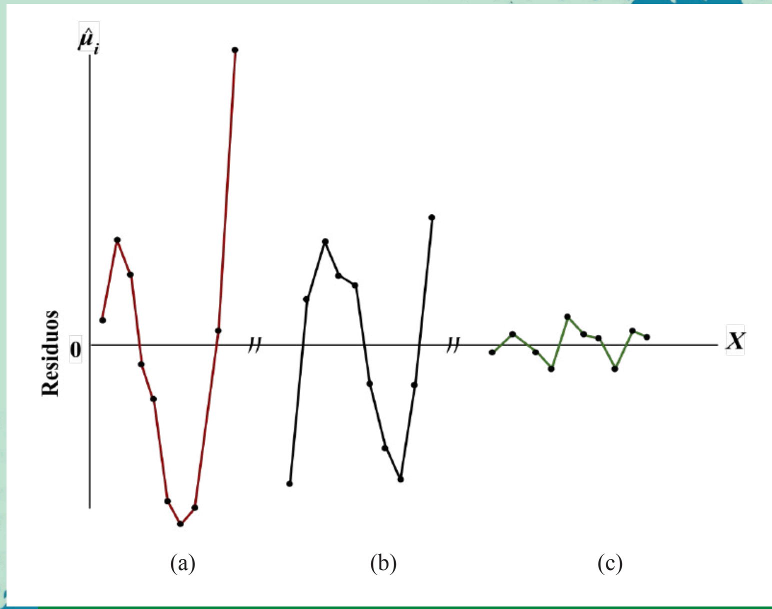
UNIVERSIDAD NACIONAL AGRARIA
LA MOLINA

TÓPICOS ESPECIALES DE ECONOMETRÍA APLICADA



Jorge Alarcón Novoa

TÓPICOS ESPECIALES DE ECONOMETRÍA APLICADA

Jorge Alarcón Novoa

UNIVERSIDAD NACIONAL AGRARIA
LA MOLINA

JORGE ALARCÓN NOVOA

Tópicos especiales de econometría aplicada

Lima: 2022; 148 p.

ISBN: N° 978-612-5086-09-9

Hecho el Depósito Legal en la Biblioteca Nacional del Perú N° 2022-13046

Primera edición: diciembre de 2022 - Tiraje: 500 ejemplares

Impreso en Perú - Printed in Perú

Diseño y diagramación:
Daniella Luna Barrios

Se terminó de imprimir en diciembre de 2022 en:
QyP Impresores s.r.l.
Av. Ignacio Merino 1546 - Lince
E-mail: qypimpresores2005@yahoo.com

CONTENIDO

ABREVIATURAS

AIC	Criterio Akaike para la comparación de la bondad de ajuste de modelos
BIC	Criterio Schward para la comparación de la bondad de ajuste de modelos
COV	Covarianza
DW	Estadístico Durbin-Watson
ELIO	Estimador lineal insesgado y óptimo
FRP	Función de regresión poblacional
FRM	Función de regresión muestral
"*F*"	Estadístico de la distribución *F*; Fc es el valor calculado y Ft es el valor tabular o crítico
FIL	Función intrínsecamente lineal
Ho	Hipótesis nula
Ha	Hipótesis alterna
Ln	Logaritmo natural
MCI	Método de estimación por mínimos cuadrados indirectos
MCG	Método de estimación por mínimos cuadrados generalizados
MCO	Método de estimación por mínimos cuadrados ordinarios
MC2E	Método de estimación por mínimos cuadrados en dos etapas
MDA	Modelo con datos agrupados
MDG	Modelo panel de efectos fijos intra-grupos
MEAL	Modelo panel de efectos aleatorios
MEF	Modelo panel de efectos fijos inter-grupos (con variables ficticias)
MFE	Modelo en su forma estructural (ecuaciones múltiples)
MFR	Modelo en su forma reducida (ecuaciones múltiples)
ML	Multiplicador de Lagrange
MNR	Modelo no restringido
MR	Modelo restringido
MRL	Modelo general de regresión lineal
MV	Método de estimación de máxima verosimilitud
PRV	Prueba de razón de verosimilitud
pLim	Probabilidad de consistencia de un estimador, en relación al parámetro.
R^2	Coeficiente de determinación
RESET	Prueba de error de especificación propuesto por Ramsey
t	Estadístico de la distribución *t* de Student

SCE	Suma de cuadrados explicada
SCR	Suma de cuadrados residual
SCT	Suma de cuadrados total
VI	Variable instrumental
Wald	Estadístico que sigue una distribución Ji-cuadrada y se usa para probar hipótesis con modelos restringidos.
χ^2	Estadístico de la distribución Ji-cuadrada

PRÓLOGO

Este libro está pensado básicamente para ayudar a los estudiantes en sus cursos y en apoyo a sus investigaciones; también puede ayudar a profesionales, de diversos sectores, en cuanto a la aplicación de técnicas de modelística en sus investigaciones. No hay nada inventado, la idea es rescatar aspectos prácticos en Econometría Aplicada, pero que son poco usados, haciendo una presentación sencilla y didáctica.

En cada capítulo el libro contiene ejemplos de uso de los temas tratados; también ejercicios y sugerencias de solución. Los ejercicios están hechos con base en información tomada de la propia experiencia de investigación del autor, así como ejercicios adaptados de otros libros, a menudo con adaptación a circunstancias de nuestra propia universidad, ciudad y país.

Dado que éste no es un libro básico, contiene un primer capítulo conceptual que puede ayudar con la revisión de temas que son importantes para un mejor entendimiento de los instrumentos tratados. También hay recuadros que son “apoyos” importantes en la explicación de aspectos conceptuales específicos.

Además del primer capítulo, conceptual, el segundo capítulo se dedica a la presentación de diferentes formas funcionales comunes en economía y que se basan en la transformación que ofrece el proceso de linealización de funciones originales no lineales en los parámetros. El capítulo 3 estudia el tratamiento de problemas relacionados con inferencia estadística en el contexto de modelos restringidos. El capítulo 4 se refiere a la especificación y diagnóstico de un modelo lineal general; es importante, pues los errores de especificación usualmente conllevan otro tipo de violaciones de supuestos y es un tema que usualmente es soslayado. Los errores de especificación aparecen frecuentemente cuando se omiten variables importantes o también cuando se usan variables irrelevantes, así como por el uso de funciones matemáticas distintas a las que representan los datos o la misma teoría.

El capítulo 5 está dedicado al manejo de modelos estáticos con datos panel, en que se combinan datos de series de tiempo con corte transversal; estos modelos son poco usuales en la práctica por el costo relativamente alto en la obtención de la data y por la complejidad en su especificación, tratamiento y estimación. El capítulo 6, que es el último, trata el caso de modelos de ecuaciones simultáneas, los contextos de su uso, los requerimientos para su resolución y las modalidades de estimación. En este capítulo final se analizan también los problemas que surgen en la estimación de los parámetros y las alternativas de solución existentes.

Finalmente es importante agradecer a los estudiantes de los cursos de Econometría que con su participación en las clases contribuyeron a la revisión y realización, sobretodo, de los ejercicios de cada capítulo. También gracias a dos lectores anónimos que se tomaron el tiempo de leer el manuscrito y contribuir, con sus comentarios, a mejorarlo. Indudablemente los errores que subsistan son de entera responsabilidad del autor.

CAPÍTULO I. CONCEPTUALIZACIÓN GENERAL

La econometría se emplea en todas las áreas de la economía, para probar teorías, permitir información para políticas públicas, predecir escenarios alternativos; incluso es utilizada en otras disciplinas de las ciencias sociales para efectos de predicción y también de política, entre otros.

Los objetivos del análisis econométrico son estimar los parámetros β_j de un modelo y probar hipótesis acerca de ellos; los valores y signos de los parámetros determinan la validez de una teoría económica y los efectos de políticas.

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \beta_3 X_{3i} + \cdots + \beta_k X_{ki} + \mu_i$$

Donde los parámetros β_j son estimados mediante coeficientes representados por: $\hat{\beta}_j$

Al respecto, algunas definiciones importantes son las siguientes:

Estimación insesgada
Un estimador $\hat{\beta}$ es insesgado cuando su valor esperado es igual al verdadero parámetro[2]. Es decir $E(\hat{\beta}) = \beta$

Una propiedad deseable es que un parámetro de regresión estimado tenga un estimador cuya distribución tenga al parámetro como su Valor Medio. Ésta es una definición absoluta.

Eficiencia de un estimador
Una segunda propiedad deseable de un estimador es que tuviera una dispersión muy pequeña alrededor del verdadero parámetro.
Entonces, si $V(\hat{\beta}_1) < V(\hat{\beta}_2)$ ....

Decimos que $\hat{\beta}$ es un estimador insesgado y eficiente si para un tamaño muestral dado, la varianza de $\hat{\beta}$ es menor que la varianza de cualquier otro estimador insesgado que hubiera. Ésta es una definición relativa.

---

[2] Un *Estimador* es un número, mientras que una *Estimación* es considerada una regla o procedimiento.

Error cuadrático medio

Algunas veces es necesario "equilibrar" el deseo entre insesgamiento y eficiencia, dependiendo del propósito: por ejemplo, cuando la meta de un modelo es maximizar la precisión de las predicciones, un estimador con una varianza muy baja y un pequeño sesgo puede ser más deseable que un estimador insesgado con una varianza alta.

Eventualmente, un criterio útil es minimizar el Error Cuadrático Medio (*ECM*):

$$ECM(\hat{\beta}) = E(\hat{\beta} - \beta)^{\mathbf{2}}$$

Se puede demostrar que la ecuación puede también escribirse como[3]:

$$ECM(\hat{\beta}) = [Tendencia(\hat{\beta})]^{\mathbf{2}} + V(\hat{\beta})$$

Cuando el estimador es insesgado, el *ECM* del estimador es igual a su varianza:

$$ECM(\hat{\beta}) = V(\hat{\beta})$$

Consistencia de un estimador

Tiene que ver con la propiedad asintótica de un estimador: comportamiento con el crecimiento del tamaño de muestra. Es un concepto que incorpora un criterio probabilístico:

Plim de $\hat{\beta} = \beta$ significa convergencia en probabilidad; es decir que conforme el número de observaciones ("*n*") crece hasta el infinito, la probabilidad de que $|\hat{\beta} - \beta|$ sea menor que cualquier número positivo arbitrariamente pequeño**,** se aproximará a uno (la probabilidad).

En rigor, $\boldsymbol{\hat{\beta}}$ es consistente si converge con $\boldsymbol{\beta}$ en el límite de probabilidad; es decir si para cualquier $\gamma > 0$, *Lim* Prob $(|\hat{\beta} = \beta| < \gamma) = 1$

Una regla práctica es que los economistas se han enfocado más en estimadores consistentes en lugar de insesgados. En particular es de preocupación el uso del concepto de Consistencia del error cuadrático medio: "que el ECM del estimador se aproxime a CERO conforme $\boldsymbol{n}\uparrow\infty$", implica que el estimador es asintóticamente insesgado y que la varianza se aproxima a cero conforme el número de observaciones ("*n*") crece.

Es una ventaja que, en la práctica, los estimadores consistentes a menudo tienden a tener una varianza (VAR) que tiende a CERO conforme $\boldsymbol{n}\uparrow\infty$".

[3] La demostración puede revisarse en la nota a pie de página 32 de Pyndick y Rubinfeld (1991).

Inferencia estadística

Conjunto de métodos y técnicas que permiten inducir, a partir de información empírica muestral, el comportamiento de una determinada población con un riesgo de error medible en términos de probabilidad. La estadística paramétrica comprende: (i) la estimación de parámetros y (ii) el contraste de hipótesis. Ambos métodos se basan en el conocimiento teórico de la distribución de probabilidad del estadístico muestral que se utiliza como estimador de un parámetro. La estadística inferencial es un instrumento medular en la estimación y uso de modelos econométricos.

En específico una prueba de hipótesis o contraste tiene como objetivo comprobar si determinado supuesto referido a un parámetro poblacional, o a parámetros análogos de dos o más poblaciones, es compatible con la evidencia empírica contenida en la muestra. El uso de estadísticos más frecuentes son el estadístico "*t*" para contrastes simples y "*F*" para contrastes más complejos o para restricciones de exclusión múltiple. Es usual también requerir de un margen de error en cada prueba, lo que en la práctica se hace mediante el uso del denominado "p-valor"

Mínimos cuadrados ordinarios (MCO)

Mínimos cuadrados ordinarios es el método más común para la estimación de los parámetros de los modelos econométricos. Bajo determinados supuestos, los estimadores MCO resultan ser estimadores lineales insesgados y óptimos (ELIO), incluyendo también un estimador de la $V(\mu_i)$. La violación de supuestos conlleva, entre otros, la existencia de heteroscedasticidad, autocorrelación, multicolinealidad, etc. En alguna literatura especializada, a un estimador ELIO también se le denomina MELI, que significa "Mejor Estimador Lineal Insesgado".

Estructuras de datos

Los tipos comunes de estructuras de datos pueden ser básicamente tres: datos de corte transversal, que se refieren a aquellos que describen las actividades de individuos, empresas u otras unidades en un punto dado en el tiempo o en períodos relativamente cortos; en cambio, los datos de series de tiempo describen el movimiento de las variables a lo largo del tiempo. Los datos que combinan los dos antes descritos se denominan datos panel o datos combinados, a menudo se usan para estudiar el comportamiento de un grupo de individuos o empresas a lo largo del tiempo.

Coeficientes de determinación[4]

R^2 es el coeficiente de determinación y establece la proporción de la variación muestral de la variable dependiente que es explicada por las variables independientes

[4] En este documento se utiliza, en representación del coeficiente de determinación, el término R^2 o también R2.

y se usa como medida de la bondad de ajuste. La R^2 ajustada (R_a^2) ha mostrado ser un mejor indicador de la bondad de ajuste del modelo, pues mientras que R^2 nunca puede disminuir cuando se agrega variables independientes a la regresión, el R_a^2 si penaliza por la cantidad de variables regresoras adicionales y puede disminuir cuando se incrementa el número de regresoras. Existen otras medidas de bondad de ajuste alternativas para la comparación entre modelos con la misma variable dependiente.

Variables ficticias
Se usan variables "ficticias" o binarias para distinguir entre grupos o categorías de población involucradas en el modelo. El coeficiente estimado de la variable ficticia estima la diferencia (*ceteris paribus*) entre dos o más grupos. Cuando las variables ficticias participan en forma aditiva en el modelo, los coeficientes de las variables ficticias expresan diferencias en interceptos de cada grupo o categoría; por el contrario, cuando las variables ficticias participan en forma aditiva, sus coeficientes representan diferencias en las pendientes.

Funciones de regresión poblacional y muestral
Una función de regresión poblacional (FRP) es la expresión de la media condicional del valor de una variable dependiente en función de una o más variables independientes: $E(Y/Xi)= f(X)$; también se le conoce como función de esperanza condicional y representa a un grupo definido y completo de unidades en los que se enfoca un análisis econométrico.

La función de regresión muestral (FRM) expresa el valor ajustado de la FRP a partir de una muestra observada de individuos que son parte de la población en cuestión. Se representa por: $\hat{Y} = f(X_s)$; donde el valor $\hat{Y}$ implica una estimación de los parámetros poblacionales.

Teorema de Gauss Markov
Formulado por Carl F. Gauss y Andréi Markov, el teorema se refiere a un conjunto de cinco supuestos que debe cumplir un estimador obtenido mediante el método de Mínimos Cuadrados Ordinarios, para que se considere un estimador lineal insesgado óptimo (ELIO); es decir, el estimador con menor varianza que cualquier otro estimador insesgado obtenido con un método de estimación distinto. El Teorema de Gauss-Markov involucra: linealidad en los parámetros, media nula y exogeneidad estricta de las variables independientes, homocedasticidad, no multicolinealidad ni autocorrelación.

Coeficientes de regresión parcial

En un modelo de regresión múltiple, como el siguiente:

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \beta_3 X_{3i} + \cdots + \beta_k X_{ki} + \mu_i$$

Los coeficientes de regresión de las pendientes, es decir $\beta_1, \beta_2, \beta_3, \dots, \beta_k$ se denominan coeficientes de regresión parcial, pues a diferencia de la pendiente de un modelo bivariado, miden el cambio en el valor promedio de Y, por cada unidad de cambio en c/u de las variables regresoras, cuando todas las demás variables se consideran constantes; en otras palabras el solo efecto de una determinada variable regresora, manteniendo bajo control el efecto de las otras variables independientes.

