeno disuelto en el hypolimnion es mayor que 5 ppm y cero de otra forma. CLEAR que es igual a 1 si la profundidad promedio es al menos de 3 metros y cero de otra forma. Como puede observarse, la mayoría de las variables fueron significativas dados sus valores t(entre paréntesis). Parsons y Kealy consideraron para el análisis de cambio en el bienestar las variables DONO y DOYES que median el cambio en la calidad del agua como se mencionó anteriormente. Específicamente cuando el número de lagos es de 24, se encontró que el valor fue de US$ 0.50 para la pesca, manteniéndose el patrón en todas las alternativas [37].

9.2 El método de los precios hedónicos

Cuando los individuos adquieren un bien en el mercado, su adquisición se realiza en tanto tiene una serie de atributos que el consumidor desea [38]. Sin embargo, como se observó en capítulos anteriores, algunos bienes podrían tener más de un atributo ¿Quién usa el tiempo de ocio sólo para ver televisión? Como bien lo plantean Atkinson y Halvorsen (1984) muchos bienes pueden ser vistos como canastas de atributos individuales que tienen mercados explícitos. En el capítulo 5 se encontró que los atributos de los bienes entraban directamente en la función de producción de hogares, que en adelante será nuestra función de utilidad, aunque no tenían un mercado explícito pues lo que observaban los agentes eran los precios de los bienes. En esta sección, se presentará una línea de investigación que pretende avanzar en algunas de las ideas planteadas en dicho capítulo.

Rosen (1974), propone una técnica de estimación de atributos en dos etapas: Primero, el precio de una bien se regresa en términos de sus atributos. Y, la derivada parcial del precio del bien con respecto a un atributo se interpreta como el precio marginal implícito. En la segunda etapa los precios implícitos estimados son usados para estimar las demandas inversas de los atributos.

La segunda etapa de Rosen, puede producir algunos "riesgos" como la multicolinealidad entre los atributos, generando un cambio en los signos esperados [Hogarty(1975),Deaton y Muelbauer(1980), Atkinson y Halvorsen(1984)].

El modelo de Precios hedónicos, puede plantearse de la siguiente forma: Supongamos un consumidor con un vector de características socioeconómicas α que deriva su utilidad de consumir varias características de un bien g que tiene una serie de atributos z_1, z_2,.., z_n (por supuesto, algunos atributos son medio ambientales como la polución, etc.) y de un bien numerario x. Sea la función de utilidad

(9.20) $\mu = \mu\,(z_1,z_2,\ldots,z_n\ , x, \alpha\ )$

Que se maximiza con respecto a la restricción presupuestal

(9.21) $Y = x + P(z_1,z_2,\ldots,z_n\ )$

[37] Los autores estiman el modelo también cuando los visitantes usan el lago para natación, pesca, vela, y por paisaje. Los resultados aquí presentados son específicamente para pesca, el excedente para natación es de US$0.83, para vela de US$0.19 y para paisaje de US$0.15.

[38] Ver al respecto el capitulo 7.

De las ecuaciones anteriores deberá quedar claro que la función de utilidad es débilmente separable en el sentido de Maler(ver cap. 5), esto es los atributos $Z_{i's}$ son débilmente separables de los otros bienes. Dada la débil separabilidad una elección por los atributos puede ser analizada de maximizar la función de subutilidad sujeta a las restricciones de gasto del bien en cuestión

(9.24) $Y_g = C(g) + \sum_{t=1}^{T} \frac{1}{(1+r)^t} Pg$

Se supone que Y_g es la parte del gasto asignada al bien g. Además, existen unos costes fijos de consumir g, C(g). Y, el bien se deprecia a una tasa de descuento de r, una mejor formalización podría incluir una tasa de preferencia ρ en la especificación. El lagrangiano para este problema de maximización será

(9.25) $L = \mu(g,x) + \lambda[Y_g - C(g) - \sum_{t=1}^{T} \frac{1}{(1+r)^t} P_t g_t]$

De las condiciones de primer orden, encontramos que la demanda para el bien g depende de las características socioeconómicas y que $\frac{\partial P_g}{\partial z_{gj}}$ indicaría la disposición marginal a pagar por una unidad adicional de la misma, esto es, su precio implícito.

Por otro lado, la existencia de restricciones lineales o no lineales, podría volver algo complejo el problema como en Palmquist(1984). En últimas, una función lineal implicaría que los precios implícitos de los diferentes atributos permanecieran constantes cualquiera que fuese el nivel de partida, implicando una combinación aditiva entre estos. En cuanto a las restricciones no lineales, el precio implícito cambiará en tanto cambien las características con relación a la cantidad consumida, esto significa que la importancia marginal del atributo cambiará de acuerdo al tipo de especificación (Logarítmica, Semilogaritmica, Cuadrática, Exponencial o Box-cox)[Ver Azqueta p138].

Un problema adicional surge en la estimación: La identificación como bien menciona Palmquist, si la forma funcional correcta fuese conocida, no existirían problemas de identificación. Sin embargo si algunos atributos no son incluídos, obviamente existe un problema de identificación. Atkinson y Halvorsen(1984) asumen funciones de utilidad Homotéticas y, de este forma, ecuaciones hedónicas no lineales nos darían los cambios en los precios marginales. La homoteticidad asignada escala las compras de los individuos con diferentes ingresos, lo que da el número de observaciones necesarias sobre la curva de indiferencia. Brown y Mendelsohn(1984), Brown y Rosen(1982) y Palmsquist(1984) presentan como método alternativo usar datos de mercados espacial o temporalmente diferentes, de esta forma, separan las ecuaciones hedónicas a ser estimadas en cada mercado. La variación entre los precios de mercado en los diferentes mercados permite identificar las funciones de demanda. A continuación, se presentará el procedimiento realizado por Atkinson y Halvorsen(1984).

Sea W = w(a, X) la función de utilidad, donde a es un vector con n componentes de atributos de un automóvil además de la eficiencia del mismo. Y, X es un vector de los otros bienes. Asumiendo separabilidad débil en W = W[μ(a),X] donde μ(a) la función de subutilidad del automóvil. En cumplimiento de la separabilidad débil, la restricción sobre el gasto en un automóvil, vendrá definida por

(9.26) $Z = C(a) + \sum_{t=1}^{T} \frac{1}{(1+r)^t} \frac{P_t M_t}{E(a)}$

Donde Z es el valor presente de los gastos de los servicios que presta el automóvil sobre la vida del automóvil, C(a) son los costos del capital del automóvil, r es una tasa de descuento, P_t es el precio esperado de la gasolina en el año t, M es el número de Millas conducidas en el año t, E(a) es la eficiencia del automóvil expresada en millas por galón, y T es la vida esperada del automóvil. El lagrangiano para este problema de maximización de la subutilidad del automóvil será

(9.27) $L = U(a) + \mu\, [Z - C(a) - \sum_{t=1}^{T} \frac{1}{(1+r)^t} \frac{P_t M_t}{E(a)}\,]$

Las condiciones de primer orden, implican

(9.28) $\frac{\partial U}{\partial a_i} = \mu\, [\frac{\partial C}{\partial a_i} - \sum_{t=1}^{T} \frac{1}{(1+r)^t} P_t M_t \frac{\partial E}{\partial a_i} E^{-2}\,]$

(9.29) $Z = C + \sum_{t=1}^{T} \frac{1}{(1+r)^t} P_t M_t E^{-1}$

En (9.28) se muestra que la utilidad marginal de cada atributo deberá ser igual al costo marginal del capital más los costos marginales de los gastos en la gasolina. Los autores asumen que la eficiencia de la gasolina puede incrementarse solamente cuando decrecen algunos de los atributos deseados por los consumidores. Las elecciones tecnológicas que incrementan la eficiencia de la gasolina sin que decrezcan los otros atributos, formalmente pueden definirse de acuerdo a la siguiente relación c = c(a,F) y E = E(a,f) donde F representa la extensión del ahorro de gasolina por la tecnología incorporada en el automóvil. En últimas, las millas recorridas son una función de los atributos del automóvil y de la eficiencia en la gasolina, M= M[a,E(a)]. Diferenciando totalmente se encuentra