CAPÍTULO II. FORMAS FUNCIONALES ALTERNATIVAS

2.1. Linealidad del Modelo General

Expresado el modelo de regresión lineal de la siguiente forma:

$$Y_i = f(X_{1i}, X_{2i}, .., \beta_0, \beta_1, \beta_2 \dots) + \mu_i,$$

los supuestos asociados con la especificación de las variables independientes o regresoras y el término de error, son elementos fundamentales para que la técnica de Mínimos Cuadrados Ordinarios (MCO) cumpla con la provisión de estimadores lineales, insesgados y óptimos (ELIO). Según aval del Teorema de Gauss Markov, las propiedades estadísticas de los estimadores MCO se basan en los supuestos del Modelo clásico de Regresión Lineal (MRL), dentro de los cuales destaca la relación lineal en los parámetros.

La teoría económica nos presenta varias formas de relación entre variables, pero no nos indica la forma funcional de las mismas. El supuesto fundamental para la aplicación del método de MCO es la linealidad en los parámetros. En general, el reconocimiento práctico de la linealidad de una función[5] en los parámetros (o en las variables) es mediante el reconocimiento de las siguientes condiciones:

(a) El parámetro (o variable independiente) debe tener exponente uno.
(b) El parámetro (o variable independiente) no debe ser exponente.
(c) El parámetro (o variable independiente) no debe ir interactuando con ningún otro parámetro (o variable).

Una función será originalmente lineal en los parámetros cuando puede aplicarse directamente el método de MCO para obtener estimadores ELIO; es decir no se requiere transformación alguna. Comúnmente la transformación de una función no lineal (FNL) a otra intrínsecamente lineal se logra utilizando logaritmos; ocasionalmente algunas funciones transformadas se obtienen utilizando manipulación matemática más compleja. Ejemplos de funciones originalmente lineales, en la Tabla 2.1, son las funciones (1) y (2).

---

[5] Teóricamente, la evaluación de la linealidad en parámetros (o también variables) requiere la utilización de técnicas de derivación matemática que escapan al propósito de este texto.

Una función es intrínsecamente lineal cuando es originalmente no-lineal en los parámetros, pero es posible "linealizarla" o transformarla para permitir el uso del método de MCO con el fin de obtener los mejores estimadores de los parámetros. La técnica, denominada de "linealización", nos permite finalmente obtener múltiples formas funcionales, consideradas como FRL en las que es posible usar el método de MCO con el fin de obtener los mejores estimadores lineales e insesgados (MELI o ELIO). La función (3) de la Tabla 2.1, es un buen ejemplo de este tipo de funciones.

Existen también funciones que no son originalmente lineales pero que tampoco es posible linealizarlas o transformarlas con el propósito de utilizar MCO apropiadamente; a tales funciones se les conoce como Funciones Intrínsecamente No Lineales (FINL). En este caso los métodos de estimación mayoritariamente usados incluyen procesos iterativos y cálculos numéricos. A diferencia de los casos en que se utiliza MCO, los parámetros no son estimados en forma explícita[6]

La teoría, o la dispersión de los puntos muestrales, frecuentemente sugieren el uso de funciones tanto lineales originales como intrínsecamente lineales (FIL). La teoría sugiere, incluso, la especificación de modelos intrínsecamente no lineales, como es el caso de la *función de producción con elasticidad constante de sustitución*, que es un caso especial de una función de producción Cobb-Douglas.

Tabla 2.1. Funciones y relaciones más usadas en Economía

<table>
<tr><th>Función</th><th>Naturaleza</th><th>Aplicabilidad del método de MCO</th></tr>
<tr><td>1. $Y = \beta_0 + \beta_1 X_1 + \mu_i$</td><td>La función es lineal en las variables y lineal en los parámetros</td><td>(Aplica MCO)</td></tr>
<tr><td>2. $Y = \beta_0 + \beta_1 X_1 + \beta_2 X_1^2 + \mu_i$</td><td>La función es no lineal en las variables y lineal en los parámetros</td><td rowspan="2">(MCO es aplicable)</td></tr>
<tr><td>3. $Y = \beta_0 {X_1}^{\beta_1} \mu_i$</td><td>Función no lineal en las variables y NO lineal en los parámetros (Intrínsecamente lineal)</td></tr>
<tr><td>4. $Y = \beta_0 {X_1}^{\beta_1} + \mu_i$</td><td>Función no lineal en las variables y no lineal en los parámetros (NO es intrínsecamente lineal)</td><td>(No aplica MCO)</td></tr>
</table>

[6] En el capítulo 14 del libro de Gujatati & Porter (2010) se presentan métodos comunes utilizados en la estimación de modelos intrínsecamente no lineales.

5. $Y = \beta_0 + \beta_1^2 X_1 + \mu_i$	Función no lineal en los parámetros	MCO no es aplicable

2.2. Funciones más usadas en Economía

La matemática nos brinda la oportunidad de contar con un gran número de funciones que podrían ser usadas en el contexto de las ciencias sociales y otras disciplinas. De las más usadas en economía, tenemos las siguientes:

2.2.1. Función lineal

$$Y = \beta_0 + \beta_1 X_1 + \mu_i \qquad (2.1)$$

Propiedades

(i) Modelo que cumple con la propiedad de linealidad en parámetros y que, por tanto, puede ser estimado por el método de MCO.

(ii) Es una función de pendiente constante: $\frac{dY}{dX} = \beta_1$

(iii) Es una función de elasticidad variable, pues depende de los valores de *X* y *Y*

$$\varepsilon = \frac{dY}{dX}\left(\frac{X}{Y}\right) = \beta_1\left(\frac{X}{Y}\right)$$

Ejemplo. Un bien con función de demanda lineal.

$$\hat{Q}_i = 8 - 0.5(P_i)$$

Q = cantidad demandada de computadores, en cientos de unidades
P = precio de computadores, en miles de soles
Entonces $\hat{\beta}_1$=0.5 indica que, por ejemplo, si se incrementara el precio de los computadores en 1 mil soles, entonces disminuirá la cantidad demanda de computadores en 100.

$\hat{\beta}_0$= 8 indica que, si los computadores fueran gratis, la cantidad demandada sería de 800 unidades (la unidad de medida del intercepto es la misma que la de la variable dependiente). Ver Figura 2.1, en página posterior, que muestra funciones en escala lineal y semilogarítmica.

2.2.2. Función doble –logarítmica

Partiendo de una función exponencial que expresa una relación originalmente no lineal en los parámetros:

$$Y_i = e^{\beta_0} X_i^{\beta_1} e^{\mu} \qquad (2.2)$$

Y que puede ser linealizada o transformada en:

$$LnY = \beta_0 + \beta_1 LnX_i + \mu_i \qquad (2.3)$$

La ecuación original es no lineal en las variables y en los parámetros. La ecuación transformada es lineal en los parámetros, por lo que es aplicable el método de MCO.

Propiedades

(i) Si los supuestos del modelo de regresión se cumplen, los parámetros de (2.3) pueden ser estimados mediante MCO y los estimadores serán lineales, insesgados y óptimos (ELIO).

(ii) El modelo transformado (2.3) tiene pendiente variable: depende de los valores de *Y* y de *X*.

$$\frac{dLnY}{dLnX} = \beta_1 \qquad \mapsto \qquad \beta_1 = \frac{(\frac{dLnY}{dX})}{(\frac{dLnX}{dX})}$$

$$\beta_1 = \frac{\frac{1}{Y}(\frac{dY}{dX})}{\frac{1}{X}(\frac{dX}{dX})} \qquad (2.4)$$

$$\frac{1}{Y}\left(\frac{dY}{dX}\right) = \beta_1\left(\frac{1}{X}\right)$$

$\mapsto \ \frac{dY}{dX} = \beta_1 \left(\frac{Y}{X}\right)$; la pendiente es variable, depende de los valores de *Y* y *X*.

(iii) El modelo linealizado es un modelo de elasticidad constante.

De la fórmula (2.4): $\boldsymbol{\varepsilon} = \frac{dY}{dX}\left(\frac{X}{Y}\right) = \hat{\beta}_1$; representa una elasticidad constante.

Una buena pista para la verificación es graficar la función doble-LOG en ejes logarítmicos y evaluar la linealidad gráfica de la función transformada.

Ejemplo 2.1. Considere el modelo:

$LnQ_{\boldsymbol{i}} = \beta_0 + \beta_1 LnP_i + \mu_i$; cuyos estimadores MCO resultan en:

$\widehat{LnQ} = 3.96 - 0.25 LnP_i$

Donde:
Q = cantidad demandada de carne de cerdo (en Kg)
P = precio del bien (en S/. x Kg).

Interpretación:
$\beta_1 = -0.25$ (elasticidad) significa que si aumentara el precio del bien en 100%, la cantidad demandada de carne de cerdo disminuye en 25%. O si duplica el precio de la carne, entonces la cantidad demandada de carne de cerdo disminuye en 25%.

β_0= 3.96 (antilog= 52.5) sería la cantidad demanda si el bien sería "donado" (gratuito).

2.2.3. Función semi-logarítmica Log-Lin

Parte de un modelo de regresión original como el siguiente:

$$\boldsymbol{Y_i} = Y_0 \boldsymbol{e^{\beta_1 X} e^{\mu}} \qquad (2.5)$$

Que puede ser expresado como:

$$Ln\boldsymbol{Y_i} = Ln\boldsymbol{Y_0} + \beta_1 X_i + \mu_i$$
$$\text{o} \quad Ln\boldsymbol{Y_i} = \beta_0 + \beta_1 X_i + \mu_i \qquad (2.6)$$

La ecuación (2.6) representa una ecuación lineal en los parámetros y lineal en la variable independiente, logarítmica en la variable dependiente; el método MCO es aplicable.

Propiedades

(i) Si los supuestos del modelo de regresión se cumplen, los parámetros pueden ser estimados mediante MCO y los estimadores serán ELIO.

(ii) Función de pendiente variable:

$\beta_1 = \frac{d(LnY)}{dX} = \frac{1}{Y}\left(\frac{dY}{dX}\right)$; la pendiente depende del valor de Y.

$\left(\frac{dY}{dX}\right) = \beta_1(Y)$

(iii) β_1 es un parámetro importante, representa una semi-elasticidad de Y respecto a X : el estimador mide el cambio proporcional relativo (%) en $\boldsymbol{Y}$ para un cambio unitario absoluto en el valor de la variable regresora X:

$$\beta_1 = \frac{(\text{cambio relativo en } Y)}{(\text{cambio absoluto en } X)}$$

Para la comprobación, de (2.6):

$$\beta_1 = \frac{d(LnY)}{dX} = \frac{1}{Y}\left(\frac{dY}{dX}\right) = \frac{dY/Y}{dX}$$

Cuando la variable independiente cambia en una unidad, la variable dependiente cambia aproximadamente en $\%\Delta Y$: $100*(\hat{\beta}_i)$.

Ejemplo 2.2. Considere la función:

$$Ln(P) = \beta_0 + \beta_1 Ln(nox) + \beta_2 R + \mu_i$$

Donde:
P= precio promedio de viviendas en área metropolitana de la ciudad de Ica- Perú
Nox= cantidad de óxido de nitrógeno en el aire del entorno
R= número de habitaciones promedio.

La relación es doble-logarítmica entre *P* y *Nox,* pero es semi-logarítmica entre *P* y *R*. El coeficiente $\hat{\beta}_1$ representa la elasticidad precio respecto a la contaminación, mientras que el coeficiente $\hat{\beta}_2$ *expresa una semi-elasticidad.* Así, si mediante MCO se obtiene:

$$\widehat{Ln(P)} = 9.23 + 0.718.Ln(nox) + 0.306.R \qquad n=506; \; R^2 = 0.514$$

Con todos los estimadores significativos, cuando "*nox*" aumenta 10%, entonces el precio disminuye en 7.2%, manteniendo constante el número de habitaciones promedio. Asimismo, Cuando el N° de habitaciones incrementa en UNO, el precio de las viviendas aumenta aproximadamente en $100*\hat{\beta}_2$. Es decir, el precio incrementaría aproximadamente en 100*(0.306) = 30.6%

Ejemplo 2.3. Se puede usar un modelo semi-logarítmico (Log-lineal) para la estimación del crecimiento temporal de alguna variable económica, cuya propiedad es que reproduce la tasa de crecimiento de la variable dependiente *Y* ante un cambio a lo largo del tiempo (*t*). En este caso *X* es una variable de tendencia ($\boldsymbol{t}$), que aumenta de unidad en unidad (es decir anualmente, mensualmente, etc.), permitiendo obtener la tasa de crecimiento de *Y* (*Y* podría ser el PBI, por ejemplo).

Si se parte de la fórmula:

$PBI_t = Y_0(1 + r)^t \mu_i$

$Y_t = Y_0(1 + r)^t \mu_i$

Tomando logaritmos se obtiene:

$Ln\boldsymbol{Y_i} = LnY_0 + tLn(1 + r) + \mu_i$, es decir resulta en una ecuación del tipo:

$Ln\boldsymbol{Y_i} = \beta_0 + tLn(1 + r) + \mu_i$

$\widehat{Ln\boldsymbol{Y_t}} = \hat{\beta}_0 + \hat{\beta}_1 X_t$

Donde $X = t$ (tendencia temporal- años) y $\hat{\beta}_1$ sería $Ln(1 + r)$

Suponga que los resultados de la estimación MCO son:

$\widehat{LnPBI_t} = 12.007 + 0.025\ \boldsymbol{t}$

Interpretación:
$\hat{\beta}_1 = 0.025$ implica que: $(1 + r) = e^{0.025} = EXP\ (0.025)$

$r = EXP\ (0.025) - 1 = 0.0253$

"***r***" es la tasa de crecimiento del PBI: en el período muestral, el *PBI* ha crecido a una tasa del 2.5% anual. El intercepto, eventualmente no tendría demasiada importancia económica.

Ajuste de la función Log-Lin

En el caso de esta función, existe una propuesta de "ajuste" de la semi-elasticidad que permite una mejor precisión de la elasticidad (Wooldridge, 2009, pp.). Consiste en los siguiente**:**

%ΔY: 100*[exp$(\hat{\beta}_i)$-1]

Entonces, cuando la variable independiente cambia en una unidad (%ΔX=1), la variable dependiente (Y) cambia en: %ΔY: 100*[exp$(\hat{\beta}_i)$–1]

Así, en el ejemplo previo, el valor "ajustado" indica que si el N° de habitaciones incrementa en uno, el precio de las viviendas aumentará en 100*[exp$(\hat{\beta}_i)$-1]=100*[(exp(0.306)]=35.8%, que es un valor mayor al encontrado previamente (de 30.6%).

Es importante precisar, sin embargo, que tal ajuste no es crucial con cambios porcentuales pequeños (digamos menores a 10%)- es decir cuando $\hat{\beta}_i$< 0.10.

2.2.4. Función semi-logarítmica (Lin-Log)

Parte del modelo: $e^{y} = e^{\beta_0} X^{\beta_1} \mu_t$

$$\boldsymbol{Y_t} = \beta_0 + \beta_1 LnX_t + \mu_t \qquad (2.7)$$

El modelo transformado es lineal en los parámetros y logarítmico en la variable X_i. El método de estimación MCO aplica del siguiente modo:

$$\frac{dY}{dX} = \beta_1 \left(\frac{d(LnX)}{dX}\right) = \beta_1 \left(\frac{1}{X}\right)\left(\frac{dX}{dX}\right) \longmapsto \beta_1 = \frac{dY}{dX/X}$$

Propiedades

(i) Función transformada que es lineal en los parámetros, logarítmica en la variable X_i. MCO es aplicable.

(ii) La función tiene pendiente variable: depende del valor de X.

De (2.6) $$\frac{dY}{dX} = \beta_1 \left(\frac{d(LnX)}{dX}\right) = \beta_1 \left(\frac{1}{X}\right)\left(\frac{dX}{dX}\right)$$

entonces: $\frac{dY}{dX} = \frac{\beta_1}{X}$

(iii) β_1 es un parámetro importante, representa una semi-elasticidad de *Y* respecto a *X*: el estimador mide el cambio proporcional absoluto en *Y* para un cambio porcentual unitario en el valor de la variable regresora (variable *X*):

$$\frac{dY}{dX} = \beta_1 \left(\frac{d(LnX)}{dX}\right) = \beta_1 \left(\frac{1}{X}\right)\left(\frac{dX}{dX}\right) \longmapsto \beta_1 = \frac{dY}{dX/X}$$

β_1 = (cambio absoluto en *Y*) / [cambio relativo en *X*]
β_1 mide el cambio absoluto en *Y* como consecuencia de un cambio relativo en el valor de la variable regresora (variable *X*).