(9.30) $\frac{\partial M}{\partial a_i} = \frac{\partial M}{\partial a_i} + \frac{\partial M}{\partial E}\frac{\partial E}{\partial a_i}$; $\frac{\partial M}{\partial E} > 0, \frac{\partial E}{\partial a_i} < 0$

Los cambios en las cantidades óptimas de los atributos como resultados de cambios en los precios de la gasolina, se analizan a través de estática comparativa de la siguiente forma

$$(9.31)\begin{bmatrix} A_{11} & \cdots & A_{1n} & \theta E_1 - C_1 \\ \cdot & \cdots & \cdot & \cdot \\ \cdot & & \cdot & \cdot \\ A_{n1} & \cdots & A_{nn}. & \theta E_n - C_n \\ \theta E_1 - C_1 & \cdots & \theta E_n - C_n & 0 \end{bmatrix} \times \begin{bmatrix} \frac{\partial a_1}{\partial P_0} \\ \cdot \\ \cdot \\ \frac{\partial a_n}{\partial P_n} \\ \frac{\partial \mu}{\partial P_0} \end{bmatrix} = \begin{bmatrix} -\mu E_1 \frac{\partial \theta}{\partial P_0} \\ \cdot \\ \cdot \\ -\mu E_n \frac{\partial \theta}{\partial P_0} \\ E \frac{\partial \theta}{\partial P_0} \end{bmatrix}$$

Siendo P_o el período base del precio de la gasolina, $A_{ij} = U_{ij} - \mu(C_{ij} - \theta E_{ij} + 2\theta E^{-1}E_iE_j$ con i,j=1,…..,N, y los subíndices representan las primeras y segundas derivadas parciales con respecto a los atributos, esto es, $C_{ij}=\partial^2 C/\partial a_i \partial a_j$ y $\theta=\sum_{t=1}^{T}\frac{1}{(1+r)^t}P_t M_t E^{-2}$. De esta forma, θ es igual a la derivada parcial negativa del valor total de los gastos de la gasolina con respecto a la eficiencia, y también puede interpretarse como el beneficio marginal de un incremento en la eficiencia de la gasolina ante un cambio en el precio del período base de la gasolina. Por lo tanto, la magnitud del efecto sobre los atributos, y de aquí sobre la eficiencia, podría depender sobre el nivel esperado de los precios.

Para estimar el anterior modelo, los autores seleccionan una serie de atributos que explican la variación en el costo de capital y la eficiencia en el combustible como la aceleración, A, el confort de un paseo, R, su estilo tradicional, S, y el confort de los asientos delanteros, C. Atributos como el prestigio del propietario y la calidad del trabajador podrían afectar el costo de capital sin tener efecto directo sobre la eficiencia del combustible. Como proxis de estas variables, los autores proponen incluir las siguientes variables falsas: si es un carro importado, I, si es un carro de lujo, L, y si es un carro especial, H. Estas variables falsas fueron incluidas en la ecuación de costos.

Los datos tomados incluyen 158 automóviles nuevos en 1978. Se propone entonces, especificar una función de subutilidad tipo Cobb-Douglas de la forma

(9.32) Ln U = $\gamma_0 + \sum_i \gamma_i Lna_i$; i=A,R,S,C; $\sum_i \gamma_i = 1$

La forma funcional para el costo de capital y la eficiencia del combustible se desarrolla a partir de la metodología Box-Cox. La forma estimada, fue

(9.33) $C^{(\psi)} = \alpha_0 + \sum_i \alpha_i a_i^{(\lambda)} + \sum_j \alpha_j D_j$

(9.34) $E^{(\phi)} = \beta_0 + \sum_i \beta_i a_i^{(\gamma)}$; i = A , R, S, C; j = I, L, H.

Las transformaciones Box-Cox $C^{(\psi)}$, $a_i^{(\lambda)}$,$E^{(\phi)}$,$a_i^{(\gamma)}$, tienen la forma $V^{(\delta)} = (V^{\delta} - 1) / \delta$ con $\delta \neq 0$ y $V^{(\delta)}$ =Ln V con δ=0. Estas transformaciones son continuas alrededor de δ=0 dado que el límite en el caso de que $\delta \neq 0$ cuando $\delta \to 0$ es Ln V. La forma funcional Box-Cox tiene varios casos especiales, el log-lineal cuando $\psi = \lambda = 0$, el lineal cuando $\psi = \lambda = 1$, la forma semi-log cuando $\psi = 0$ y $\lambda = 1$ y, la inversa semi-log cuando $\psi = 1$ y $\lambda = 0$. Los autores eligieron la forma Log-Lineal después de desechar las otras formas a través de los valores de las funciones de verosimilitud.

Tabla 9.5. Parámetros estimados del modelo de precios hedónicos

Variables	Función de costo de capital	Función de eficiencia del combustible	Función de subutilidad del automóvil
Intercepto	4.5215	7.0532	
	(11.0998)	(28.8119)	
Estilo tradicional	0.7040	-1.2678	0.3123
	(2.5056)	(6.8549)	
Aceleración	0.5638	-0.2940	0.2501
	(5.4632)	(3.6804)	
Confort del Asiento delantero	0.8364	-0.6783	0.3710
	(3.2849)	(3.3822	
Confort del Viaje	0.1502	-0.0996	0.0666
	(1.7098)	(1.5615)	
Hecho en el Extranjero	0.3079	-----------	
(7.4093)			
Carro especial	0.1458	-----------	
(2.9509)			
Carro lujo	0.5597		
(10.1551)			
Núm. observaciones	158	158	
R^2	0.71	0.71	

(Razones t entre paréntesis)

Los coeficientes de las variables estilo tradicional, aceleración, confort del asiento delantero el intercepto, hecho en el extranjero, carro especial y carro de lujo son significativos al 99% mientras la variable confort del viaje es significativa al 90%. Los coeficientes de los atributos en la función de subutilidad son iguales a las partes de los gastos. Estos valores fueron calculados a partir de la función de costo de capital imponiendo la restricción de homogeneidad $\gamma = \dfrac{\alpha_i}{\sum_i \alpha_i}$. Diferenciando la ecuación estimada del costo del capital, la ecuación de eficiencia del combustible y la función de subutilidad, se encuentra una aproximación a (9.31), esto es

(9.32) $C_i \equiv \dfrac{\partial C}{\partial a_i} = \dfrac{\partial LnC}{\partial Lna_i}\dfrac{C}{a_i} = \dfrac{\alpha_i C}{a_i}$

(9.33) $C_{ij} \equiv \dfrac{\partial^2 C}{\partial a_i \partial a_j} = -\dfrac{\alpha_i C}{a^2{}_i}; \forall\ i = j$

$= 0; \forall\ i \neq j$

(9.34) $E_i \equiv \dfrac{\partial E}{\partial a_i} = \dfrac{\partial LnE}{\partial Lna_i}\dfrac{E}{a_i} = \dfrac{\beta_i E}{a_i}$

(9.35) $E_{ij} \equiv \dfrac{\partial^2 E}{\partial a_i \partial a_j} = -\dfrac{\beta_i E}{a^2{}_i}; \forall\ i = j$

$= 0; \forall\ i \neq j$

(9.36) $U_{ij} = \dfrac{\partial^2 LnU}{\partial a_i \partial a_j} = \dfrac{\gamma_i}{a^2{}_i}; \forall\ i = j$

$= 0; \forall\ i \neq j$

Para calcular el valor del beneficio marginal de la eficiencia del combustible, θ, se debe especificar las relaciones entre los precios esperados de la gasolina y el precio del período base. Los autores eligen una especificación que implica una elasticidad unitaria gasolina-precios esperados con respecto al período de base, de la forma siguiente

(9.37) $P_t = P_0(1+f)^t$

Donde P_0 es el precio del período de base y f es la tasa esperada de un incremento en el precio real de la gasolina. Sustituyendo P_t en θ, se obtiene la expresión

(9.38) $\theta = \sum_{t=1}^{T} P_0(1+b)^t M_t E^{-2}$

Siendo (1+b) = (1+f)/(1+r). Con el fin de calcular θ se asumió que P_0 = U$0.70 y b=0.2. Por otro lado, el valor de μ es calculado usando la fórmula

(9.39) $\mu = \dfrac{\sum_i \dfrac{\gamma_i}{a_i}}{\sum_i \left[\dfrac{\alpha_i C}{a_i} - \theta\dfrac{\beta_i E}{a_i}\right]}$

Esta ecuación es simplemente la suma de la ecuación (9.28) sobre i y luego se despeja μ.