Ejemplo 2.4. en una muestra de familias de Huancayo se ha obtenido los siguientes resultados de corte transversal:

$$\hat{\boldsymbol{Y}}_{\boldsymbol{t}} = -1283.9 + 257.27LnX_t$$
R2= 0.49

Donde:
$\hat{\boldsymbol{Y}}_{\boldsymbol{t}}$= gastos familiares mensuales en alimentos (S/.)
X_t= gastos familiares totales del mes (S/.)

Siendo tales estimadores significativos estadísticamente, el coeficiente 257.3 implica que un incremento de 1% en el gasto total de las familias, en promedio, conlleva un incremento en casi 257.3/100 = S/ 2.6 soles en el gasto de alimentos de las familias incluidas en la muestra (coeficiente dividido entre 100).

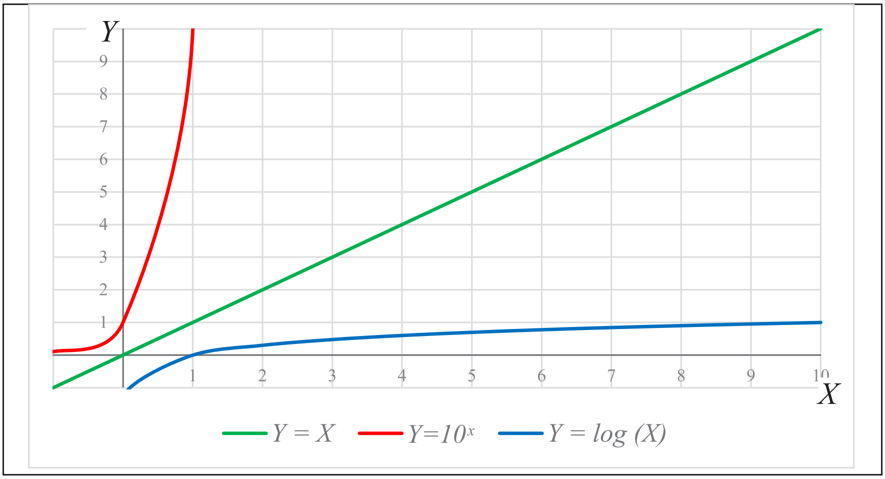


Figura 2.1. Formas funcionales lineal y semi-logarítmicas en escala lineal

2.2.5. Función recíproca

Función original: $Y = \beta_0 + \beta_1 \left(\frac{1}{X_1}\right) + \mu_t$ (2.8)

La función recíproca, a pesar de no ser lineal en la variable *X*, es originalmente lineal en los parámetros; por tanto, representa un modelo de regresión lineal que no necesita transformación en el marco de aplicabilidad del método de MCO.

Una característica importante del modelo recíproco es que a medida que *X* crece indefinidamente, el término $\beta_1 (\frac{1}{X_1})$ se acerca a cero, entonces *Y* se aproxima a un valor asintótico o valor límite β_0. El valor asintótico lo tomará la variable dependiente *Y*, cuando el valor de *X* crece indefinidamente.

Propiedades

(i) La función recíproca no necesita transformación; es originalmente lineal en los parámetros.

(ii) La función tiene pendiente variable pues depende del valor de *X*.

De (2.8) se deriva: $\frac{dY}{dX} = -\beta_1 X^{-1-1} \implies \frac{dY}{dX} = \frac{-\beta_1}{X^2}$ (2.9)

El hecho que, en la identidad (2.9), la pendiente sea variable implica que si β_1 es positivo, la pendiente siempre es negativa; mientras que si β_1 es negativo, la pendiente siempre es positiva. Véase Figura 2.2, casos (a) y (c).

(iii) A medida que *X* crece indefinidamente, $\beta_1 (\frac{1}{X_1})$ se acerca a cero, entonces *Y* se aproxima al valor asintótico o límite β_0 (casos (a) y (c) de la Figura 2.2). El valor asintótico lo tomará la variable dependiente *Y*, cuando el valor de *X* crece indefinidamente. Importante de esta función es realmente la existencia del valor asintótico, que es negativo en el caso (b) de la Figura 2.2.

$\beta_1 = -\frac{dY}{dX}(X^2)$ económicamente no significaría mucho.

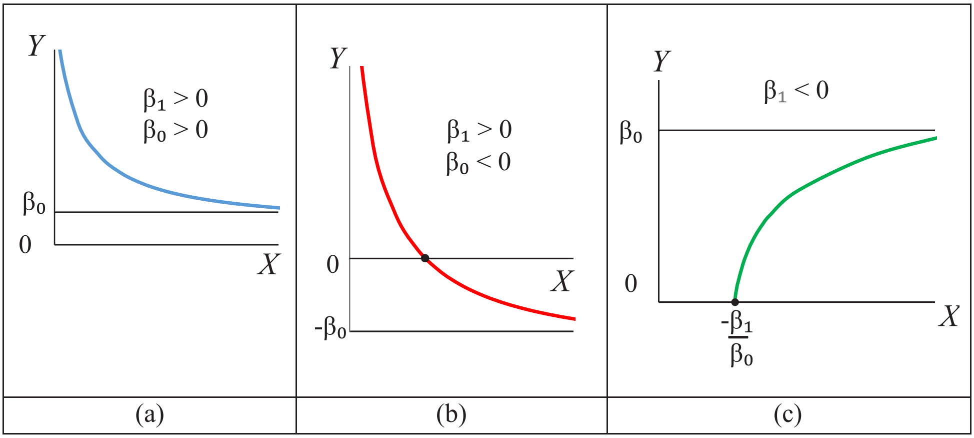


Figura 2.2. Gráficos de la función recíproca

Ejemplo 2.5. Curva de Phillips

Si la idea es determinar el efecto, en el nivel salarial, de una reducción porcentual en la tasa de desempleo, la relación puede especificarse mediante un modelo recíproco:

Modelo: $W = \beta_0 + \beta_1 \left(\frac{1}{D}\right) + \mu_t$

Estimación: $\widehat{W_t}$ = - 0.2594 + 20.588 $(\frac{1}{D_t})$

Donde
W = porcentaje de variación de los salarios (%)
D = porcentaje de desempleo (%).

Interpretación:
Asumiendo que los estimadores son estadísticamente significativos, el porcentaje de variación de los salarios depende del nivel de desempleo y del coeficiente 20.6.

β_0 = – 0.2594 (piso salarial), significa que a medida que aumenta la tasa de desempleo, la variación porcentual de los salarios puede llegar a ese valor límite, incluso reducirse hasta 0.25%.

2.2.6. Función cuadrática

La cuadrática es un tipo de función polinomial que, en economía, a menudo se emplea para captar efectos marginales, crecientes o decrecientes.
Considere la ecuación:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_1^2 \qquad (2.10)$$

β_0, β_1 y β_2 son parámetros a estimar.

Cuando $\beta_1 > 0$ y $\beta_2 < 0$, la relación entre Y y X tiene forma parabólica, con existencia de un valor máximo, como se muestra gráficamente, a continuación. Éste es un caso típico en economía.

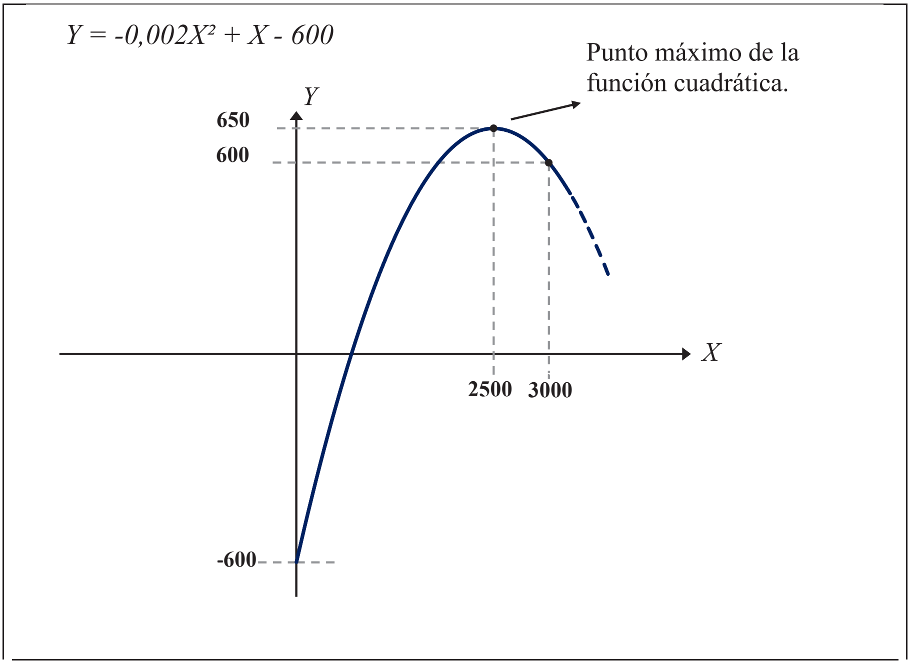


Figura 2.3. Máximo de una función cuadrática cuando $\beta_1 > 0$ y $\beta_2 < 0$.

Menos frecuente en Economía es el caso de la existencia de un mínimo, que se produce cuando los coeficientes de la ecuación son: $\beta_1 < 0$ y $\beta_2 > 0$.

Propiedades

(i) La función cuadrática de forma parabólica (cuando $\beta_1 > 0$ y $\beta_2 < 0$) es común en la representación de economía. Por ejemplo: una función de producción.

(ii) La relación original entre Y y X es originalmente lineal en los parámetros; no necesita transformación para la aplicación de MCO.
(iii) La función tiene una pendiente que es variable; depende del valor de X.

A partir de la ecuación (2.10): $\frac{dY}{dX} = \beta_1 + 2\beta_2 X$

(iv) Es posible demostrar que el máximo de la función tiene un lugar en un punto de inflexión del valor de X, denominado X^*:

$$X^* = \left| \frac{\beta_1}{(-2\beta_2)} \right|$$

El signo de β_2 permite identificar la existencia de un máximo o un mínimo (asociado a la forma gráfica de la función).

(v) La segunda derivada de Y respecto a X permite obtener la “tasa de cambio” o “tasa de crecimiento” de la función.

$$\frac{d^2Y}{dX^2} = 2\beta_2$$

En conclusión, se tiene un efecto marginal que es variable; tal efecto marginal es creciente antes que la función alcance el valor de X^*, luego de ello, cuando se produce un incremento en X el valor de Y disminuye.

Ejemplo 2.6. Considere la estimación de la siguiente ecuación:

Si $\widehat{W} = 3.73 + 0.298X - 0.061X^2$

Donde:
W= salario de ejecutivos en Chiclayo, US$ x hora
X= experiencia de los ejecutivos, en años

El resultado se visualiza en el Figura 2.4, que indica que si la experiencia (X) tiene relación directa con el salario (W), pero tal influencia se produce con un efecto decreciente sobre el salario. Es decir, por ejemplo, el primer año de experiencia vale US$ 0.298 la hora; el segundo año tal valor será: 0.298 – 2*(0.0061)*(1) = 0.286

centavos la hora, y así sucesivamente. Al pasar de 10 a 11 años de experiencia, por ejemplo, el salario por hora será de 0.298-2*(0.0061)*(10)=0.176 centavos por hora.

También es importante indicar que el mayor valor salarial (W) ocurre cuando la experiencia es de: $X^* = |\frac{\beta_1}{(-2\beta_2)}| = \frac{0.298}{-2*(-0.0061)} = 24.4$; ello conlleva finalmente a un valor de $W = 7.37$. Es decir, el salario de ejecutivos de Chiclayo, alcanza su máximo valor en US$ 7.37 la hora y cuando la experiencia es de X = 24.4 años.

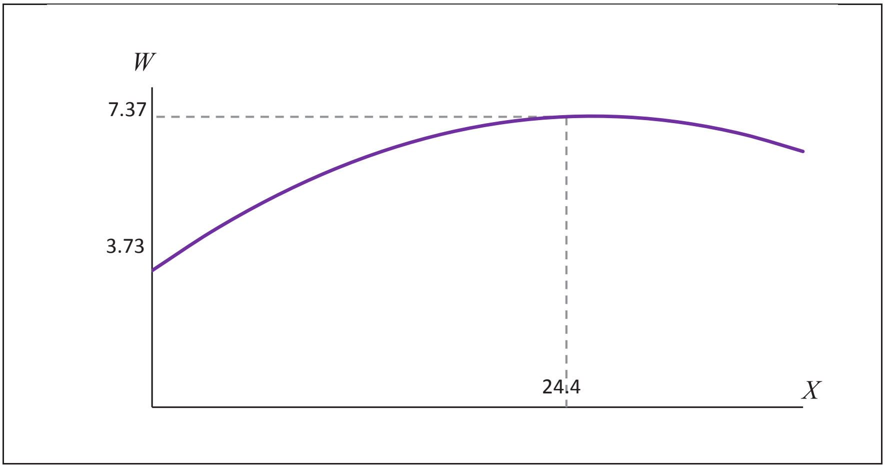


Figura 2.4. Función cuadrática correspondiente al ejemplo 2.6

En resumen, las principales características de algunas formas funcionales importantes en Economía son presentadas en la Tabla 2.2.

Tabla 2.2. Características de formas funcionales más usadas en Economía

Nombre de la Función	**Forma Funcional**	**Efecto Marginal (dY/dX)**	**Elasticidad (dY/dX).(X/Y)**
Lineal	$Y = \beta_0 + \beta_1 X_1$	β_1	$\beta_1(\frac{X_1}{Y})$
Lineal-Log	$Y = \beta_0 + \beta_1 LnX_1$	$\frac{\beta_1}{X_1}$	$\frac{\beta_1}{Y}$
Log-Lineal	$LnY = \beta_0 + \beta_1 X_1$	$\beta_1(Y)$	$\beta_1(X_1)$
Reciproca	$Y = \beta_0 + \beta_1(\frac{1}{X_1})$	$-\frac{\beta_1}{X_1^2}$	$-\frac{\beta_1}{X_1 Y}$
Log-Log	$LnY = \beta_0 + \beta_1 LnX_1$	$\frac{\beta_1 Y}{X_1}$	β_1
Cuadrática	$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_1^2$	$\beta_1 + 2\beta_2 X_1$	$[(\beta_1 + 2\beta_2 X_1)X_1]/Y$

Recomendaciones prácticas respecto a la Forma Funcional

(1) La teoría subyacente (por ejemplo, la teoría de la curva de Phillips) puede sugerir una determinada forma funcional.

(2) Los coeficientes del modelo escogido deberán satisfacer determinadas expectativas útiles, a priori. Por ejemplo, si se está considerando una función de demanda, se esperaría un coeficiente negativo para el parámetro de la relación entre precio y cantidad.

(3) A veces más de una forma funcional se ajusta razonablemente "bien" a los datos. Es pertinente y posible utilizar algunos indicadores de bondad de ajuste, que nos permite la selección de una determinada y precisa función.

(4) En general no se debe sobrevaluar el uso del coeficiente de determinación (R2). Lo que reviste mayor importancia son varias cosas juntas: la justificación teórica del modelo elegido, los signos de los coeficientes estimados, importancia estadística, etc.

(5) Es buena práctica calcular la tasa de cambio de *Y* sobre *X*, para fines comparativos. Una propiedad deseable es obtener estimadores cuya distribución tenga al parámetro como su "valor medio". Ésta es una definición absoluta**.**

2.3. Modelos con Variables Interactivas

Ocasionalmente, en modelos de regresión múltiple es importante considerar que el efecto parcial de una variable explicativa, respecto a la variable dependiente, es función de la magnitud de otra variable explicativa (algo así como un efecto parcial "doble", ver recuadro). Entonces estamos frente a un modelo de especificación que requiere el uso de un término de interacción de variables.