Dados los valores estimados de (9.31) el sistema de ecuaciones se soluciona para las derivadas parciales de los atributos con respecto al período de base de la gasolina. De esta forma, la elasticidad de demanda para el atributo i respecto al precio de período de base se calcula como

$$(9.40)\quad \frac{\partial a_i}{\partial P_0}\frac{P_0}{a_i}; \qquad i = A, R, S, C$$

Los resultados encontrados, se muestran en la siguiente tabla:

Tabla 9.6. Elasticidades precio-gasolina de las demandas estimadas

Elasticidades de demanda para los atributos.

Modelo	Eficiencia gasolina	Estilo tradicional	Aceleración	Confort del asiento delantero	Confort del viaje	Elasticidad de demanda para la eficiencia del combustible
Cadillac El Dorado	11	-2.19	0.05	1.97	0.60	1.36
Oldsmobile Tornado	15	-1.90	0.06	1.64	0.52	1.23
Ford Fairmont	19	-2.22	0.25	1.96	0.77	1.33
Mercury Bobcat	24	-3.32	0.28	3.24	1.13	1.82
Chevrolet Chevette	28	-2.01	0.13	1.74	0.61	1.27
Plymouth Arrow	33	-2.22	0.06	2.00	0.61	1.38
Toyota Corolla	39	-1.94	0.07	1.68	0.54	1.25

Como se puede observar, un incremento en el precio de la gasolina decrece la demanda por estilo tradicional e incrementa la demanda por confort en los asientos delanteros y en el confort del viaje. El efecto neto de un cambio inducido en los atributos de los automóviles será un incremento en la eficiencia del combustible. Dados los valores de la elasticidad de la eficiencia del combustible, es posible pensar que en el largo plazo la elasticidad propia de la demanda por gasolina es mayor que uno, esto significa que responde más que proporcionalmente a un cambio en los precios esperados del combustible.

9.3 El Método de la valoración contingente

El método de la valoración contingente busca obtener la valoración que otorga un individuo ante un cambio en el bienestar, como producto de una modificación en las condiciones de oferta de un bien, como podría ser el bien ambiental. Es un método directo, en tanto, la única forma posible de encontrar dicha valoración es preguntándosela al individuo. En este sentido, el método de la valoración contingente busca que el individuo revele lo que estaría dispuesto a pagar por una mejora(o por evitar un empeoramiento), o la cantidad exigida como compensación por un daño(o a renunciar a una mejora). El mecanismo de encuesta, como ya han mencionado Azqueta(1995), Mitchel y Carson(1989) tiene, entre otros problemas, el punto de partida, el problema del tiempo, el tipo de sesgo generado en la respuesta, el sesgo de información y el sesgo de hipótesis. Sin embargo, a partir de los informes presentados por Kenneth Arrow y Robert Solow(1993) a la National Ocean and Atmospheric Administration(NOAA) se concluye que el método proporciona una estimación confiable, siempre y cuando, se pregunte por la disposición a pagar, se use el formato binario (o de referendum) y se recuerde constantemente al entrevistado la gran cantidad de mejoras al medioambiente que compiten por una serie de recursos financieros escasos, dada la limitación presupuestaria.

Dadas las diferencias entre la disponibilidad a pagar o la compensación exigida, los modelos de valoración contingente se centran en las funciones de utilidad indirectas o las funciones de gasto. Aquí se presentarán ambas versiones, desarrolladas por Hanemann(1984) y Cameron(1987) y luego la versión presentada por MacConell(1988).

9.3.1 La función de gasto y la función de utilidad

El modelo de referéndum se basa en respuestas binarias (si o no) de los individuos y es usado como una medida del cambio de riqueza. El supuesto implícito consiste en que las respuestas individuales, en forma discreta, provienen de la maximización de la utilidad. Dicha maximización implica una respuesta acorde a la función de utilidad típica. Considere la respuesta a la pregunta ¿ Aceptaría usted un cheque por $ X para renunciar a los derechos de uso de este recurso durante un año?. Suponga que la función de utilidad es la siguiente

(9.41) $u = v_j(y) + \varepsilon_j$

Sea j = 1 para la situación inicial, cuando existe acceso al bien, y j = 0 para la situación cuando no existe acceso al bien. Sea y el ingreso y ε_j el término aleatorio de error. Un individuo (i) podría responder sí, a la pregunta, cuando

(9.42) $v_1(y+x) \geq v_0(y) + \varepsilon_1 - \varepsilon_0$

Para Hanemann la respuesta depende del nivel de utilidad indirecta en ambos estados, y de esta forma, la función de respuesta es la diferencia en las funciones indirectas de utilidad.

La interpretación de Cameron de la respuesta, parte de la función de gasto: Sea $m_j(u_0 + \eta_j)$, η la cantidad de dinero necesaria para que un individuo alcance el nivel de utilidad corriente (u_0) y, sea η_y el término aleatorio de error, j = 1 en la situación corriente y J = 1 para la situación sin acceso al recurso. Una respuesta sí, implica

(9.43) $X > m_1(u_0) - m_0(u_0) + \eta_1 - \eta_0$

A partir de (9.43) MacConnell define la función de variación, como

(9.44) $s(m_i, u_i) = m_1(u_0) - m_0(u_0)$

La función (9.44) se denomina función de variación debido a que puede ser considerada como la variación equivalente o compensatoria dependiendo de la pregunta realizada. Sin el elemento aleatorio, η_j, los modelos que parten de la función de utilidad indirecta o del gasto serán idénticos, esto es, estrictamente serian iguales si las partes estocásticas fuesen cero, entonces m será igual a u. De forma, cuando la utilidad marginal del ingreso es constante e independiente del constante en todos los estados $\left[\frac{\partial u_1}{\partial y} = \frac{\partial u_0}{\partial y}\right]$ y, constante a través de todos los individuos en la muestra [$u_{rl} = u_{sl} = 0$], las distribuciones $\eta_1 - \eta_0$ y $\varepsilon_1 - \varepsilon_0$ son transformaciones lineales entre sí. Sin embargo, cuando no existe una utilidad marginal del ingreso constante y cuando se adicionan los términos de error, m y u no son iguales.

Suponga, que la pregunta de partida consiste en ¿ Estaría usted dispuesto a aceptar $ X por renunciar al uso de un recurso por un año? Entonces, el valor del acceso al recurso, consistiría en la variación compensatoria para un cambio en los precios, entre la situación inicial y un precio de choque para aquel bien cuyo acceso ha sido eliminado o restringido. Si el consumidor responde que no, entonces es lógico pensar que la variación compensatoria seria superior a la cantidad $ X propuesta. Como la variación compensatoria se puede calcular directamente de la función de gasto(Cap. 3, pp 40-43) entonces necesariamente

(9.45) $m(p^*, q, u) - m(p, q, u) > X$
$u(p, q, y) > u(p^*, q, y + X)$

Siendo p^* el vector de precios de choque, q un vector de calidad de los bienes consumidos. Dado que la función de utilidad no es observable y varía entre los individuos con diferentes ingresos, se puede hacer que la variación compensatoria sea igual a (9.44), esto es, igual a la función de variación. Asumiendo que ésta, no es determinística, entonces

(9.46) $CV = s(p, q, y) + \eta$

De esta forma, la probabilidad de responder Sí será

(9.47) Probabilidad [aceptar X] = Probabilidad [X > s(p, q, y) + η_j]
= Probabilidad [X - s(p, q, y) $\geq \eta_y$]

De igual forma, se podría estimar s(p, q, y) – X a través de estimar la probabilidad de responder No. Por otro lado, $\frac{\partial s}{\partial y}$ será mayor que cero cuando se trata de un bien normal, igual a cero cuando sea un bien neutral, y menor que cero. Es de esperar, cuando se refiere a un bien superior que $\frac{\partial s}{\partial q}$ sea mayor que cero, ya que el valor marginal de un incremento en la calidad del bien será positivo.