Coeficiente de regresión parcial

En regresión múltiple o multivariada se utiliza el concepto de "coeficientes de regresión parcial". Por ejemplo, en el modelo con dos variables regresoras: $Y = \beta_0 + \beta_{1/2}X_1 + \beta_{2/1}X_2$, significa que $\beta_{1/2}$ mide el efecto de X_1 en *Y*, considerando contantes los valores de X_2; o también que mide el efecto parcial de X_1 en *Y*, después de haber descontado el efecto de las demás variables regresoras (solo X_2 en este caso).

Usualmente, la presencia de interacciones en un modelo obedece a la propia teoría.

Ejemplo 2.7. Se tiene el siguiente modelo de corte transversal (adaptado de Wooldridge, 2010)

$$Precio = \beta_0 + \beta_1(Supt) + \beta_2(Dorm) + \beta_3(Dorm * Supt) + \mu_i \qquad (2.11)$$

Donde:
Precio = precio del alquiler de viviendas en zonas residenciales de Piura- Perú
Supt = superficie de las viviendas
Dorm = número de dormitorios.

El efecto de la cantidad de dormitorios (*Dorm*) en el precio de las residencias (*Precio*) será:

$$\frac{d(Precio)}{d(Dorm)} = \hat{\beta}_2 + \hat{\beta}_3(Supt) \qquad (2.12)$$

O sea, en este ejemplo, el efecto "real" de $Dorm$ en el *Precio* de las viviendas va más allá de solo $\hat{\beta}_2$ (que sería el efecto parcial). Si $\hat{\beta}_3$ es estadísticamente significativo y positivo, un dormitorio adicional implicaría un aumento mayor en el precio de alquiler de las viviendas, que el mero valor reportado por $\hat{\beta}_2$, pues requiere tomar en cuenta la superficie total de las viviendas.

La ecuación (2.11) quiere decir que el efecto que tiene el número de dormitorios (en el precio) depende no solo de sí mismo, sino también de la superficie de las viviendas: así, *en casas más grandes*, o de mayor superficie, un dormitorio adicional implicaría un aumento diferente en el precio de alquiler de las viviendas, que en el caso de *casas más pequeñas*. En otras palabras, existe un efecto interacción entre la superficie total de las casas y la cantidad de dormitorios. Será útil medir el efecto de dormitorios en el precio, pero evaluando (2.12) en valores -por ejemplo- promedio o en cuartiles superior o inferior de la muestra.

Ejemplo 2.8. Considerando un modelo para una clase universitaria de nivel inicial, que mediante MCO arroja los siguientes resultados:

$$\widehat{Exmat} = 2.05 - 0.0067(Atend) - 1.63(Adm) + 0.296(Adm^2) + 0.056(Atend)(Adm)$$

Estad. "*t*" (- 2.823) (- 4.328) (2.851) (1.994)

Donde:
Exmat = calificación examen final del curso Matemática I en la Universidad (0-20)
Atend = asistencia a clases en el semestre, %
Adm = calificación en el examen de admisión a la Universidad, (0-20)

Con todos los estimadores estadísticamente significativos, entonces la interpretación parcial (simple) del coeficiente asociado a *Atend* indicaría que la asistencia a clases tiene un efecto negativo en la calificación final del curso. Pero en realidad este coeficiente mide tal efecto solo cuando *Adm*= 0, lo cual no tiene interés, pues el coeficiente asociado con *Adm* es estadísticamente significativo (el valor muestral de *Adm* < 0).

En realidad, el efecto de *Atend* en la variable dependiente *Exmat* no es el mismo que el efecto parcial, siendo importante considerar los valores de *Adm*. En la muestra el valor promedio de *Adm* = 2.59, de forma que, en promedio, el efecto real de la variable *Atend* en *Exmat* (representado por $\bar{\bar{\beta}}_1$) es:

$$\frac{d(Exmat)}{d(Atend)} = \hat{\beta}_2 + \hat{\beta}_3(Adm)$$

$$\bar{\bar{\beta}}_1 = -0.0067 + 0.0056(2.59) = 0.0078.$$

Como *Atend* se mide en porcentaje, significa que, si *Atend* aumenta 1%, *Exmat* aumentará 0.078 desviaciones estándar, a partir de la puntuación media en el examen final. En otras palabras, con una mayor calificación en el Examen de admisión, la asistencia a clases incrementa MÁS la calificación final, a partir del promedio en el Examen de Admisión.

Ejercicios del capítulo II: Formas Funcionales Alternativas

2.1. (a) Responda a las preguntas del cuadro (con SI o NO), en base a los modelos adjuntos

(b) Sustente sus respuestas en base a las demostraciones econométricas pertinentes

Modelo	¿Es lineal en los parámetros?	¿Es intrínsecamente lineal?	¿Puede estimarse mediante uso de MCO?
(i) $Y_i = \frac{1}{1+\ e^{\beta_0+\beta_1 Xi+ui}}$			
(ii) $Y_i = \ \beta_o X_i^{\beta_1}\mu_i$			
(iii) $Y_i = \ \beta_o X_i^{\beta_1} + \mu_i$			

Sugerencias*: *Revisar la bibliografía pertinente.

2.2. Complete el cuadro siguiente y explique el significado de la pendiente y elasticidad**:**

	Forma Funcional Original	Forma Lineal (transformada)	Pendiente	Elasticidad
Lineal		$Y = \beta_0 + \beta_1 X_1$		
Logarítmica doble		$LnY = \beta_0 + \beta_1 LnX_1$		
Semi-logarítmica (LOG- LINEAL)		$LnY = \beta_0 + \beta_1 X_1$		
Semi-logarítmica (LINEAL- LOG)		$Y = \beta_0 + \beta_1 LnX_1$		
Recíproca		$Y = \beta_0 + \beta_1(\frac{1}{X_1})$		

Sugerencias: *revisar pp. 50-51 del libro de J. Pichihua (2002) y pp. 173 del libro de Gujarati & Porter (2010).*

2.3. Responda en relación a la interpretación requerida en la última columna

	Forma Lineal (transformada)	**Si en la función transformada, $\hat{\beta}_1$ es:**	**Interprete el valor de $\hat{\beta}_1$**
Lineal	$Y = \beta_0 + \beta_1 P_1$	*0.12*	
Doble logarítmica	$LnY = \beta_0 + \beta_1 LnP_1$	*0.25*	
Semi-logarítmica (LOG- LINEAL)	$LnY = \beta_0 + \beta_1 P_1$	*0.030*	
Semi-logarítmica (LINEAL- LOG)	$Y = \beta_0 + \beta_1 LnP_1$	*25*	
Recíproca	$Y = \beta_0 + \beta_1 (\frac{1}{P_1})$	*20.588*	

Sugerencias: *Revisar la sección 6.2 del libro de Wooldridge (2010) y pp. 50-51 de J. Pichihua (2002).*

2.4. En un modelo de regresión: $Y_i = \beta_0 + \beta_1 X_i + \beta_2 X_i^2 + \mu_i$

Interprete los resultados de la estimación, en lo que concierne a los estimadores de $\boldsymbol{\beta_1}$ y $\boldsymbol{\beta_2}$**;** así como a su significancia.

Sugerencias: *revisar sección 6.2 (pág192) del libro de Wooldridge (2010).*

2.5. Considere los siguientes dos modelos**:**

$$Ln(Y_t^*) = \alpha_1 + \alpha_2 Ln(X_t^*) + U_t^* \qquad (1)$$
$$Ln(Y_t) = \beta_1 + \beta_2 Ln(X_t) + U_t \qquad (2)$$

Donde Y_t^*=$w_{1*}Y_t$ y X_t^*=$w_{2*}X_t$, con las $\boldsymbol{w_s}$ constantes.

(a) Establezca las relaciones entre los dos conjuntos de coeficientes de regresión y sus errores estándar.

(b) ¿Es diferente el $\mathbf{r^2}$ en los dos modelos? Explique.

***Sugerencias*:**
(a) Ambos modelos tienen la misma pendiente, pero diferentes interceptos (grafique).

(b) Los coeficientes de determinación serán los mismos.

2.6. Considere los siguientes resultados de una función de crecimiento:

$\widehat{Ln(GS)} = \ 8.33 \ + \ 0.0071(T)$
(*EE*) = (0.0016) (0.00018) $r^2 = (0.982)$
Estad. ***t*** = (5202)* (39.2)*

Donde:
GS significa gasto en servicios, ***T*** representa el tiempo,
El símbolo (*) expresa alta significancia estadística del estimador.

Pruebe la hipótesis que el coeficiente de la pendiente no es significativamente diferente de 0.005.

***Sugerencias*:**
Ver páginas 162-165 y 177 del libro de Gujarati & Porter (2010), 5ª Ed. También sección 10.5 del libro de Wooldridge (páginas 360-364).

2.7. Considere el modelo:

$$Y = \frac{e^{\beta_0+\beta_1 Xi}}{1+ e^{\beta_0+\beta_1 Xi}} \qquad (1)$$

Tal como se presenta, ¿es un modelo de regresión lineal- soluble por MCO?

Si no fuera un modelo soluble por MCO ¿Que mecanismo podría utilizarse (si acaso) para convertirlo en un Modelo de Regresión Lineal que pueda resolverse utilizando MCO? ¿Cómo interpretaría el modelo resultante y en qué circunstancias sería adecuado dicho modelo?

Sugerencia:
La función no es lineal en los parámetros; por tanto, no es un modelo de regresión lineal. Sin embargo, se puede "linealizar" el modelo original y lograr la conversión:

$Ln\left[\frac{Yi}{1-Y_i}\right] = \beta_0 + \beta_1(X_i).$

Revisar el ejercicio 6.11 del libro de Gujarati & Porter (2010), pp. 177, y elabore las pruebas de hipótesis de los estimadores de ambos parámetros.

2.8. Considere el modelo de regresión**:**

$$Y = \beta_0 + \beta_1\left(\frac{1}{X_i}\right) + \mu$$

(a) ¿Es un modelo de regresión lineal que puede ser estimado mediante el uso de MCO?

(b) En caso (a) fuera negativa ¿Cómo se puede estimar el modelo?

(c) ¿Cuál es el comportamiento de la variable *Y* a medida que *X* tiende a infinito?

(d) Use un ejemplo económico apropiado y explique.

***Sugerencias*:**
La función es lineal en los parámetros. Por tanto:
(a) Si es un modelo de regresión lineal

(b) Se puede aplicar MCO al modelo original.
Revisar el caso 6.3 del libro de Gujarati & Porter (2010), pp. 176

2.9. Con un conjunto de datos sobre casas vendidas en Lurín (1988), en la siguiente ecuación se estima la relación entre el precio de las casas (*pr*) respecto a la distancia a un relleno sanitario (*dist*), en km.:

$\widehat{Ln(pr)} = 9.4 + 0.312\, Ln(dist)$
n =135, R^2= 0.162

(a) ¿Cuál sería el cambio que ocurriría en el precio de una casa si se incrementara en un diez por ciento la distancia promedio de una casa, respecto al relleno sanitario? ¿Es el signo de este estimador el esperado?

(b) ¿Cuál sería el precio promedio de una casa ubicada a un kilómetro de distancia del relleno?

(c) ¿Qué otros factores relacionados con una casa afectarían su precio? ¿Pueden estar estas variables correlacionadas con la distancia al incinerador? Si la respuesta fuera positiva, ¿que podría suceder en términos de los supuestos del modelo?

> ***Sugerencias*:**
> *Es un caso de una función de precios hedónicos. Si parecería adecuado el signo del estimador de β_1- Otro factor podría ser: la superficie de la casa. ¿Qué más?*

2.10. Suponga el modelo: $Y_i = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_1 X_2 + \varepsilon$
(adaptado de Gujarati & Porter (2010), ejercicio 8.9)

Donde:
Y_i = gasto de consumo personal (S/.)
X_1= ingreso personal (S/.)
X_2 = riqueza personal (S/.)

(a) ¿Qué significado tiene la interacción $X_1 X_2$?.

(b) Plantee las hipótesis y compruebe si la propensión marginal a consumir es independiente de la riqueza del consumidor?

> ***Sugerencias*:**
> *(a) Para un mejor entendimiento, estime los valores marginales de la V. dependiente (Y) con respecto a X_1 y X_2.*
>
> *(b) Cuando β_3=0, entonces la PMC será independiente del nivel de riqueza.*

Ejercicios para uso de Software

2.11. Utilizando los siguientes datos:

Y_i	86	79	76	69	65	62	52	51	51	48
X_i	3	7	12	17	25	35	45	55	70	120

(i) Ajuste el modelo: $\frac{100}{100-Y_i} = \beta_1 + \beta_2(\frac{1}{X_i})$

(ii) Obtenga las estadísticas de regresión e interprete.

Sugerencias:
$\hat{Y}^* = 2.0675 + 16.266 X^*$
EE (0.1596) (1.3232) $r^2 = 0.9497$

2.12. Empleando los datos del archivo ***RDCHEM*** (Wooldridge, 2010), utilice MCO para obtener los resultados del efecto de las ventas (*sales*) sobre la cantidad de residuos sólidos desechados (en %), de acuerdo a la siguiente ecuación:

$$(\widehat{rdintens}) = \beta_0 + \beta_1(sales) + \beta_2(sales)^2$$

Suponiendo significativos los estimadores de los parámetros,

(a) ¿En qué punto se vuelve negativo el efecto de "*sales*" sobre "*rdintens*"?

(b) ¿Conservaría el término cuadrático del modelo? Explique

(c) Defina una nueva variable "*salesbil*" como las ventas medidas en miles de millones de dólares: *salesbil*=(*sales/1000*). Escriba de nuevo la ecuación cuadrática, pero con "*salesbil*". No olvide considerar los errores estándar y el coeficiente de determinación.

(d) Con propósito de presentación de resultados, ¿qué ecuación preferiría?

***Sugerencias*:**
(a) El punto de "quiebre" es dado por $\frac{\beta_1}{(2\lvert\beta_2\rvert)}$ = *21.43* (ventas en millones de USD)
(b) Probablemente conservaría el término cuadrático.
(c) La ecuación estimada es: $(\widehat{rdintens}) = 2.613 + 0.3\,(salesbil) - 0.007(salesbil)^2$ *(coeficientes significativos).*
(d) La ecuación en (iii) es más fácil de interpretar que la primera.

2.13. Considere datos de viviendas vendidas durante 1991, en el Estado de Carolina del Norte (USA), que fue el año que se inició la construcción de un incinerador local de basura (tomado del ejercicio C6.1 del libro de Wooldridge, 2009). Utilice el archivo *KIELMC* para el siguiente planteamiento:

Para estudiar el efecto de la ubicación del incinerador sobre los precios de la vivienda, considere el modelo de regresión:

$Log(PRICE_t) = \beta_0 + \beta_1 Log(DIST_t) + U_t$

Donde:
PRICE es el precio de la vivienda (US$)
DIST es la distancia de la vivienda al incinerador (medida en metros).

(a) Interpretando esta ecuación de forma causal, ¿qué signo espera que tenga $\boldsymbol{\beta_1}$ si la presencia del incinerador hace decrecer el precio de la vivienda? Estime esta ecuación e interprete resultados.

(b) Al modelo de regresión simple anterior agregue las variables *Ln*(*INTST*), *Ln*(*AREA*), *Ln*(*LAND*), *ROOMS*, *BATHS*, *AGE*. Donde *INTST* es la distancia de la vivienda a la carretera principal, *AREA* es el área de la vivienda (m2), *LAND* es el tamaño del terreno (m2), *ROOMS* es el N° de habitaciones, *BATHS* es la cantidad de baños y *AGE* es años de antigüedad de la vivienda. Entonces, ¿qué concluye acerca del efecto del incinerador? Explique ¿por qué los incisos (i) y (ii) dan resultados contradictorios?

> ***Sugerencias*:**
> (*a*) *El efecto causal de la "distancia" sobre el "precio" significa que es mejor tener un incinerador lejos de la casa. La ecuación estimada es:* $\boldsymbol{L(\widehat{price}) = 8.05 + 0.365\, L(dist)}$ [*ambos coeficientes estadísticamente significativos y n=142, R2=0.18*].
>
> (*b*) *Cuando las variables Log(INTST), Log(AREA), Log(LAND), ROOMS, BATHS, AGE, son añadidas al modelo anterior (i), entonces el coeficiente de Log(DIST) es 0.55* (*con error estándar de 0.58*)*: efecto no significativo. Esto es así porque hemos controlado explícitamente por muchos otros factores que determinan la "calidad", precio de la casa (tal como "tamaño" y "número de baños") y su ubicación.*

2.14. Utilice el archivo *GPA2* para plantear y resolver las 4 preguntas planteadas debajo (adaptado del ejercicio C6.4 del libro de Wooldridge, 2009):

$Sat = \beta_0 + \beta_1(hsize) - \beta_2(hsize)^2$

Donde:
Sat = puntaje en el examen de admisión a la Universidad.
Hsize = tamaño promedio del grupo que termina sus estudios de secundaria (en miles).

(a) Estime *Sat* en función a *Hsize* en forma cuadrática. Comente.

(b) Empleando la ecuación planteada en (a) ¿cuál es el tamaño óptimo para un grupo de secundaria? Justifique.

(c) Encuentre el tamaño óptimo de una escuela secundaria, empleando *Ln*(sat) como variable dependiente. ¿Es muy distinto el resultado a lo obtenido en (a)? Comente.

***Sugerencias*:**

(a) La ecuación estimada:

$$\widehat{(sat)} = 997.98 + 19.81\ (hsize) - 2.13(hsize)^2$$

(EE) (6.20) (3.99) (0.55) n=4137

R2=0.0076

(b) Queremos el valor de (hsize) en que "sat" alcanza su máximo: 19.81/(2*2.13)=4.65. Dado que "hsize" está en miles, entonces 465 estudiantes es el tamaño óptimo de la clase.*

(c) Con Log(sat) como variable dependiente tenemos:

$$Lo\widehat{g(sat})= 6.896 + 0.0196\ (hsize) - 0.0021(hsize)^2$$

(EE) (0.006) (0.004) (0.00054) n=4137

R2=0.0078

Aunque el coeficiente de determinación sigue bajo, el tamaño óptimo de clase es ahora 469, que es muy cercano al valor obtenido antes.

CAPÍTULO III. INFERENCIA ESTADÍSTICA CON MODELOS RESTRINGIDOS

En un determinado modelo lineal general multivariado y en el contexto de uso de inferencia estadística para probar determinadas hipótesis acerca de los parámetros, la distribución ***F*** conjunta permite desarrollar un contraste de hipótesis de que todos los coeficientes de regresión son CERO, frente a otra hipótesis alterna de que al menos uno de los coeficientes sea diferente de cero. Incluso en el caso más simple, cuando deseamos probar la hipótesis nula de que un solo coeficiente de regresión es significativamente igual a cero, la prueba ***F*** se reduce a una prueba ***t*-**student, mediante la cual se contrasta un estadístico ***t*** calculado como la proporción del coeficiente estimado en relación al error estándar estimado, contra el estadístico ***t*** crítico, asociado a una distribución simétrica de probabilidad.

Contrates de hipótesis comunes en regresión multivariada
(tomado de Alarcón y Nolazco, 2014)

Las pruebas de hipótesis sobre coeficientes individuales son las más sencillas y sirven también para determinar la significancia estadística de los estimadores de los parámetros del modelo de regresión, en forma individual. En E-views los resultados de estas pruebas son automáticamente presentados en la primera regresión utilizando MCO. La idea es básicamente utilizar el estadístico "***t***" de Student para tomar decisión respecto a las hipótesis presentadas líneas debajo. $\boldsymbol{\beta_j}$ es el parámetro del modelo que se va a estimar, $\boldsymbol{\widehat{\beta}_j}$representa al estimador obtenido por MCO, $\boldsymbol{S(\widehat{\beta}_j)}$se refiere al error estándar del estimador. Es importante añadir que en el caso de la Ho: $\beta_j = 0$, el valor calculado de "t" es $t_c = \frac{\widehat{\beta}_j - 0}{S(\widehat{\beta}_j)}$

Resumen de Hipótesis para Parámetros Individuales

Hipótesis	Estadístico de Prueba	Criterio de Decisión
Ho: $\beta_j = 0$	$t_c = \frac{\hat{\beta}_j - \beta_j}{S(\hat{\beta}_j)}$	Se rechaza la Ho si: ***tc >*** $t_{(n-k)}$ ***gl***
Ha: $\beta_j \neq 0$	$S(\hat{\beta}_j) = \sqrt{V(\hat{\beta}_j)}$	No se rechaza Ho si: ***tc ≤*** $t_{(n-k)}$ ***gl***

La prueba de hipótesis conjunta utiliza el estadístico "***F***" asociado a una distribución "***F***", que sirve para probar si todos los parámetros son estadísticamente iguales a cero, o al menos uno es diferente. Este caso se resume en el siguiente Cuadro 4.3. Tanto los estadísticos "t" como "*F*" sirven para determinar la significancia (individual o conjunta) de los estimadores obtenidos mediante el método MCO.

Resumen del uso del Estadístico *"F"* en contrastes conjuntos

Hipótesis	Estadístico de Prueba	Criterio de Decisión
Ho: $\beta_2 = \beta_3 = ... = \beta_k = 0$	$F_{(k-1,n-k)} = \frac{\sum \hat{y}_i^2/(k-1)}{\sum \hat{e}_i^2/(n-k)}$	Se rechaza la Ho si: $Fc > F(k-1, n-k)\ gl$
Ha: al menos un $\beta_j \neq 0$		No se rechaza Ho si: $Fc \leq F(k-1, n-k)\ gl$

El estadístico ***"F"*** expresa la relación entre el Cuadrado Medio debido a la regresión en relación al Cuadrado Medio debido al Residual.

Sin embargo, en el marco de la estimación de los parámetros de un modelo lineal general, existen circunstancias en que deseamos realizar otro tipo de contrastes estadísticos, entre otros, los siguientes:

(i) Probar la significancia estadística de un *sub-conjunto* del total de coeficientes de regresión; esto es algo que se produce cuando se presume que sólo un grupo específico de variables regresoras explica los cambios en la variable dependiente.

(ii) Probar la veracidad estadística referida a la igualdad entre uno o más de los coeficientes o estimadores del modelo.

Tales circunstancias singulares son las que se presentarán a continuación.

3.1. Pruebas Estadísticas con la Distribución *F* de Fisher- Snedecor

3.1.1. Significancia estadística de un sub-conjunto de coeficientes de regresión

Suponga un modelo original de regresión múltiple

$$Y_i = \beta_1 X_{1i} + \beta_2 X_{2i} + \beta_3 X_{3i} + \beta_4 X_{4i} + \cdots + \beta_k X_{ki} + \mu_i \qquad (3.1)$$

Y suponga también que se desea probar si un grupo "***q***" de todos los coeficientes de regresión es conjuntamente igual a CERO; por ejemplo:

Ho: β_3 y $\beta_4 = 0$
Ha: al menos un parámetro es diferente de cero.

Es decir que considerando el modelo 3.1 como no-restringido[7], de ser cierta la hipótesis nula tendríamos el siguiente modelo (3.2) restringido:

$$Y_i = \beta_1 X_{1i} + \beta_2 X_{2i} + \mu_i \qquad (3.2)$$

(el número de coeficientes involucrados, en este caso, es ***q=2***).

La lógica detrás de la prueba de hipótesis es sencilla: cuando eliminamos las 2 variables (q) y estimamos el modelo restringido, entonces la suma de cuadrados residual (SCR) es mayor a la de SCR del modelo no-restringido ($SCR_r > SCR_{nr}$), debido a que dos variables de la ecuación (3.2) salieron del error y pasaron a ser variables explícitas en el modelo (3.1). Entonces si la hipótesis nula (Ho) es correcta, la eliminación de q variables tendrá "poco" efecto en el poder explicativo de la ecuación y SCR_r será ligeramente mayor a SCR_{nr}

El estadístico crítico apropiado es: $\boldsymbol{F} = \dfrac{(SCR_r - SCR_{nr})/(q)}{(SCR_{nr})/(n-k)}$

Donde:

SCR_r= suma de cuadrados residual del modelo restringido
SCR_{nr}= suma de cuadrados residual del modelo no restringido
q = número de restricciones; (n-k) = N° de parámetros del modelo NO-restringido.

F es un estadístico que también puede presentarse en función del coeficiente de determinación, como:

[7] Es un modelo no-restringido pues no se han hecho supuestos acerca de los coeficientes del modelo original especificado.

$$\boldsymbol{F} = \frac{(R_{nr}^2 - R_r^2)/(q)}{(1 - R_r^2)/(n-k)}$$

Ejemplo 3.1. se tiene la siguiente función de demanda de viviendas:

$$LnV = \beta_0 + \beta_1 Ln(P) + \beta_2 Ln(I) + \mu_i \qquad (3.3)$$

Donde:
V= cantidad de viviendas en un año
P= precio por unidad de vivienda en la localidad de la familia (en miles de soles)
I= ingreso familiar anual (en miles de soles)

Los coeficientes de regresión representan elasticidades. Se han obtenido los siguientes resultados de la estimación:

$$\widehat{LnV} = 4.17 - 0.247Ln(P) + 0.96Ln(I) \qquad R^2 = 0.371$$
$$EE \qquad (0.11) \quad (0.017) \qquad (0.026)$$

Entonces ambas elasticidades de demanda (– 0.247 y 0.96) son estadísticamente significativas. Sin embargo, sería interesante saber si la elasticidad ingreso de demanda β_2 es significativamente diferente de uno.

El estadístico calculado ***t*** sería:

$$t_{n-k} = \frac{\widehat{\beta}_2 - \beta_2}{S_{b2}} = \frac{0.96 - 1}{0.026} = -1.54 \qquad n = 117$$

Dados los valores críticos, provenientes de una tabla "***t*** ", no podremos rechazar la Ho.

Ejemplo 3.2. Supongamos que usamos la siguiente ecuación modificada:

$$LnV = \beta_0 + \alpha_1 D + \beta_1 Ln(P) + \alpha_2 D.Ln(P) + \beta_2 Ln(I) + \alpha_3 D.Ln(I) + \mu_i \quad (3.4)$$

Donde:
V= cantidad de viviendas en un año
P= precio por unidad de vivienda en la localidad de la familia (en miles de soles)
I= ingreso familiar anual (en miles de soles)
D= variable ficticia que representa el caso de familias migrantes (el valor es 1 cuando la familia es migrante y 0 en caso contrario).

Si los resultados fueron:

$$\widehat{LnV} = 4.2 + 0.006(D) - 0.221[Ln(P)] + 0.92[Ln(I)] - 0.114[D.Ln(P)] + 0.341[D.Ln(I)]$$

(Estad. *t*) (0.143) (11.05) (29.68) (1.87) (2.84)

R^2= 0.38

Con margen de error de 5%, las pruebas "*t* " en los coeficientes individuales indican que el primero es insignificante (coeficiente asociado a *D*) y los otros dos significativos. Sin embargo, deseamos también probar la hipótesis nula siguiente:

Ho: α_1=α_2=α_3=0

Entonces usamos $\boldsymbol{F} = \frac{(R_{nr}^2 - R_r^2)/(q)}{(1-R_r^2)/(n-k)}$ para la decisión sobre la prueba de hipótesis.

R_{nr}^2=0.380; R_r^2=0.371; *n*=3120; *k*=6; *q*=3

$\boldsymbol{F} = \frac{(0.380-0.371)/(3)}{(1-0.371)/(3114)} = 15.1$

Como *F*c > *Ft*, en todos los niveles de significancia existe suficiente evidencia para rechazar la hipótesis nula de demanda de vivienda idénticas para migrantes y no migrantes. Es decir, existe al menos un estimador estadísticamente diferente de cero, y no existe suficiente evidencia estadística que indique que la demanda de viviendas es diferente para migrantes y no migrantes.

3.1.2. Combinaciones lineales de los coeficientes de regresión diferentes

En ocasiones es también importante probar hipótesis que implican combinaciones lineales de los coeficientes de la ecuación, para lo cual es importante utilizar la técnica de uso de MCO en modelos restringidos.

Por ejemplo, suponga que se ha estimado la siguiente función:

$$Y = \beta_0 + \beta_1 X_{1t} + \beta_2 X_{2t} + \mu \qquad (3.5)$$

Donde:
Y= consumo; X_{1t}= ingreso laboral; X_{2t}= ingreso no laboral.

Podría ser útil probar la hipótesis que las propensiones marginales a consumir β_1 y β_2 son iguales:

Ho: $\beta_1 = \beta_2$ frente a Ha: $\beta_1 \neq \beta_2$

Entonces la ecuación (3.5) representa el modelo no restringido de dos variables.

Bajo la certidumbre de la hipótesis nula, tendremos la ecuación del modelo restringido siguiente:

$$Y_i = \beta_0 + \beta_1 (X_{1i} + X_{2i}) + \mu_i \qquad (3.6)$$

El estadístico ***F*** apropiado estaría dado por:

$\boldsymbol{F} = \frac{(SCR_r - SCR_{nr})/(q)}{(SCR_{nr})/(n-k)}$ o también por: $\boldsymbol{F} = \frac{(R_{nr}^2 - R_r^2)/(q)}{(1-R_r^2)/(n-k)}$

Con el número de restricciones $q = 1$.

También, por ejemplo, en una ecuación sobre los determinantes del consumo agregado:

$C = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \mu$ (3.7)

Donde:
C= consumo; X_{1i}= ingreso laboral; X_{2i}= ingreso no laboral.

Podría ser útil probar la hipótesis de que la propensión marginal a consumir es UNO.

Es decir, Ho: $\beta_1 + \beta_2 = 1$ contra Ha: $\beta_1 + \beta_2 \neq 1$

El modelo no restringido es el de la ecuación (3.7).

Para obtener el modelo restringido, en (3.7) substituimos: $\beta_2 = 1 - \beta_1$, obteniendo:

$Y_i = \beta_0 + \beta_1 X_{1i} + (1 - \beta_1) X_{2i} + \mu_i$

$(Y_i - X_{2i}) = \beta_0 + \beta_1 (X_{1i} - X_{2i}) + \mu_i$ (3.8)

El contraste de las hipótesis planteadas, requiere el uso del estadístico que siga una distribución ***F*** en el marco del modelo restringido:

$\boldsymbol{F} = \frac{(R_{nr}^2 - R_r^2)/(q)}{(1-R_r^2)/(n-k)}$, con una sola restricción ($q = 1$).

Ejemplo 3.3. Supongamos la siguiente ecuación de demanda por viviendas:

$LnV = \beta_0 + \alpha_1 D + \beta_1 Ln(P) + \alpha_2 D.Ln(P) + \beta_2 Ln(I) + \alpha_3 D.Ln(I) + \mu_i$
(3.9)

V= cantidad de viviendas demandadas en un año
P= precio por unidad de vivienda en la localidad de la familia (en miles de soles)
I= ingreso familiar anual (en miles de soles)
D= variable ficticia que representa el caso de familias migrantes (es 1 cuando la familia es migrante y 0 en caso contrario).

Y deseamos probar la hipótesis que la elasticidad ingreso de la demanda para viviendas de emigrantes es igual a 1; es decir: Ho: $\beta_2 + \alpha_3 = 1$ frente a la Ha: $\beta_2 + \alpha_3 \neq 1$.

Substituyendo α_3 en el modelo no-restringido (3.9), tendremos el modelo restringido:

$$LnV = \beta_0 + \alpha_1 D + \beta_1 Ln(P) + \alpha_2[D.Ln(P)] + \beta_2 Ln(I) + (1-\beta_2)[D.Ln(I)] + \mu_i$$

$$LnV - D.Ln(I) = \beta_0 + \alpha_1 D + \beta_1 Ln(P) + \alpha_2[D.Ln(P)] + \beta_2 Ln(I) - (\beta_2)[DLn(I)] + \mu_i \quad (3.10)$$

La ecuación (3.9) representa al modelo no restringido y la (3.10) el modelo restringido.

Los resultados de (3.9) fueron:

$$\widehat{LnV} = 4.2 + 0.006(D) - 0.221[Ln(P)] + 0.92[Ln(I)] - 0.114[D.Ln(P)] + 0.341[D.Ln(I)]$$

(Estad. *t*) (0.143) (11.05) (29.68) (1.87) (2.84)

R^2= 0.380

Mientras que de la ecuación (3.10) se ha estimado R^2= 0.3785. Con solo una restricción, la prueba arroja como resultado el valor calculado:

$$\boldsymbol{F} = \frac{(0.380-0.378)/(1)}{(1-0.378)/(3114)} = 7.56$$

A un nivel de significancia de 0.05, rechazamos la hipótesis nula, concluyendo que la elasticidad ingreso de la demanda para viviendas de migrantes no es igual a 1.