9.3.2 Estimación por máxima verosimilitud con datos de "referéndum"

Gran parte de los trabajos de valoración contingente usan modelos de elección dicotómica tipo Logit con datos de referéndum, y luego se integra el área bajo la curva [Cameron(1987a,b), Bishop y Heberlein(1979), Haneman (1984)]. Cameron y Huppert(1991) proponen que el modelo de regresión sea censurado normalmente.
Suponga que la verdadera valoración de aquel individuo que responde es Y_i y que Log $Y_i = X'_i\beta + u_i$, siendo u_i normalmente distribuido con media cero y varianza σ. Bajo un escenario de disponibilidad a pagar, al individuo se le ofrece un valor singular de umbral t_i. Si el individuo está dispuesto a pagar esta cantidad, entonces la disponibilidad a pagar, Dp_i, será igual a 1 y cero en caso contrario. De esta forma, se puede asumir

(9.48) Probabilidad (Dp_i =1) = Probabilidad (Log Y_i > Logt $_i$) = Probabilidad(u_i > Logt $_i$ - $X'_i\beta$)
= Probabilidad (u_i / σ > (Logt $_i$ - $X'_i\beta$) / σ)
= 1 - Φ[(Logt $_i$ - $X'_i\beta$) / σ]

De donde se deduce que la función de verosimilitud viene definida por

$$(9.49)\ \text{LogL}=\sum_{i=1}^{n}\left\{Dp_i Log\left[Dp-\Phi\left(\frac{Logt_i - X'_i\beta}{\sigma}\right)\right]+(1-Dp_i)Log\left[\Phi\left(\frac{Logt_i - X'_i\beta}{\sigma}\right)\right]\right\}$$

Cuando los datos provienen de un tipo de encuesta que pregunta sobre intervalos, generalmente se asigna un punto medio del intervalo relevante como proxi de la variable sobre el intervalo, de esta forma se usa mínimos cuadrados ordinarios donde dichos puntos medios son la variable dependiente. Cameron y Huppert(1991) encuentran la función de máxima verosimilitud para dichos intervalos, eliminando la subvaluación o sobrevaluación que proviene de escoger este punto medio. Dado que Y_i se conoce que pertenece al intervalo (t_{Bi} , t_{Ai}) , entonces el Log Y_i se encontrara entre Log t_{Bi} y Log t_{Ai}.

Sí Log $Y_i = X'_i\beta + u_i$ y si u_i tiene una distribución normal con media cero y varianza σ, entonces se puede estandarizar el rango de valores del Log Y_i de tal forma que

(9.50) Probabilidad ($Y_i \subseteq (t_{Bi}, t_{Ai})$)= Probabilidad [(Logt $_{Bi}$ - $X'_i\beta$) / σ < z_i < (Logt $_{Ai}$ - $X'_i\beta$) / σ]

$$= \left[\Phi\left(\frac{Logt_{Bi} - X'_i\beta}{\sigma} \right) - \Phi\left(\frac{Logt_{Ai} - X'_i\beta}{\sigma} \right) \right]$$

Dado que z_i es la normal estándar y que Φ es la función normal de densidad acumulada, y que z_{Bi} y z_{Ai} son respectivamente la normal para los límites bajos y altos, la correspondiente función de Log-verosimilitud para una muestra de n observaciones independientes viene definida como

(9.51) $$LogL(\beta,\sigma | t_{Bi}, t_{Ai}, x_i) = \sum_{i=1}^{n} Log\left[\Phi(z_{Bi}) - \Phi(z_{Ai}) \right]$$

Cameron y Hupper(1991) usan el siguiente tipo de pregunta, para realizar la valoración contingente ¿Cuánto es lo máximo que usted estaría dispuesto a pagar cada año para apoyar las actividades de restauración del hábitat cuyo resultado duplicase las tasas de captura de salmón y pescado rayado en la bahía de San francisco y, el área de océano, de tal forma que sin estos esfuerzos usted esperaría en esta área que los niveles de captura permanecieren en los niveles corrientes? (coloque un circulo en la cantidad). La lista de los valores es \$0, \$5, \$10, \$15, \$20, \$25, \$50, \$75, \$100, \$150, \$200, \$250, \$300, \$350, \$400, \$450, \$500, \$550, \$600 y “\$750 o más”. Se usaron 342 observaciones y las estadísticas son las siguientes:

Tabla 9.7. Estadísticas descriptivas.

Variable	Descripción	Media y Desviación Estándar
MIDPT	Punto medio del intervalo de respuestas.	57.98 (132.96)
Log(MIDPT)	Logaritmo de MIDPT.	3.115 (1.371)
TRIPS	Número de viajes para capturar salmón y striped bass en los 12 meses anteriores.	4.416 (5.449)
Log(INC)	Logaritmo del ingreso del hogar en miles de dólares usando el punto medio del intervalo reportado.	3.602 (0.6544)
S-TRIP	1 si todos los viajes fueron para pescar salmón, 0 de otra forma.	0.4188
B-TRIP	1 si todos los viajes fueron para pescar pescado rayado, 0 de otra forma.	0.3339
ADVCD	1 si tiene habilidades avanzadas de pesca, 0 de otra forma.	0.2812
OWNBOAT	1 si es el propietario del bote, 0 de otra forma.	0.3317
S-TARG	Número de Salmones/ Número de Viajes cuya finalidad es capturar salmón.	0.7341 (1.046)

De acuerdo a los anteriores datos se estimó (9.51) por máxima verosimilitud, los resultados encontrados fueron

Estimación por máxima verosimilitud. Variable dependiente: Log(Disponibilidad a pagar[definida en intervalos]). Numero de observaciones: 200.

Tabla 9.8. Parámetros estimados del modelo de valoración contingente

		Límite bajo		Límite alto	
Variable	Estimación puntual	Media	Máximo mínimo	Media	Máximo Mínimo
Constante	2.207 (5.151)	2.415 (2.762)	5.097 -0.3835	2.165 (2.752)	3.878 -1.351
TRIPS	0.03008 (2.023)	0.03802 (1.806)	0.1024 -0.01193	0.03430 (1.783)	0.1028 -0.02676
Log(INC)	0.3537 (3.177)	0.2804 (1.404)	0.9979 -0.2457	0.3414 (1.769)	1.160 -0.1092
S-TRIP	-0.4753 (-2.511)	-0.2653 (-0.7022)	0.6333 -1.313	-0.3460 (-1.048)	0.3893 -1.247
B-TRIP	-0.5884 (-2.956)	-0.6132 (-1.524)	0.4971 -2.171	-0.5908 (-1.759)	0.3826 -.1580
ADVCD	0.5824 (3.355)	0.4121 (1.405)	1.449 -0.4570	0.4664 (1.835)	1.220 -0.2779
OWNBOAT	0.4770 (2.982)	-0.2873 (1.073)	0.3374 -1.082	-0.3615 (-1.290)	0.4177 -1.166
S-TARG	-0.1798 (-2.175)	-0.1590 (-1.246)	0.1157 -0.5174	-0.1867 (-1.553)	0.05930 -0.5460
σ	1.292 (22.13)	1.166 (4.572)	2.214 0.7080	1.197 (5.302)	2.091 0.7435
Max Log L		-150.70 (14.88)		-157.16 (17.09)	

Los autores realizaron una serie de réplicas de la muestra, encontrando que mientras el promedio de la disponibilidad a pagar era de $58.50 con una desviación estándar de (34.78) en el límite bajo los valores estaban entre $28.42 y $384.54 y, en el límite alto los valores se encontraban entre $200 y $32.30. Como bien lo señalan los autores, las discrepancias encontradas se deben a la estrategia en los incentivos, a las interpretaciones de las preguntas y a la selección de la muestra. Este resultado no es sorprendente, si se tiene en cuenta que Bishop, Heberlein y Kealy(1983) ya habían mencionado las diferencias que se podían encontrar usando diferentes estimadores de esta valoración. Por otro lado, Alberini(1995) muestra que el investigador deberá hacer supuestos fuertes en torno a la distribución de la disponibilidad a pagar, y por lo tanto, supuestos distribucionales incorrectos generan funciones asimétricas de disponibilidad a pagar. Las simulaciones de Alberini muestran que el test de la ji-cuadrada (χ^2) tiene bajo poder a no ser que se tengan muestras grandes, incluso con mil observaciones el poder no es mayor del 30%. Por lo cual, una prueba de significancia del modelo (Probit en el trabajo de Alberini) podría rechazar la hipótesis nula solamente una tercera parte de las veces. Y podría llevar al investigador, a concluir que la valoración contingente es un modelo razonable cuando la verdadera disponibilidad a pagar es muy diferente de la que se ha asumido (pág. 92).