3.2. Pruebas Estadísticas Adicionales con Modelos Restringidos

En la mayoría de casos no hay razón para ir más allá de los estadísticos "*t*" y "*F*", dado que con el cumplimiento de los supuestos básicos del Teorema de Gauss Markov, tales estadísticos ("*t*" y "*F*") han mostrado ser adecuados en la provisión de estimadores insesgados y consistentes; las pruebas se justifican mejor si sumamos a ello el supuesto de normalidad de la distribución del término de error[8]. Sin embargo, en el campo del análisis asintótico, es decir con existencia de muestras "grandes"[9], existen otras formas alternativas, eventualmente mejores, de probar restricciones de **exclusión múltiple** o la significancia estadística de subconjuntos de parámetros. Son los casos del Multiplicador de Lagrange (ML) y el estadístico de Razón de Verosimilitud, que han tenido importante popularidad en Economía y que son brevemente presentados a continuación.

---

[8] Son casos especiales del estadístico Wald.

[9] Muestra grande es un concepto "relativo"; se entiende aquí como: $n \longmapsto$ infinito

Eventualmente, en análisis asintótico los contrastes que usan el Multiplicador de Lagrange (ML) o el estadístico de Razón de Verosimilitud, han probado ser superiores a los de la prueba *F,* en casos especiales de modelos intrínsecamente lineales y modelo no-lineales.

3.2.1. Prueba del multiplicador de Lagrange (*ML*)

El nombre "*estadístico de multiplicador de Lagrange*" (ML) proviene de un proceso de *optimización restringida de una función*. Se apoya también en los supuestos de Gauss Markov, que son los mismos que justifican el uso del estadístico "***F***" en muestras grandes, pero se basa en la utilización de una función de máxima verosimilitud; el supuesto de normalidad no es necesario.

A diferencia del caso previo, en que se usa la distribución "*F*", el procedimiento de uso del estadístico ML comienza con la hipótesis nula proporcionada por el modelo restringido (MR) y culmina con la incorporación de la información proporcionada por el modelo no-restringido (MNR); la prueba examina si un movimiento en la dirección de la hipótesis alternativa puede mejorar, en forma significativa, el poder explicativo del modelo restringido[10].

Consideremos que $\hat{\beta}_{nr}$ es el estimador de máxima verosimilitud de un parámetro del modelo sin restricción, y supongamos que $\hat{\beta}_r$ representa los parámetros del modelo restringido. Entonces nuestro objetivo es maximizar la función restringida [Ln L(β_r)] – λ (β_{nr}= β_r); es decir:

Max [Ln L(β_r)] – λ (β_{nr}= β_r) (3.10)

Donde:
Ln L(β_r) es el logaritmo de la función de verosimilitud
λ = multiplicador de Lagrange
β_{nr}= β_r expresa la restricción.

Intuitivamente, el valor máximo de la función de verosimilitud se logrará cuando la restricción se cumpla en forma exacta. El multiplicador de Lagrange mide la valoración marginal asociada con la restricción: a mayor valor del término λ (β_{nr}= β_r), mayor será la reducción en el valor máximo de la función LnL(β_r), en la medida que la restricción se cumpla. Es decir, el multiplicador de Lagrange es:

[10] Desde la perspectiva de la prueba de WALD (uso del estadístico *F*) el proceso comienza con el modelo NO restringido y pregunta si la imposición de un *conjunto de restricciones* disminuye de manera significativa el poder explicativo del modelo original de regresión.

$$\frac{\partial Ln\ \mathrm{L}(\beta_r)}{\partial \beta_{nr}} = \lambda$$

La prueba LM se usa frecuentemente en el caso que se considere la posibilidad de agregar variables explicativas adicionales a un modelo inicialmente restringido.

Ejemplo 3.4. Suponga el modelo restringido:

$$Y_i = \beta_1 X_{1i} + \mu_r \qquad (3.11)\ \ \mathrm{MR}$$

Y se considera la posibilidad de agregar dos variables adicionales contenidas en un modelo sin restricciones:

$$Y_i = \beta_1 X_{1i} + \beta_2 X_{2i} + \beta_3 X_{3i} + \mu_{nr} \qquad (3.12)\ \ \mathrm{MNR}$$

Ho: $\beta_2 = \beta_3 = 0$ contra
Ha: al menos alguna de las β_s es diferente de cero.

Entonces la prueba de la hipótesis que c/u de las q variables adicionales (2 en este caso) tiene un coeficiente cero, se realiza del siguiente modo:

(i) Calculando primero los residuales del modelo restringido (3.11); específicamente:

$$\hat{\mu}_{ir} = Y_i - \beta_1 X_{1i}$$

(ii) Considerando luego la regresión de la serie de valores residuales ($\hat{\mu}_{ir}$) con respecto a todas las variables explicativas "adicionales", provenientes del modelo no-restringido:

$$\hat{\mu}_{ir} = \delta_0 + \delta_1 X_{1i} + \cdots + \delta_k X_{ki} + error \qquad (3.13)$$

(iii) Decisión: si todas las variables "adicionales" fueran irrelevantes, entonces sus coeficientes serían cero cuando se pasa del modelo restringido al no-restringido. Sin embargo, si alguna o todas las variables adicionales son significativamente determinantes de Y, esperamos que sus coeficientes sean estadísticamente significativos; por tanto, que la ecuación (3.13) se estimará con un buen ajuste.

La prueba del ML se determina con base en una prueba de significancia de la regresión en la ecuación (3.13). Específicamente, mediante el uso de un indicador que sigue una distribución Ji-cuadrada:

$$LM = nR_r^2 \sim \chi_q^2$$

Donde "q" es el N° de restricciones; "n" es el tamaño de muestra.
R_{nr}^2 es el coeficiente de determinación de la ecuación 3.13.

Las condiciones de rechazo (o no) de la hipótesis nula (Ho), son las comunes: con un nivel de significancia dado, determinados los grados de libertad y el planteamiento adecuado de las hipótesis, se determina un valor crítico frente al que se contrasta el estadístico en cuestión. Eventualmente es más sencillo y práctico calcular el valor de "p", el menor nivel de significancia al que puede rechazarse la hipótesis nula, de manera tal que la hipótesis pueda probarse a cualquier nivel de significancia.

Es importante añadir también que, asintóticamente, el estadístico de multiplicador de Lagrange y el estadístico *F* proveen resultados muy semejantes (con muestras grandes). También, como ocurre con el estadístico "*F*", se debe tener cuidado de utilizar el mínimo número de observaciones en los pasos (i) y (ii) del procedimiento antes indicado.

Por ejemplo, suponga el siguiente modelo no-restringido:

$$Y_i = \beta_1 X_{1i} + \beta_2 X_{2i} + \beta_3 X_{3i} + \beta_4 X_{4i} + \mu_i \qquad (3.14)$$

Deseamos usar el estadístico ML para probar si X_{3i} y X_{4i} tienen, o no, efecto significativo en *Y*, luego de controlados los efectos de las demás variables regresoras. Es decir:

Ho: β_3 y $\beta_4 = 0$ frente a
Ha: al menos un parámetro es diferente de cero.

Paso 1. Estimamos el modelo restringido:
Obteniendo $\hat{\mu}_{ir}$ con "n" observaciones muestrales.

Paso 2. Estimamos la ecuación: $\hat{\mu}_{ir} = \delta_0 + \delta_1 X_{1i} + \cdots + \delta_k X_{ki} + error$

Suponiendo n=2723, R_μ^2 = 0.015, entonces: $LM = nR_r^2$ = 2725(0.015) = 40.875.

El estadístico χ_q^2 con dos grados de libertad es aproximadamente 4.6 (con 10% como margen de error). Por tanto, a ese nivel de significancia, se puede rechazar la hipótesis nula y concluir que las variables X_{3i} y X_{4i} si tienen efecto significativo sobre la variable dependiente *Y*.

3.2.2. Prueba de razón de verosimilitud (PRV)

Al igual que en el caso del estadístico *Multiplicador de Lagrange*, en este caso la PRV se usa también si se requiere probar que ciertas restricciones de parámetros están apoyadas por los datos. También es una prueba asintótica; usa un estadístico elaborado en base a funciones derivadas del método de Máxima Verosimilitud. No requiere de la condición de normalidad de la distribución del término de error.

En este caso, si creemos que un modelo apropiado es:

$$Y_i = \beta_1 X_{1i} + \beta_2 X_{2i} + \beta_3 X_{3i} + \mu_i \qquad (3.15)\ \text{MNR}$$

Además, tenemos la convicción que la variable X_{1i} está bien especificada y es significativa, pero no tenemos seguridad sobre las demás variables regresoras X_{2i} y X_{3i}. En tal caso se tendría:

Ho: $\beta_2 = \beta_3 = 0$
Ha: al menos alguna de las β_s es diferente de cero.
(sería un caso de un modelo original "sobre-especificado" por la presencia de variables irrelevantes)

$$Y_i = \beta_1 X_{1i} + \mu_i \qquad (3.16)\ \text{MR}$$

La PRV utiliza el coeficiente: $\lambda = \frac{L(\beta_r)}{L(\beta_{nr})}$

Donde: $L(\beta_{nr})$ representa el valor máximo de la función de verosimilitud cuando se aplican restricciones; $L(\beta_r)$ representa el valor máximo de la FV cuando no se aplican restricciones.

Como el denominador se basa en un modelo no restringido, tal cantidad debe ser al menos tan grande como el numerador; entonces $0 < \lambda < 1$. Si la Ho es verdadera entonces esperamos $\lambda \doteq 1$; por tanto, podríamos rechazar la Ho cuando λ es "suficientemente" pequeña, en función del valor crítico o valor tabular.

En resumen, la prueba se basa en el hecho que, para muestras grandes, se tiene la siguiente identidad:

$$PRV = -2[L(\beta_r)\text{-}\ L(\beta_{nr})] \sim \chi^2_m$$

Donde "*m*" es el número de restricciones impuestas al modelo no-restringido.

Con cierto nivel de significancia (α= 0.05 por ejemplo), procedemos entonces con el procedimiento clásico de prueba de hipótesis.

En la mayor parte de las situaciones que implican modelos lineales, con muestras grandes, las pruebas "*F*" y PRV deberían generar resultados muy parecidos, pero la PRV podría ser más atractiva en la medida que no requiera del supuesto de normalidad del error.

Ejemplo 3.5. Considere el siguiente modelo de regresión no lineal:
(adaptado de Greene (2012), sección 14.6.4, pp. 148-149)

$$f(Y_i/X_i,\beta) = \frac{1}{\beta+X_i}\left[EXP\left(\frac{Y_i}{\beta+X_i}\right)\right] \qquad (3.17)$$

Donde: Y_i es renta familiar anual, X_i es el nivel de educación de padres de familia.

El modelo (3.17) es un caso especial (MR) de una función de distribución probabilística gamma $\Gamma(\rho)$[11], que sería entonces el modelo no restringido:

$$f(Y_i/X_i,\beta,\rho) = \frac{\beta_i^{\rho}}{\boldsymbol{\Gamma}(\boldsymbol{\rho})}\; Y_i^{\rho-1}[EXP(-Y_i\beta_i)] \qquad \text{MNR}$$

(En el MNR se ha simplificado: $\beta_i = \frac{1}{\beta+X_i}$)

La restricción considerada es:

Ho: $\rho = 1$; Ha: $\rho \neq 1$.

Utilizando el método de máxima verosimilitud, los resultados relevantes reportados fueron los siguientes:

Tabla 3.1. Estimaciones de máxima verosimilitud, ejemplo 3.4

Cantidad	Estimación no restringida	Estimación restringida
β	- 4.719	15.60
ρ	3.151	1.000
$Ln(L)$	- 82.916	- 88.436
V($\hat{\beta}$)	5.773	46.164
V($\hat{\rho}$)	0.663	0.000

El estadístico de contraste de razón de verosimilitud resultó ser:

$\lambda = -2[-88.44 - 82.92)] = 11.04$

Con un grado de libertad y 5% de margen de error, el valor crítico de las tablas es 3.842, entonces tenemos suficiente evidencia para rechazar la hipótesis nula; concluyendo que: $\rho \neq 1$.

Comparación resumida de pruebas

En su forma general, la prueba de Wald guarda estrecha similitud con las pruebas del multiplicador de Lagrange y el de Razón de Verosimilitud, dado que todas se basan en la diferencia entre dos modelos, no restringido y restringido. En el caso

[11] La función de distribución probabilística GAMMA se describe, en detalle, en el Apéndice B, sección B4.5, de W. H. Greene (7ª edición, 2012).

especial del modelo de regresión lineal (en los parámetros), la prueba de WALD se simplifica a una prueba "*F*"

$$\boldsymbol{F} = \frac{(R_{nr}^2 - R_r^2)/(q)}{(1 - R_r^2)/(n-k)}$$

Asintóticamente las 3 pruebas son prácticamente idénticas. Sin embargo, como regla general, difieren dentro de las propias muestras, y pueden generar pruebas de significancia diferentes y en ocasiones conflictivas. En la medida que el modelo sea lineal, la prueba de Wald siempre será la estadística de prueba "mayor" y la prueba LM proporcionará la "menor".

Cuando está implicados modelos lineales, es fácil aplicar la prueba de WALD debido a que es más simple estimar ambos modelos, restringidos y no restringidos. Sin embargo, cuando están implicados modelos más complejos o modelos no-lineales, la prueba LM puede proporcionar una alternativa atractiva, dado que depende en forma directa solo de la estimación del modelo restringido. LM ha mostrado también amplia versatilidad en múltiples aplicaciones econométricas.

Ejercicios del capítulo III: Inferencia estadística con modelos restringidos

3.1. Considere el siguiente modelo: $\boldsymbol{Y_t = \beta_0 + \beta_1 X_{1t} + \mu_t}$

(a) Si utiliza el método de máxima verosimilitud, escriba la función objetivo explícita, a optimizar, y explique las características del procedimiento.

(b) ¿Bajo qué supuestos los estimadores $\hat{\beta}_0$ y $\hat{\beta}_1$ serían de varianza mínima? Indique brevemente tres casos en que estos supuestos no se satisfacen y las consecuencias que ello tendría.

(c) Suponiendo que se cumple la hipótesis que β_1= 2, ¿cómo sería el modelo restringido y cómo sería el valor restringido de β_0: β_{r0}.

> ***Sugerencia****s: Revisar literatura pertinente.*
>
> *(a) La función objetivo sería: Ln [L(β_0, β_1, σ^2)]. Faltaría determinar el lado derecho de la ecuación.*
>
> *(b) El teorema de Gauss-Markov indica que si $\boldsymbol{X_{1t}}$ no es estocástica y se cumple que μ_t ~iid **(0,** σ^2) entonces los estimadores β_0 y β_1, obtenidos por MCO, son ELIO. También los estimadores obtenidos por máxima verosimilitud que hemos usado acá, coinciden con los de MCO.*
>
> *(c) Incorporando la restricción, en el modelo no restringido, se obtiene:*
> $Y_t = \beta_0 + 2X_{1t} + \mu_t$ ► $Y_t - 2X_{1t} = \beta_0 + \mu_t$
> *Entonces utilizando MV será posible obtener el estimador restringido:* $\beta_{r0} = \bar{Y} - 2\bar{X}_{1t}$

3.2. Considere el siguiente modelo para un caso hipotético de la economía peruana: (ejercicio adaptado de Gujarati & Porter, ejercicio 8.7 pp 263)

$$\boldsymbol{Y_t = \beta_0 + \beta_1 X_{1t} + \beta_2 X_{2t} + \mu_t}$$

Donde:
$\boldsymbol{X_{1t}}$ significa el gasto en bienes importados (millones de S/.), $\boldsymbol{X_{2t}}$ representa el ingreso personal disponible y "***t*** " es una variable de tendencia temporal.

Considerando que se han estimado los siguientes dos modelos, en el período 2000-1017:

$\widehat{\boldsymbol{Y}}_{\boldsymbol{i}} = -859.92 + 0.647\boldsymbol{X}_{\boldsymbol{1t}} - 23.195X_{2t}$ R2=0.977

$\widehat{\boldsymbol{Y}}_{\boldsymbol{i}} = -261.10 + 0.2452\boldsymbol{X}_{\boldsymbol{1t}}$ R2=0.940

Indique si es cierto o falso que el error estándar ($\boldsymbol{S_2}$) asociado a la variable X_{2t} es 4.7 aproximadamente.