10 Bibliografía

Abbott, M and O, Ashenfelter. (1976). "Labor supply, commodity demand, and the allocation of time", Review of economics studies, vol. 43, pp.389-411.

Aigner, D and Goldberger, J. (1977). Latent variables in socioeconomic models. North Holland.

Alberini, A. (1995). "Testing willingness-to-pay models of discrete choice contingent valuation survey data", Land economics, vól.71, núm.1, pp.83-95.

Amemiya, T. (1994). Introduction to statistics and econometrics, Harvard University Press.

Amemiya, T. (1981). "Qualitative response models: a survey", Journal of economics history, vol. XIX, pp. 1483-1536.

Amemiya, T. (1974). " Tobit models : A Sorvey " Journal of econometrics, 24, pp- 3-63.

Anderson, G.J. (1987). "Prediction test in limited dependent variable models", Journal of econometrics, 34, pp.253-61.

Anderson, S.P., Palma, A and J.F, Thisse. (1995). Discrete choice theory of product differentiation, MIT press, Cambridge Mass.

Anderson, S.P., Palma, A and J.F, Thisse.(1992). "Interpretations of the logit discrete choice model in the theory of product differentiation" en J.M.E. Gee and G.Norman (comps.) Market structure and strategy, Hemmel Hempstead: Harvester Wheatsheaf.

Andrés, Javier; Jaume García y Sergio Jimenez. (1989). "La incidencia y la duración del desempleo masculino en España", Moneda y credito,189, pp.75-124.

Archibald, G.C and B.C. Eaton and R.G. Lipsey. (1986). "Address models of value" en J.E.

Arrow, K.J.(1971). Essays in the theory of risk beaning. Chicago:Markham.

Arrow, K.J et-al. (1968). "Labor substitution and economic efficiency", *Review of Economics and Statistics*, vól.43, pp.225-250, August, 1968.

Arrow, K.J. (1982). "Risk perception in psychology and economics", *Economic inquiry*, Vol. 20, pp.1-9.

Arrow,K., Colombato, E., Perlman, M y Schmidt,Ch. (1996). The Rational Foundations Of Economic Behaviour.

Arrow, K.J (1970a). “Exposition of the theory of choice under uncertainty”, in K. J. Arrow, Ensays in the Theory of Risk-Bearing, Amsterdam : North-Holland.

Arrow, K.J.(1970b). “The theory of risk aversion”, in K. J. Arrow, Essays in the Theory of Risk-bearing, Amsterdam : North-Holland.

Atkinson, S.E y R, Halvorsen. (1984). “A new hedonic technique for estimating attribute demand: an application to the demand for automobile fuel efficiency”,The review of economics and statistics, vól.66, núm.3, pp.417-426.

Azqueta, D. A. (1994). Valoración económica de la calidad ambiental, McGraw-Hill, España.

Balkan, E and J.R, Kahn. (1988). “The value of changes in deer hunting quality: travel cost approach”, Applied economics, 20, pp. 533-39.

Azqueta, O.D.(1994).Valoración económica de la calidad ambiental, Mc Graw-Hill.

Battalio, (1973). “A test of consumer demand theory using observations of individual purchases”, Western economic journal, Dec, pp.411-428.

Barten, A. P. (1964). “Family composition, prices and expenditure patterns”, in P.E. Hart, G. Mills, Barten, A. P and J. K Whitaker (eds.), Econometric Analysis for National Economics Planning, London : Butterworth.

Barten, A. P. (1969). “Maximum likelihood estimation of complete system of demand equations”. *European economic review*, Vol. 1, pp. 7-73.

Barten, A. P. (1966). “Theorie en empire van een volledig stelsel van vraaguergelijkingen”, Doctoral dissertation, Roterdam : University of Roterdam.

Barten, A. P and J. K Whitaker. (1977). “The systems of consumer demand functions approach: A review”, Econometrica, 45, pp.23-51.

Becker, G. S. (1965). “A theory of the allocation of time”, Economic Journal, vol.75, pp.493-517.

Bera, A.K., Jarque, C.M. y L.F, Lee. (1984). “Testing the normality assumption in limited dependent variable models”, International economic review, 25, pp.563-78.

Biddle, J.E and D.S, Hamermesh. (1990). “Sleep and the allocation of time”, Journal of political economy, vól.98, núm.5, pt.1, pp.922-943.

Bishop, R.C and T.A, Heberlein. (1979). "Measuring values of extra-market goods: are indirect measures biased", American journal of agricultural economics, núm.61, pp.926-930.

Blackorby, I . (1978). "Measures of quality and their meaning in terms of social welfare", Journal of economic theory, vól.18, pp. 59 - 80.

Blundell, R.W and Walker, I. (1982). "Modeling the joint determination of household labor variation", Economic journal, 92, pp.351-64

Blundell, R. and C, Meghir. (1986). "Selection criteria for microeconomic model of labor supply", Journal of applied econometrics, 1, pp.55-82.

Bollen, K. (1989). Structural equations with latent variables, John Wiley Sons.Inc.

Bockstael,N.E., Strand,I.E and M, Hanemann. (1987). "Time and recreational demand model", American journal of agricultural economics, 69, pp. 293-302.

Bockstael, N.E. and K.E, McConnell. (1983). "Welfare measurenment in the household production framework", American economic review, vól.66, pp.799-812,Sep.

Brannas, K and T, Laitila.(1989). "Heterocedasticity in the tobit model", Statistical papers, 30, pp.185-196.

Breush, T.S and A.R, Pagan. (1980). "The lagrange multiplier test and its applications to model specification in econometrics", Review of economic studies, 47, pp. 239-53.

Brown, G.Jr and R, Mendelsohn. (1984). "The hedonic travel cost method", The review of economics and statistics, núm. 6, August.

Brown, G.Jr and H.S, Rosen.(1982). "On the estimation of structural hedonic price models", Econometrica, núm.50, May, pp.765-768.

Brush, R.R and Galante, E. (comps.) Handbook of mathematical psychology, New York.

Byron, R.A. (1970). "A simple method for estimating demand systems under separability assumptions", Review of economics studies, 37, pp.261-74.

Cameron, T.A and D.D, Huppert. (1991). "Referendum contingent valuation estimates: sensitivity to the assignment of offered values", Journal of the american statistical association, vól.86, núm.416, pp.910-920.

Cameron, T.A and M.D, James. (1987). "Efficient estimation methods for use with ´closed-ended´ contingent valuation survey data", The review of economics and statistic, núm.69, pp.269-276.

Cameron, T.A.(1987). “The impact of grouping coarseness in alternative grouped-data regression models”, Journal of econometrics, (annals), núm.35, pp.37-57.

Cesario, F.J. and J. L, Knetsh. (1970). “Time bias in recreation benefit estimates”, Water resources research, 6, pp.700-704.

Chambers, E.A and Cox, D.R. (1967).“Discrimation between alternative binary response models”. Biometrika. Vol 4, nums (3-4), pp. 573-78.

Chamberlain, G. (1980). “Analysis of covariance with qualitative data”. Review of economic studies, Jan, pp. 225-38.

Chesher, A.D. (1984). “Testing for neglected hetoregeneity”, Econometrica, vol. 52, pp.865-72.

Chesher, A.D and Lancaster, T. and M. Irish. (1985). “On detecting the failure of distributions assumptions”, Annals de L´INSEE, 59, pp. 7-44.

Chipman,J.S. (1974). “Homothetic preferences and aggregation”, Journal of economic theory, vól 8, pp.26-38.

Christensen, L. R., Dale, W.J and Lawrence, J. l. (1975). “Transcendental logarithmic utility function”. *American Economic Review*, No. 65, pp.367-383, June.

Christensen, L, D Jogerson, L. Lau. (1971). “Conjugate duality and the trascendental logarithmic production function”. *Econometrica,* Jul, 255-56.

Costa, D.L. (1998). “The unequal workday: a long-term view”, The American economic review, vol.88, num.2, pp.330-334.

Cox, D.R. (1983). “Some remarks on overdispersion”, Biometrika , 70,pp.269-74.

Davidson, R and J.G, Mackinnon. (1989). “Testing for consistency using artificial regressions” Econometric Theory, 5, pp.363-84.

Davidson, R and J.G, Mackinnon..(1984). “Convenient specification test for Logit and Probit models”, Journal of econometrics, 25, pp. 241-62.

Deaton, A. (1981). Theoretical and empirical approaches to consumer demand under rationing, en

Deaton, A (comp,) Essays in the theory and measurement of consumer behavior, N.Y, Cambridge University Press.

Deaton, A and Muellbauer, J. (1989). Economics and consumer behavior. Ny. Cambridge University Press.

Deaton, A. (1974). "The analysis of consumer demand in the united kingdon,1900-1970", *Econometrica*, Vol.42, pp.341-67.

Deaton, A y Muellbauer, J. (1980). Economics and consumer behavior. *Cambridge University Press* (1989).

Deaton, A y Muellbauer, J (1980) "Almost ideal demand system", *American Economic Review*, Vol. 70, pp. 312-336.

Deaton, A. (1989). *El consumo*, Alianza Editorial.

Diewert, W.E. (1982). Duality approaches to microeconomic theory, en K Arrow and M.D.

Domencich, T.A and D, Mcfadden. (1975). Urban travel demand, Amsterdam, North holland.

Dudley, L. and Montmarquetle, C. (1976). "A model of the supply of bilateral foreign aid". American economic review, March pp.132-42.

Durbin, J. (1954). "Errors in variables", Review of the international statistical institute, 22, pp.22-32.

Dubin, J.A and D.L, Macfadden. (1984). "An econometric analysis of residential electric appliance holdings and consumption", Econometrica, vol.52, March, num.2, pp.345-362.

Edgell, S.E and W.S, Geisler. (1980). "A set-theoretic random utility models of choice behavior". Journal of mathematical psychology, 21, pp.265-278.

Edwards, B.K and M.R, Starr. (1987). "A note on indivisibilities, specialization, and economics of scale", American Economic Review, march, pp.192-195.

Englin, J and J.S, Shonkwiler. (1995). "Modeling recreation demand in the presence of unobservable travel cost: toward a travel price model", Journal of environmental economics and management, 29, pp. 368-77.

Eugene, S. (1985). "Nutrition and the demand for tastes", Journal of political economy, vól.93, núm.5, pp.881-900.

Eye Von, A and C, Clogg. (1994). Latent variable analysis: Applications for development research, Sage Publications.

Freeman III, A.M. (1993). The measurement of environmental and resource values: theory and methods, Resources for the future.

Frish, R. (1959). "A complete scheme for computing all direct and cross demand elasticities in a model with many sectors", Econometrica, 27, pp.367-97.

Gary, B. (1965). "A theory of allocation of the time", Economic journal, núm.75, pp.493-517, Sep.

Gauriéroux, C.A, Monfort, E.R y A, Trognon. (1987). "Generalized residuals", Journal of econometrics, 34, pp.5-32.

Georgescu-Roegen, W. (1958). "Threshold in choice and the theory of demand", Econometrica, vol.26, pp.157-168.

Goldberger, A.S. (1983). "Abnormal selection bias" en Karlin, Amemiya y Goodman (comps.), Studies in econometrics, Time series and multivariate analysis, New York, Academic Press.

Goldman, S.M and Usawa. H. (1964). "A note on separability in demand analysis", Econometrica, vól 32, jul.

Greehalg, C. (1980). "Participation and hours of work for married women in Great Britain", Oxford Papers, vól.32, pp. 296 - 318.

Green, W.H. (1999). Análisis econométrico, Tercera edición, Prentice Hall Iberia.

Grogger, J.T and R.T, Carson. (1991), "Model for truncated counts", Journal of applied econometrics, vól. 6, pp. 225-38.

Grogger, J. (1990). "A simple test for exogeneity in probit, logit and poisson regression models", Economics letters, 33, pp. 329-32.

Gronau R. (1976). "The allocation time of Israel women". Journal of Political economic, Aug, pp.201-20.

Gronau, R. (1973). "The effects of children on the housewife: value of time", Journal of political economy, vol. 81, supplement, pp. S168 - 99.
Hall, R.E. (1973). "Wages income and hours of work in the U.S. labor price" en Cain, B y Watts, H (Comps.), Income maintenance and labor supply, Chicago.

Ham, J.C. (1977). "Rationing and the supply of labor: an econometric approach" en Deaton, A and Muellbauer, J (Comps.) Economics and consumer behavior, (1989), Cambridge, Cambridge University Press.

Hamermesh, D.S. (1998). “When we work”, The American economic review, vol.88, num.2, pp.321-325.

Hanemann, M.W. (1984). “Discrete/continuous models of consumer demand”, Econometrica, vol.52, May, num.3, pp.541-561.

Hanemann,W.M. (1991). “Willingness to pay and willingness to accept: how much can they differ?” American economic review, vól.81, núm.3, pp.635-647.

Hanemann, W.M. (1984). “Welfare evaluations in contingent valuation experiments with discrete responses”, American journal of agricultural economics, núm.66, pp.332-341.

Hausman, J.A. (1981). “Exact consumer surplus and deadweight loss”, American economic review, vól.71, pp.662-676, Sep.

Heckman, J.J and T.E, MaCurdy. (1980). “A life cycle model of female labor supply”, Review of economic studies, vól.47, Jan, pp.47-74

Hogarty, T.F. (1975). “Price quality for automobiles: a new approach”, Applied economics, vól. 7, march, pp.42-51.

Holtz-Eakin, D., D, Joulfaian and H.S, Rosen. (1993). “The Carnegie conjecture: some empirical evidence”, The quarterly journal of economics, may, pp.413-435.

Houthakker, H.S. (1965). “A note on self-dual preferences”, *Econometrica*, No. 33, pp. 797-801, Oct.

Howe H R:A Pollack and T.J Wiles (1979). “Theory and time series estimation of the quadratic expendiure system”. *Econometric*, Vol. 47, No.5, pp.1231.

Intriligator (comps,) Handbook of mathematical economics, Vl II, North Holland Publishing.

Jorgenson, D.W and Lau, T.S. (1975). “The structure of consumer preferences”, Annals of economic and social measurement, 4, pp.49-101.

Kahneman, Knetsch and Thaler. (1990). “Experimental test of the endowment effect and the coase theorem” Journal of political economy, vól.98, núm.6, pp.1325-1348.

Kahneman, D and Tversky, A. (1992). “Advances in prospect theory: cumulative representation of uncertainty”, *Journal Of Risk And Uncertainty*, Vol. 5, pp. 46-55.

Kealy, M.J and R.C, Bishop. (1986) “Theoretical and empirical specifications issues in travel cost demand studies”, American journal of agricultural economics, 68, pp.660-67.

Kevin, L. (1966). “A new approach to consumer theory”, Journal of political economy, vól.74, pp.132-157, April.

Kiefer, N.M.(1982). “Testing for dependence in multivariate Probit models”, Bometrika, 69, pp.161-6.

Klein, LR. and H. Rubin. (1947-48) “ A constant utility index of the cost of living”. Review Of Economic Studies, 15, pp. 84-57.

Kreps, D.M. (1995). *Curso de Teoría Microeconómica*. Mcgraw-Hill.

Lancaster, K. J. (1966a). “A new approach to consumer theory”, Journal of Political Economy , vol.74, pp132-57.

Lancaster, K. J. (1966b). “ Change and innovation in the technology of consumption”, American Economic Review, vol.56, pp14-23.

Lancaster, T. (1984). “Test of specification in econometrics”, Econometrics review, 3, pp.211-42.

Lee, L.F. (1978). “Unionism and wage rates: A simultaneous equations model with qualitative and limited dependent variables. International economic review Jun, pp. 415-34.