> ***Sugerencia:*** *utilice la relación entre "t", F y R2*
>
> *1° Estime F para modelos restringidos y obtenga $F_{(1,17)}$ en el modelo (1)*
>
> *2° Utilice la relación $F_{(1,17)}=t_{17}^2$ y $t_{17}^2=(\frac{\widehat{\beta}_2}{S_2})$ para despejar y estimar finalmente el valor de $\boldsymbol{S_2}$ aproximadamente 4.46.*

3.3. Si usted tuviera los siguientes resultados de regresión:
(ejercicio tomado de Gujarati Ejercicio 8.7, pp. 264)

$\widehat{\boldsymbol{Y}}_{\boldsymbol{t}} = 16899.0 - 2978.5\boldsymbol{X}_{\boldsymbol{1t}}$ (1) R2=0.615
"*t*" (8.515) (−4.728)

$\widehat{\boldsymbol{Y}}_{\boldsymbol{t}} = 9734.2 - 3782.3\boldsymbol{X}_{\boldsymbol{1t}} + 2815X_{2t}$ (2) R2=0.771
"*t*" (3.371) (−6.61) (2.971)

Estime el tamaño de la muestra en base al que se han obtenido los resultados indicados.

> ***Sugerencias*:**
> *Utilice la relación $F=t^2$ para:*
> *(i) Obtener, de la ecuación (1): $F_{(1,n-k)}=t_{n-k}^2$. F=22.35*
>
> *(ii) Utilice la ecuación de F como función de R2, n y K, para despejar "n" y comprobar que n= 16 aproximadamente.*

3.4. A menudo, la función de Cobb-Douglas se utiliza, en microeconomía, para representar la relación entre la producción de una empresa, *P*, y los factores de producción *L* y *K*.

(a) Plantee las dos hipótesis requeridas para contrastar la hipótesis de rendimientos constantes a escala frente a la alternativa de rendimientos decrecientes en la producción. También establezca el estadístico y la distribución de probabilidad a usar para el contraste.

(b) Suponiendo existen rendimientos constantes a escala, a fin de obtener apropiados estimadores ¿cómo incorporaría esta información a la ecuación original?

(c) Para el caso hipotético de una empresa peruana de producción de espárragos, la estimación MCO ha permitido obtener:

$$lnP = 1.37 + 0.632(lnK) + 0.452(lnL)$$

(*EE*) (0.257) (0.220) *R^2= (0.908)*

Utilizando tales resultados y también lo obtenido en (a) y (b), para contrastar las dos siguientes hipótesis:
Ho: $\beta_k + \beta_l = 1$ frente a Ha: $\beta_k + \beta_l \neq 1$.
Ho: $\beta_k = \beta_l$ frente a Ha: $\beta_k \neq \beta_l$.

(d) El problema no proporciona tamaño de muestra. ¿Afecta esto a las conclusiones? ¿Cómo caracterizaría esta función de producción de espárragos?

***Sugerencias*:**

(a) La función de producción original es: $Y_t = e^{\beta_0} X_t^{\beta_1} X_t^{\beta_2} e^{\mu}$. Se sugiere usar logaritmos para linealizar el modelo, luego plantear un contraste de hipótesis de una sola cola, estadístico "t".

(b) Si hubieran rendimientos constantes a escala, entonces los parámetros del modelo se deberían estimar así: $p_t = \beta_0 + \beta_k X_{kt} + (1 - \beta_k) l_t + \mu_t$ (valores de las variables representan logaritmos).

(c) Se contrasta "rendimientos constantes a escala" frente a "rendimientos no constantes a escala" y se puede usar una prueba "t".

3.5. Considere el modelo de producción adaptado del ejercicio 8.5 del libro de Gujarati, (pp 262).

$$Y_t = \beta_0 L_t^{\beta_1} K_t^{\beta_2} \qquad (1)$$

Donde ***Y***, ***L*** y ***K*** representan el nivel de producción y niveles de uso de los factores capital y trabajo, respectivamente.

Al dividir la ecuación (1) entre *K* obtenemos:

$$\frac{Y_t}{K} = \beta_0 \left(\frac{L}{K}\right)^{\beta_1} K^{\beta_2 - 1} \qquad (2)$$

Al aplicar logaritmos obtenemos:

$Ln\,(\frac{Y_t}{K}) = \beta_t + \beta_1 Ln\left(\frac{L}{K}\right) + \beta_2 - 1(LnK) + \mu_t$ (3)

Nota: $\beta_t = Ln\ \beta_o$.

(a) Con la ecuación (3) plantee e indique cómo probaría la existencia de rendimientos constantes a escala.

(b) Si hubiera "*rendimientos constantes a escala*", ¿cuál sería la interpretación adecuada de la ecuación (3)?.

(c) Si la ecuación (1) se divide entre "*L*" en lugar de "*K*" , como en (2). ¿Cuál sería la diferencia?

Sugerencias:

(a) Se puede usar una distribución "t" para probar la hipótesis nula: $\beta_1 + \beta_2 - 1 = 0$.

(b) Si definimos la relación producto/capital (Y/K) como una medida de la productividad del capital, y luego (L/K) como la relación entre trabajo y capital, entonces la pendiente en la ecuación (3) nos ofrece una relación de causalidad interesante que debe usted decidir.

3.6. Una empresa ha estimado dos modelos de previsión de sus ventas (***V***, en soles):

$V_t = \beta_0 + \beta_1 P_t + \beta_2 C_t + \mu_t$ (1) SCR= 43.75

$V_t = \beta_0 + \beta_1 P_t + \beta_2(-2P_t + C_t) + v_t$ (2) SCR=230.0

P_t representa los precios de venta, en S/., C_t representa la calidad del producto (36 meses últimos).

Contraste la restricción impuesta en el modelo (2).

***Sugerencias*:**

La hipótesis a contrastar es: $\beta_1 = -2\beta_2$. *El estadístico a usar será* ***F*** *en función de las Sumas de Cuadrados.*

3.7. Suponga que se utiliza un modelo de regresión para evaluar los determinantes de la producción de anchoveta en Perú (*Y, en Tm),* en función al *nivel de la biomasa*

($\boldsymbol{X_{1t}}$ en Tm.) y el número de embarcaciones dedicadas ($\boldsymbol{X_{2t}}$). Se han obtenido los siguientes resultados:

$\hat{\boldsymbol{Y}}_t = 1.43 + 2.80\boldsymbol{X_{1t}} + 1.54\boldsymbol{X_{2t}}$ SCE = 29004.9 SCT = 29460 n=10

Si se ha vuelto a estimar el modelo, pero bajo la Ho: β_2= 0, habiéndose obtenido los siguientes resultados**:**
$\hat{\boldsymbol{Y}}_t = -4.06 + 3.61\boldsymbol{X_{1t}}$ R2=0.9835

Mediante un contraste de hipótesis verifique la adecuación de introducir la restricción en el modelo.

> ***Sugerencias*:**
> *(i) se obtiene la SC del modelo sin restringir, (ii) se calcula SC del modelo restringido, utilizando:* $R^2 = 1 - \frac{SCR}{SCT}$
> *(ii) entonces la hipótesis se prueba utilizando el estadístico **F**, en función de las Sumas de Cuadrados.*

Ejercicios para uso de software

3.8. Considere el modelo: $\boldsymbol{Y_t = \beta_0 + \beta_1 X_{1t} + \beta_2 X_{2t} + \mu_t}$

Donde:
$\boldsymbol{X_{1t}}$ representa el gasto en bienes importados (millones de S/.), $\boldsymbol{X_{2t}}$ expresa el ingreso personal disponible (millones de S/.).

Utilizando los siguientes datos hipotéticos:

N°	$\boldsymbol{Y_t}$	$\boldsymbol{X_{1t}}$	$\boldsymbol{X_{2t}}$
1	10	1	9
2	6	2	4
3	9	3	7
4	7	4	2
5	11	5	6
6	15	6	9

(a) Contrastar la hipótesis nula Ho: $\boldsymbol{\beta_1 + \beta_2}$**=1**

(b) Incluya el resultado de la hipótesis nula y estime el modelo restringido

***Sugerencias*:**
(a) Estimar, mediante software, los coeficientes para el modelo no-restringido. Entonces especifique y estime el modelo restringido por la hipótesis nula. Replantee primero los datos de las variables transformadas (del modelo restringido).

(b) Utilice F en el marco de ambos modelos, restringido y no-restringido, a fin de probar la validez de la hipótesis nula.

3.9. Considere el siguiente modelo (tomado de Wooldridge, ejercicio C5.3, pp.182), que explica el peso de un niño al nacer (***W***, en Kg.) en términos de 5 variables regresoras:

$$W = \beta_0 + \beta_1 CIG + \beta_2 PARITY + \beta_3 FAMINC + \beta_4 MOTHE + \beta_2 FATHE + \mu_t$$

Donde:
CIG: promedio diario de cigarrillos que fumo la madre durante el embarazo
PARITY: orden de nacimiento del niño
FAMINC: ingreso anual de la familia
MOTHEDUC: años de escolaridad de la madre
FATHEDUC: años de escolaridad de la madre.

Calcule el estadístico ML (multiplicador de Lagrange) para probar si las variables *MOTHEDUC* y *FATHEDUC* son conjuntamente significativas.

***Sugerencias*:**
Al obtener los residuales del modelo restringido, poner cuidado con que se empleen solo aquellas observaciones para las que estén disponibles todas las variables en el modelo no restringido.

Primero ejecutar la regresión de los determinantes del peso (W) en función de CIG, PARITY, FAMINC, MOTHEDUC, FATHEDUC. Luego se obtienen los residuales con 1,197 observaciones y se sigue el procedimiento de uso del estadístico LM; el valor de Ji-cuadrado será 2.86, para efectos de la comparación respectiva.

3.10. Para el caso norteamericano (USA), período 1980-1998, se estimó el siguiente modelo:
(tomado de Gujarati y Porter, 2010, ejercicios 8.29 y 7.21).

$M2_t = \beta_0 Y_t^{\beta_1} r_t^{\beta_2} e^{\mu_t}$ (1)

Donde
*M*2***:*** *demanda real de dinero*
Y: PIB real
r: tasa de interés de corto plazo.

(a) Haga uso del fichero ***Table7_10***, para estimar los parámetros del modelo, incluyendo la tasa de interés de corto plazo. Tome nota que para la conversión de cantidades nominales a reales se divide *M* y *PIB* entre el índice de precios al consumidor (IPC), excepto la tasa de interés "*r*". En el fichero los nombres de las variables independientes son: RGDP (PIB real) y TBRATE (tasa de interés de corto plazo).

Suponga que se hubiera decidido utilizar la función:
$\frac{M}{Y} = \alpha_0 \, r_t^{\alpha_1} e^{\mu_t}$ (2)

(b) ¿Cuáles serían los resultados, y cómo los interpretaría?

(c) De las dos especificaciones previas, ¿por cuál se decidiría y por qué?

***Sugerencias*:**
Utilice software relevante para determinar:
(*a*) *Con un modelo logarítmico, las elasticidades resultarán en 0.5243 y -0.0516, para ingreso y tasa de interés de CP, respectivamente.*
(*b*) *La relación* $\frac{\boldsymbol{M}}{\boldsymbol{Y}}$ *representa la proporción de ingreso que las personas desean mantener en forma "líquida". Los resultados son: Ln* $(\frac{\boldsymbol{M2}}{\boldsymbol{Y}}) = 3.4785 - 0.172(lnTBRATE)$
(c) *Usar el estadístico "F" para modelos restringidos: la primera ecuación es no-restringida y la segunda restringida.*

CAPÍTULO IV. ESPECIFICACIÓN Y DIAGNÓSTICO DEL MODELO

Uno de los supuestos más importantes del Modelo de Regresión Lineal, a menudo subestimado, es el de la "Correcta Especificación del Modelo" (supuesto 9, según Gujarati & Porter, 2010); es decir, una vez que se asume que el modelo está correctamente especificado, las pruebas y respectiva estimación resultan relativamente sencillas. Sin embargo, en realidad nunca podemos estar totalmente seguros que un determinado modelo está especificado en forma correcta; los investigadores normalmente examinan más de una posible especificación, intentando el encuentro de la ecuación que mejor represente el proceso bajo estudio. Y el problema es que un modelo mal especificado tiene consecuencias en cuanto a la violación de otros supuestos importantes del Modelo de Regresión Lineal, como son los casos de ocurrencia "impura" de colinealidad, heteroscedasticidad, autocorrelación; es decir problemas originados en una mala especificación del modelo.

En esta sección el propósito es expresar dificultades importantes asociadas a una adecuada especificación del modelo, así como los potenciales peligros implicados en la búsqueda de un modelo bien especificado, exponiendo los costos asociados con la mala especificación del mismo. Algunos tipos importantes de incorrecta especificación del modelo, pueden agruparse fundamentalmente en dos categorías: error de especificación en la selección de variables del modelo (sobretodos las variables explicativas), y error de especificación en el uso de la forma funcional más adecuada. Es importante indicar también que la detección es un elemento clave, más allá de ello lo demás requiere de empirismo; es decir, una vez detectado el error de especificación, entonces los arreglos son casi automáticos.

4.1. Error en la Selección de Variables explicativas

Un modelo sufre de especificación incorrecta cuando no explica de manera adecuada la relación existente entre la variable dependiente y las variables explicativas observadas. Existen varios casos al respecto.

Caso 1. Modelo sub-especificado
Supongamos el verdadero modelo es:

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \mu_i \qquad (4.1)$$

Pero si en la práctica usamos el modelo:

$$Y_i = \alpha_0 + \alpha_1 X_{1i} + v_i \qquad (4.2)$$

Tenemos problemas importantes, que se expresan del siguiente modo:

i. Si las variables X_{1i} y X_{2i} están correlacionadas ($r_{12} > 0$, altamente probable), entonces los estimadores $\hat{\alpha}_0$ y $\hat{\alpha}_1$, del modelo 4.2, son sesgados e inconsistentes, pues el término de error estaría correlacionado con la variable explícita del modelo 4.1, violando supuesto importante del Modelo de regresión lineal general.

ii. Los estimadores $\hat{\alpha}_0$ y $\hat{\alpha}_1$ son ineficientes. Ejemplo, el estimador de la varianza del estimador, $S(\hat{\alpha}_1)$, resulta ser un estimador sesgado de la verdadera varianza: $V(\hat{\beta}_1)$.

Por tanto, en este caso de una o más variables regresoras omitidas, los estimadores no solo son puntualmente incorrectos, sino que es bastante probable que los respectivos intervalos de confianza (IC) y los procedimientos de contraste de hipótesis conduzcan a conclusiones erróneas. Igual, los pronósticos basados en el modelo incorrecto y los IC del pronóstico no son confiables. Una vez formulado el modelo en base a la teoría, no es aconsejable eliminar una variable importante del modelo (no "tan fácilmente").

Caso 2. Modelo sobre-especificado

Supongamos que el verdadero modelo es:

$$Y_i = \beta_0 + \beta_1 X_{1i} + \mu_i \qquad (4.3)$$

Pero si en la práctica usamos el modelo:

$$Y_i = \alpha_0 + \alpha_1 X_{1i} + \alpha_2 X_{2i} + v_i \qquad (4.4)$$

Entonces tenemos lo siguiente:

i. Todos los estimadores MCO ($\hat{\alpha}_s$) del modelo (4.4) serían insesgados y consistentes, pero el problema estaría en que las estimaciones $\hat{\alpha}_s$ son ineficientes: es decir "sus varianzas tienden a ser más grandes que las $V(\hat{\beta})$ correctas" (detalles en la sección Apéndice 13.A.2 del libro de Gujarati & Porter, 2010). Esto conlleva finalmente a un enorme perjuicio en lo que se refiere al proceso de inferencia estadística: las pruebas de hipótesis y la construcción de los Intervalos de Confianza resultan ser probablemente defectuosas.

ii. Si es que son "muchas" las variables inadecuadamente incluidas en el modelo incorrectamente sobre-especificado, también incrementaría el riesgo de multicolinealidad (MC) y la pérdida de grados de libertad.

4.2. Error de Especificación en la Forma Funcional

Caso 3. Forma funcional incorrecta

Una forma funcional no es más que una ecuación que establece una relación entre variables; en el caso de Econometría, para este libro, son una o más variables endógenas y una o más variables exógenas o independientes. Puede darse el caso que las variables incluidas en el modelo sean las adecuadas, pero la forma funcional que las relaciona sea incorrecta[12].