Leontief, W. (1947). “Introduction to a theory of the internal structure of functional relationships”, Econometrica, 15, pp.361-73.

Luce, R.D and P, Supes. (1965). “Preference, utility and subjective probability” en Luce, R.D.,

Luce, R.D. (1959). Individual choice behavior: a theoretical analysis. New York: Wiley.

Luce, R.D.(1977). “The choice axiom after twenty years”, Journal of mathematical psychology, 15, pp.215-233.

Luenberger, D.L .(1995). *Microeconomic theory*, Mc Graw-Hill.

Mäler, K.G. (1974). Environmental economics: A theoretical inquiry, Baltimore:John Hopkins Press.

MacConell, K.E. (1988). “Models for referendum data: the structure of discrete choice models for contingent valuation”, Journal of environmental economics and management, núm.18, pp.19-34.

Manski, C.F. (1977). “The structure of random utility models”, Theory and decision, 8, pp.229-254.

McConell, K.E. (1992). “On site time in the demand for recreation”, American journal of agricultural economics, 74, pp.918-25.

Mackinnon, J.G. (1992). “Model specification test and artificial regressions”, Journal of economic literature, vol. 30, pp.102-46.

McDonald, J. A, and D.A Clelland. (1984). “Textile workers and union sentimen”. Social Forces 63; pp.502-521.

McFadden, D.(1986). “The choice theory of market research”, Marketing science,5,pp.275-297.

McFadden, D. (1999). “Rationality for economists?”. *Journal Of Risk And Uncertainty*, Dec.

MacFadden, D. (1978). “Modeling the choice of residential location” en A, Karlqvis (comp.), Spatial interaction theory and planning models, Amsterdam, North-Holland.

McFadden, D. (1974). “Conditional logit analysis of qualitative choice behavior” in fronters in econometrics edited by P. Zarembka, New York: Academic Press, pp. 105-42.

McFadden, D.(1981). “Econometric models of probabilistic choice”, in structural analysis discrete date, edited by C.F Mansky and D. McFadden. Cambridge Mass: Mit Press.

McFadden, D. (1984). “Econometric analysis of qualitative response models” en Z.Griliches (comp.), Handbook of econometrics, vol. 2.

MacRealy, G and C.M, Dayton. (1994). “Latent class model for longitudinal assessment of trait acquisition”, en Alexander Von Eye, Clifford C Clogg (comps.), Latent variable analysis: application for development research.

Maddala, G.S. (1995). “Specification test in limited dependent variable models”, En Maddala, G.S.,

Mascollel, A.., Whinston,M.D y Green,J.R. (1995). Microeconomic Theory, NY, Oxford University Press.

Mincer, J. (1970). “The distribution of labor incomes: a survey with special reference to the human capital approach”, Journal of economic literature, vól.8, pp.1-26.

Michael, R.T. (1973). “Education in nonmarket production”, Journal of political economy, vól.81, núm.2, pt.1, pp.306-27

Mitchel, R.C and R.T, Carson. (1989). Using surveys to value public goods: the contingent valuation method, Resources for the future, Washington.D.C.

Melino, A. (1982). “Testing for sample selection bias “, Review of economic studies, 49, pp.151-3.

Mora, J. J. (1997). “Aspectos microeconómicos del paisaje”, Boletín socioeconómico, núm.30, pp.81-97.

Mora, J. J. (2002). “Caracteristicas Socioeconomicas y Consistencia en la toma de decisiones”, Journal of Management and Economics for Iberoamerica, núm.32 pag 57-69.

Mora, J.J. y Ulloa, P. (2011). “Calidad del Empleo en las Principales Ciudades Colombianas y Endogeneidad de la Educaciòn”, Revista de Economía Institucional, Vol. 13(25), pages 163-177.

Mora, J.J. , Bernat, L.F y Zuluaga, B. (2012). "La elasticidad ingreso del consumo cultural en Cali," Revista de Economía Institucional, Vol. 14(27), pages 165-192.

Mora, J.J. (2013). "Gender differences between remittances and labor participation in developing countries: A cross-section analysis of Colombia in year 2008," Applied Econometrics and International Development, vol. 13(1), pages 99-112.

Muellbauer, J. (1981). “Linear aggregation in neoclassical labor supply”, Review of economics studies, 41, pp.28-36

Nawata, K. (1993a). “A note on the estimation of models with sample selection biases”, Economic letters, 42, pp.15-24.

Nawata, K. (1993b). “Estimation of sample selection biases models”, Manuscript, Dept of economics, University of Western, Australia.

Nelson, F. “Censored regression models with unobserved stochastic censoring thresholds”. Econometric 6, pp.309-327.

Nelson, F y L.. Olson. “Specification and estimation of a simultaneous equation model with limited dependent variables”. International economic review, Oct , pp.695-710.

Newey, W.K. (1985). “Maximum likelihood specification testing and conditional moment test”, Econometrica, vol. 53, pp.1047-73.

Newey, W.K. (1987). “Specification test for distributional assumptions in the Tobit model”, Journal of econometrics, vol. 34, pp.125-45.

Olsen, R. (1982). “Distributional test for selectivity bias and more robust likelihood estimator”, International economic review, 23, pp. 233-40.

Orme, R.J. (1990) "The small-sample performance of the information matrix test", Journal of econometrics, 46, pp.309-31.

Orme, R.J. (1992). "Efficient score test for heterocedasticity in microeconometrics", Econometrics review, 11, pp. 235-52.

Owen, J.D. (1970). "The demand for leisure", Journal of political economy, vol. 79, pp.56-76.

Pagan, A.R and F, Vella. (1989). "Diagnostic test for models based on unit record data: a survey", Journal of applied econometrics, 4, pp. 175-94.

Pagan, A. R and Park, Y. (1993). "Testing for heteroskedasticity in G.S maddala". C.R Rao an H.D vinod (eds), Handook of Statistic Vol. 11 Amsterdam : North-Holland, pp.489-518.

Palmsquist, R.B. (1984). "Estimating the demand for the characteristic of housing", The review of economics and statistics, vól.66, núm.3, pp.394-404.

Parsons, G. R and M.J, Kealy. (1992). "Randomly drawn opportunity sets in a random utility model of lake recreation", Land economics, vól 68, núm 1, pp 93-106.

Pearce, I.F. (1964). A contribution to demand analysis, Oxford University Press.

Pearce, D.W and R.K, Turner. (1990). Economics of natural resources and the environment. Harverster, Londres.

Pencavel, J.H. (1979). "Market work decisions and unemployment of husbands and wives in the seatle and Denver income maintenance experiments". Mimeographed. April.

Pencavel, J. (1998). "Assortative mating by schooling and the work behavior of wives and husbands", The American economic review, vol.88, num.2, pp.326- 329.

Phlips, L. (1978). "The demand for leisure and money", Econometrica, vol. 46, pp.1025-43.

Phillips, P.C.B., Srinvasan, T.N (comps.), Advances in econometrics and qualitative variables.

Phillips, P.C.B., Srinvasan, T.N (1994). Econometrics methods and applications, Volume II, Edward Edgar Publishing.

Phillips, P.C.B., Srinvasan, T.N (1983) Dependent and qualitative variables in econometrics, Cambridge, Cambridge University Press.

Phillips, P.C.B., Srinvasan, T.N (1995). Advances in econometrics and qualitative variables(comps,), Basil Blackwell.

Phillips, P.C.B., Srinvasan, T. And F.D, Nelson. (1975). “Specification errors in limited dependent variable models” NBER Working paper series, No 96.

Phillips, P.C.B., Srinvasan, T. And L.F, Lee. (1985). “The common structure of test for selectivity bias, serial correlation, heterocedasticity and non-normality in the Tobit model”, International economic review, 26,pp.1-20

Plott, Ch. R. (1996). “Rational individual behavior in markets and the social choice process: the discovered preferences hypothesis” en Kenneth J. Arrow., Enrico Colombato., Mark Perlman and Christian Schmidt (comps.), The rational foundations of economic behaviour- Proceedings of the IEA conference in held Turin, Italy, St Martin´s press.inc.Usa.

Pollak, R.A and T.J, Wales. (1969). “Estimation of the linear expenditure system”, Econometrica, 37, pp. 611-28.

Pollak, R.A and T.J, Wales. (1969). “Estimation of the linear expenditure system”, Econometrica, 37, pp. 611-28.

Pollack, R. A. (1978). “Endogenous tastes in demand and welfare analysis”, American Economic Review, vol.68. pp.374-79.

Pollack and Watcher, M. (1975). “The relevance of the household production function and its implications for the allocation of time”, Journal of political economy, vól.88, núm.2, pp.255-277, April.

Pollack and T.J Wales. (1979). “Welfare comparisons and equivalence scales”, American Economic Revi, vol.69, pp.216-21.

Powell, J.L. (1986). “Symmetrically trimmed least squares estimation for Tobit models”, Econometrica, vol. 54, pp.1435-60.

Pratt,J.(1964). “Risk aversion in the small and in the large”, Econometrica, 32,pp.122-36.

Preginon, D. (1980). “ Goodness of link test for generalized linear models”, Applied statistics, 29, pp.15-24.

Pudney, S.E. (1981). “An empirical method of approximating the separable structure of consumer preferences”, Review of economics studies, 48, pp.561-77.

Quandt, R.E. (1956). “A probabilistic theory of consumer behavior”, Quarterly journal of economics, 70, pp.507-536.

Rao, C.R and Mitra, S.K. (1971). “Generalid inverses of matrices and its applications, New York : Wiley.

Rao, C.R. (1947) “Large sample test of statistical hypothesis concerning several parameters with applications to problems of estimation”, Proceedings of the Cambridge philosophical society, 44, pp.50-7.

Rao, C.R (1973). Linear statistical inference and its applications, New York, John Wiley and Sons.

Rao, C.R And Mitra, S.K. (1971). Generalized inverses of matrices and its applications, New York, John Wiley and Sons.

Rosen, S. (1974). “Hedonic prices and implicit markets: product differentiation in pure competition”, Journal of political economic, núm 82, pp.34-55.

Ruud, P.A. (1984). “Test of specification in econometrics”, Econometric reviews, 3, pp.211-42.

Samuelson, P.A. (1956). “Social Indiference curves”, Quaterly journal of economics, vól. 70, pp.1-22.

Samuelson, P.A. (1947). “Fundations of Economic Analysis”, Cambridge, Mass.: Harvard University Press.

Schmid, P. And Strauss, R.P (1975). “ The prediction of occupation using multiple logit model”. International economic review; June, pp. 47-86.

Silberberg, E. (1985). “Nutrition and the demand for tastes”, Journal of political economy, vól.95,núm.5,pp.881-900.

Silberberg, E. (1990). The structure of economics: a mathematical analysis. Mc Graw Hill.

Silvey, S.D. (1959). “The lagrangian multiplier test”, Annals of mathematical statistics, núm 7.pp.389-407.

Skeels, C.L and F, Vella. (1993). The performance of conditional moment test in Tobit and Probit models. ANU, Camberra.

Slutzky, E. (1915). “ Sulia teoria del bilancio des consomatore”, Giornale degli Economisti, vol51, pp.1-26. English trans. In Readings in Price Theory; G.J. Stigler and K.E Boulding (eds.), Chicago University Press,1952.

Small, K.A and Rosen, H.S. (1981). “Applied welfare economics with discrete choice models”, Econometrica, vol. 49, pp.105-129.

Smith, R. (1989). “On the use of distributional misspecification checks in limited dependent variable models”, Economic journal, supplement, 99, pp.178-92.

Smith, R. And Blundell, R. (1986) “An exogeneity test for simultaneous equation Tobit model with application to labor supply”, Econometrica, vol. 54, pp. 679-85.

Smith, K.V and Y, Kauru. (1990). “Signals or noise? Explaining the variation in recreation,”. American journal of agricultural economics, 72, pp. 419-450.

Smith,V.K., Desvoges, W.H and M.P, Mcgivenes. (1983). “The opportunity cost of travel time in recreation demand models”, Land economics, 59, pp. 59-77.

Sono, M. (1962). “The effect of price changes on the demand and supply of separable goods” International economic review, vól.2, pp.239-71.

Spence, M. (1973). “Job market signaling”, Quarterly journal of economics, vol.87, pp.355-79.

Srinvasan, T.N (comps.), Advances in econometrics and qualitative variables, Basil Blackwell.

Srinvasan, T.N. (1994) Econometrics methods and applications, Volume II, Edward Edgar Publishing.

Srinvasan, T.N. (1983) Dependent and qualitative variables in econometrics, Cambridge, Cambridge University Press.

Stiglitz and F.G, Mathewson (comps.) New developments in the analysis of market structure, Cambridge, MIT Press.

Stiglitz, J.E. (1975). “The theory of screening, education and the distribution of income”, American economic review, vól.65, pp.283-300.

Stone, R. (1954). “Linear expenditure systems and demand analysis: an application to the pattern of British demand”, Economic journal, núm.64, pp.511-527, Sep.

Willig, R. (1976). “Consumer´s surplus without apology”, American economic review, vol.66, pp.589-597.

Tauchen, G. (1985) “Diagnostic testing and evaluation of maximum likelihood models”, Journal of econometrics, 30, pp.415-43.

Taylor, L. (1991). “Testing exclusion restricctions for misspecified Tobit model”, Economics letters, 37, pp.411-16.

Theil, H. (1965). “The information approach to demand analysis”, Econometrica, vól. 33, pp.67-87.

Thurstone, L.L. (1927). “A locus of comparative judgement”, Psychological review, 34, pp.273-286.

Tukey, J.W. (1949). “One degree of freedom for non-additivity”, Biometrics, 5, pp. 232-242.

Tversky, A .(1996). “Rational theory and constructive choice” en Kenneth J. Arrow., Enrico Colombato., Mark Perlman and Christian Schmidt (comps.), The rational foundations of economic behaviour-Proceedings of the IEA conference in held Turin, Italy, St Martin´s press.inc.Usa.

Tversky, A and Kahneman, D. (1991). “Loss aversion in riskless choice: a reference dependent model”, Quaterly journal of economics, vól.107, pp.1039-61.

Tversky, A., Slovic, P. and Kahneman, D. (1990). “The causes of preference reversal” American economic review, vol. 80, pp.204-17.

Tversky, A. (1972). “Elimination by aspects: a theory of choice”, Psychological review, 79, pp.281-299.

Tversky, A. (1969).”Intrasitivity of Preferences”, Psych. Rev. 76, pp.31-48

Uribe, J.I. (1998). “Duración del desempleo: un modelo de determinantes y su aplicación al área metropolitana de Cali”, Tesis doctoral no publicada, Facultad de Ciencias Económicas y Empresariales, Departamento de Economía Aplicada III, Universidad Complutense de Madrid,359 páginas.

Varian,H.R. (1992). Análisis microeconomico. Antoni Bosh (tercera edición).

Von Neumann,J., and O. Morgenstern.(1944). Theory of games and economic behavior, Princeton University Press, Princeton,New York.

Webb, W.B. (1985). “Sleep in industrialized settings in the Northern Hemisphere”, Psychological reports, 57, Oct, pp.591-98.

Wedderburn. (1976). “Quasi-likelihood functions, generalized linear models, and the gauss newton method”, Biometrika, vol.61, num. 3, pp. 439.

Westin, R.B and Guillen, D.W. (1978).” Parking location and transit demand: A case study of endogenous attributes in disaggregate mode choice functions”. Journal of econometric, 8, pp. 75-101.

White, H. (1982). “Maximum likelihood estimation of misspecified models”, Econometrica, vól 50, pp.1-25.

Willing,R.D. (1976). “Consumer´s surplus without apology”, American economic review, vól.66, num.4, pp.589-597.

Willis, K.G and G.D, Garrod. (1991). “An individual travel cost method of evaluating forest recreation”, American journal of agricultural economics, vól 42, pp. 33-42.

Willig, R.D. (1976). “Consumer surplus without apology”, American economic review, vól.66, núm.5, pp. 89-591, Sep

Wu, D.M. (1973). “Alternative test of independence between stochastic regressors and disturbances”, Econométrica, vol.41, pp.134-39.