Supongamos que la correcta relación entre una variable dependiente *Y* y dos variables explicativas es la siguiente:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \gamma_1 X_1^2 + \gamma_2 X_2^2 + \delta X_1 X_2 + \mu_i \qquad (4.5)$$

Pero en la práctica se usa la ecuación siguiente:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \mu_i \qquad (4.6)$$

La ecuación (4.5) muestra una respuesta cuadrática (de *Y*) en relación a las variables regresoras, además de un efecto de interacción entre ambas variables independientes. Por tanto, el verdadero efecto esperado real de un cambio en X_1 será:

$\frac{\partial Y}{\partial X_1} = \beta_1 + 2\gamma_1 + \delta X_2$ (la pendiente –real- depende de dos parámetros y de la variable X_2)

Al usar la forma funcional incorrecta (4.6), se tendría la siguiente pendiente:

$\frac{\partial Y}{\partial X_1} = \beta_1$ (La pendiente en este caso sería constante: se subestima el valor de la real pendiente).

[12] El caso de una variable relevante omitida podría también verse como caso particular de una forma funcional incorrecta.

4.3. Detección de Errores de Especificación[13]

Tanto en el caso de omisión o exceso de variables relevantes, como en el caso que se refiere a un inadecuado uso de la forma funcional, es relevante la DETECCIÓN del error de especificación. Una vez detectado el problema, los arreglos serían casi automáticos y se basan fundamentalmente en la experiencia del investigador.

4.3.1. Pruebas preliminares

Funcionan mejor en los casos de un modelo sub-especificado, por ausencia explicita de una variable importante, así como en el caso del uso de una forma funcional incorrecta. Para detectar problemas de la forma funcional inadecuada suelen utilizarse los gráficos de los *residuos,* que ante la presencia de no linealidades, normalmente presentan pronunciadas tendencias que indican falta de aleatoriedad; ver por ejemplo tres casos presentados en la Figura 4.1, en que el caso (a) tiene menor aleatoriedad que (b) y el caso (b) menor aleatoriedad que (c).

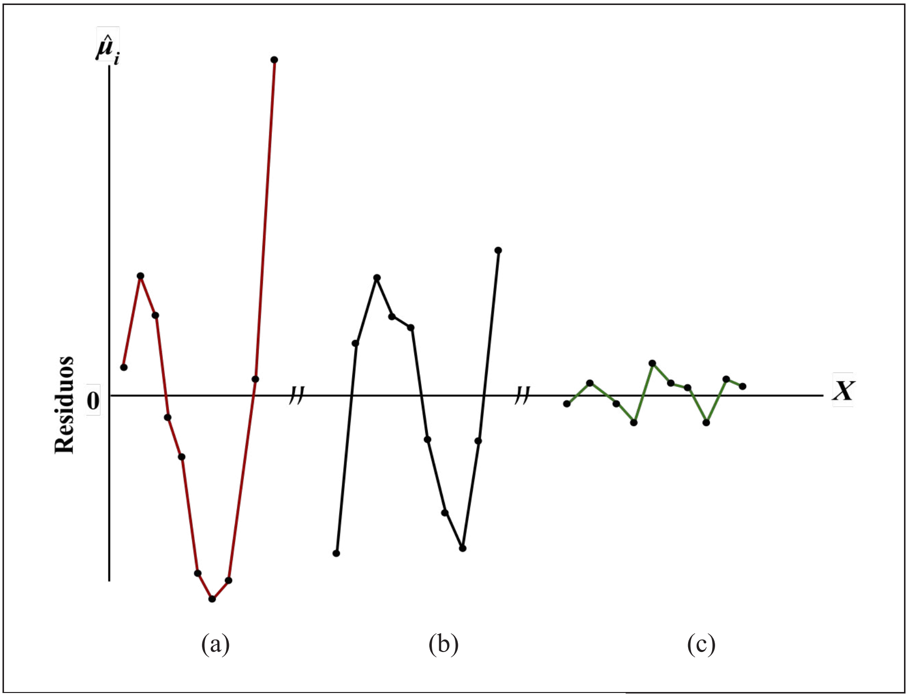


Figura 4.1. Diferentes casos de inadecuada especificación

[13] Detección en casos asintóticos, se pueden manejar mediante una de las tres opciones: Wald (F Fisher)- Razón de Verosimilitud o Multiplicador de Lagrange.

También los altos grados de autocorrelación (AC) podrían ser indicadores de un problema con la Forma Funcional; una forma sencilla de notarlo es fijándose en el estadístico Durbin-Watson (DW). Si este estadístico DW revela AC (o "indecisión") entonces podría haber un problema de mala especificación. En este caso la AC observada se dice que es de origen "impuro"[14] y reflejaría el hecho que existen una o más variables pertenecientes al modelo (el "lineal" por ejemplo) que están incluidas en el término de error μ_i y que necesitan ser "excluidas" del mismo a fin de introducirse explícitamente en la ecuación original: $Y = f(X) + \mu_i$.

En resumen, si especificamos un modelo sin problemas de multicolinealidad, ni heteroscedasticidad, pero con un R^2 no-convincente, o muy pocos coeficientes son significativos mientras si lo es el estadístico *F* conjunto, o no se tienen los signos correctos de los coeficientes, entonces miramos el valor DW y los residuos, para, en base a ello, referir la potencial existencia de un problema de autocorrelación asociado a una inadecuada especificación del modelo.

4.3.2. Pruebas formales

Prueba de error de especificación en regresión (RESET)[15]

Esta prueba fue desarrollada por J. B. Ramsey; parte de un modelo original correcto (4.7), que satisface el supuesto $E\ (\mu_i/\mathrm{X}) = 0$, pero que se presume que podría tener problemas en la Forma Funcional.

$$Y_i = \beta_0 + \beta_1 X_{1i} + \cdots + \beta_k X_{ki} + \mu_i \qquad (4.7)$$

Entonces considerando (4.7) como un modelo no restringido, Ramsey propuso estimar la ecuación (4.7) mediante MCO, para luego especificar una ecuación instrumental (4.8) que incorpora polinomios de los valores ajustados de $\hat{Y}_i$. No existe respuesta exacta sobre el número de términos a añadir, pero solo dos términos: cuadráticos y cúbicos han mostrado ser útiles en la mayoría de aplicaciones.

Con la incorporación de los polinomios, se obtiene lo siguiente:

$$Y_i = \beta_0 + \beta_1 X_{1i} + \cdots + \beta_k X_{ki} + \ \gamma_1 \boldsymbol{\hat{Y}_1^2} + \gamma_2 \boldsymbol{\hat{Y}_2^3} + error \qquad (4.8)$$

Donde $\hat{Y}_i$ es una serie de datos "ajustados", producto de la estimación de (4.7).

---

[14] Es decir, no está basada en la propia naturaleza de las variables involucradas en el modelo, sino en una inadecuada especificación del mismo.
[15] RESET es una abreviatura que proviene del nombre de la prueba en inglés, pero que en español significa: "prueba de error de especificación de regresión".

La ecuación (4.8) es instrumental**,** correspondería a un modelo no-restringido y sirve para determinar la existencia del error de especificación en el modelo original. No interesan los estimadores per se, sino la utilidad para el contraste de las hipótesis siguientes:

Ho: La ecuación (4.7) está mal especificada; frente a
Ha: No hay evidencia de mala especificación en (4.7).

También la propuesta de hipótesis podría plantearse como:
Ho: ambos parámetros $\gamma_1 \; y \; \gamma_2 = 0$ en la ecuación (4.7) frente a
Ha: $\gamma_1 \; o \; \gamma_2 \neq 0$ en la ecuación (4.7)

La prueba RESET utiliza la técnica de Mínimos Cuadrados restringidos (MCR) y un estadístico ***F*** que, si resulta significativo, nos permite rechazar la hipótesis nula y concluir que el modelo está bien especificado.

$$\boldsymbol{F} = \frac{(R_{nr}^2 - R_r^2)/(N° \; de \; nuevas \; regresoras)}{(1 - R_r^2)/(n - parámetros \; del \; modelo)}$$

Donde:
R_{nr}^2: coeficiente de determinación del modelo no restringido
R_r^2: coeficiente de determinación del modelo restringido
n: total de observaciones.

Partiendo del modelo no restringido, la idea es que si el incremento en R^2 es estadísticamente significativo, entonces *F es significativo* y rechazamos la hipótesis nula.

Dado que $\hat{Y}_i$ es también una variable aleatoria, las pruebas de significancia habituales aplican si la muestra es razonablemente grande.

Ejemplo 4.1. A fin de mostrar el procedimiento de la prueba RESET, considere la siguiente función de costo lineal:

$$Y_i = \beta_0 + \beta_1 X_i + \mu_i \qquad (4.9)$$

Nuestro interés es corroborar las siguientes hipótesis:

Ho: La ecuación (4.9) está mal especificada; frente a
Ha: No hay evidencia de mala especificación en (4.9).
El procedimiento es el siguiente:

Paso 1. Ejecutar la ecuación restringida y obtener $\widehat{\boldsymbol{Y}}_{\boldsymbol{i}}$**:**

$\widehat{\boldsymbol{Y}}_{\boldsymbol{i}}= \ 166.5 \ + \ 19.93X_i$
EE= (19.02) (3.07)
R2=0.841

Paso 2. Volver a estimar la ecuación original (4.9), pero incluyendo polinomios de $\widehat{\boldsymbol{Y}}_{\boldsymbol{i}}$ como una o más variables regresoras.

$\widehat{\boldsymbol{Y}}_{\boldsymbol{i}} = 2140.7 + 476.6X_{1i} - 0.092\widehat{\boldsymbol{Y}}_{\boldsymbol{1}}^{\boldsymbol{2}} + 0.00011\widehat{\boldsymbol{Y}}_{\boldsymbol{2}}^{\boldsymbol{3}}$ (4.10)
R2=0.998
Paso 3. Siendo la ecuación (4.10) la "no restringida", y la (4.9) la restringida, entonces:

$$F = \frac{(R_{nr}^2 - R_r^2)/(N° \ de \ nuevas \ regresoras)}{(1 - R_r^2)/(n - parámetros \ en \ el \ modelo \ no \ restringido)}$$

$$F = \frac{(0.998-0.841)/(2)}{(1-0.841)/(n-4)} = 284.40$$

Paso 4. Si el valor $\boldsymbol{F}$ calculado es significativo, con 5% como margen de error, podemos no hay suficiente evidencia para rechazar la hipótesis nula. Ho, sería como aceptar la hipótesis que el modelo restringido estaría mal especificado. En este caso, con F=284.4, altamente significativo, tenemos suficiente evidencia para rechazar la hipótesis nula y concluir que la hipótesis que el modelo lineal de costos está mal especificado.

4.4. Errores de Medición en las Variables

Hasta ahora hemos supuesto que todas las variables, independientes y dependiente, son medidas sin error. En la práctica es probable que ocurran errores de medición que conlleven a potenciales problemas de los estimadores obtenidos mediante mínimos cuadrados ordinarios (MCO).

Aunque en rigor los errores de medición son elementos "externos" al propio modelo, pues ocurren en el momento de obtener la información de las variables ya especificadas, a menudo son también considerados como errores de especificación de los modelos. En esta sección nos concentraremos en evaluar las características de los errores en la medición, su detección y eventuales alternativas de solución a los problemas ocasionados.

4.4.1. Casos de errores de medición

Caso I. Error en la variable dependiente

Suponga que el verdadero modelo, que satisface los supuestos de Gauss-Markov, es:

$$Y = \beta_1 X_i + \mu \qquad (4.11)$$

Pero si se ha originado error de medición en la variable, que conlleva a usar la variable Y^*, en lugar de Y, donde: $Y^* = Y + \varepsilon$

Agregando el error de medición $\boldsymbol{\varepsilon}$ en cada lado de la ecuación (1) obtenemos:

$$Y^* = \beta_1 X_i + (\mu + \varepsilon) \qquad (4.12)$$

Se puede concluir que siempre que la $Cov\,(\varepsilon, X_i) = 0$, entonces no habría mayor problema: el estimador de β_1 será insesgado y consistente, pero siempre el efecto secundario de error de medición en la variable dependiente es el incremento en la varianza del error. Tal incremento en la varianza del error generaría un problema en cuanto al proceso de prueba de las hipótesis y al establecimiento de los Intervalos de Confianza.

Caso II. Error de medición en la(s) variable(s) independiente(s)

Partimos de un modelo (4.11) que satisface los supuestos de Gauss-Markov, permitiendo que los estimadores MCO de los parámetros sean insesgados y consistentes:

$$Y = \beta_0 + \beta_1 X_1 + \mu \qquad (4.13)$$

Pero si X_1 no es realmente observada y lo que se tiene es una medición mediante X_1^*, que es denominada "la variable reportada", se tendría entonces:

$X_1^* = X_1 + w$ (error de medición: $\boldsymbol{w}$)

Entonces lo que estimamos realmente es:

$$Y = \beta_0 + \beta_1 (X_1^*) + \mu^* \qquad (4.14)$$

$$Y = \beta_0 + \beta_1 (X_1) + (\mu - \beta_1 w)$$

$$Y = \beta_0 + \beta_1 (X_1) + \boldsymbol{\varepsilon} \qquad (4.15)$$

Siendo $\varepsilon = \mu - \beta_1 w$

Aun suponiendo que $\boldsymbol{w}$ ~*Normal* con media 0, y es independiente del error en la ecuación verdadera, en presencia de un error de medición surgen problemas al usar MCO.

Al respecto, es posible demostrar y notar que:

$$Cov(\mu^*, X_1^*) = -\beta_1\sigma_w^2$$

También es posible demostrar que $pL\text{im}(\hat{\beta}_1) = \beta_1\left[\frac{\sigma_x^2}{\sigma_x^2+\sigma_w^2}\right]$

Y las cosas son más complicadas cuando se agregan más variables regresoras, pues confluyen también con errores en las variables dependientes.

En el caso de errores de medición en las variables regresoras, se tendría un caso particular de la presencia de variables *X* endógenas en el modelo; en tal caso, el término de error de los modelos lineales generales y las variables regresoras estarían correlacionados, lo que viola un supuesto básico de independencia entre variables regresoras y el término de error, conllevando estimadores inconsistentes. Ergo: los errores de medición constituyen un grave problema cuando están presentes en las variables explicativas sobretodo.

Es casi imposible saber si una variable tiene error de medición, la única posibilidad "a priori" es una sospecha. En tal caso sería importante tener o precisar una variable instrumental, avizorando un arreglo del problema, pues con este instrumento será posible mitigar el efecto de errores de medición.

4.4.2. Procedimiento de detección y corrección

(i) Variables instrumentales (VI)

Supongamos, en modo matricial, que tenemos un conjunto de variables explicativas $\boldsymbol{X}$ que son consideradas fijas, de modo tal que: $E(\mu/X) = E(\mu) = 0$; también expresado, en forma matricial, como $E(UU'/X) = E(UU')$, por lo que los estimadores MCO son insesgados y consistentes. Pero si las variables regresoras $\boldsymbol{X}$ fueran aleatorias y no son independientes del error μ, entonces es mejor usar variables instrumentales (VI) en reemplazo de las variables en $\boldsymbol{X}$ que presentan tal dificultad. Una VI se define entonces como una variable que no aparece en la ecuación original, no está relacionada con el error de la ecuación y se correlaciona (parcialmente) con la variable regresora endógena.

Dependiendo de cuantas variables regresoras ocasionen problemas, el procedimiento de detección y corrección comienza con la búsqueda de una o más VI; implica esencialmente sustituir las variables "problemáticas" del vector X por instrumentos muy correlacionados con ellas, pero no-correlacionados con el término de error. Luego se aplica el método de mínimos cuadrados ordinarios. Eventualmente, podría no ser sencillo el procedimiento cuando no es fácil encontrar una o más buenas VI. Acá la teoría podría ayudar con tal "encuentro".

El uso de una correcta variable instrumental permite la obtención de estimadores consistentes. Para corroborar la pertinencia de la VI se puede usar el Contraste de Hausman, también denominado "*Contraste de endogeneidad*" porque, en general, establece la naturaleza de una variable regresora. En esta sección interesa el caso particular de la presencia de error de medición en una determinada variable regresora, que invalidaría el uso de tal regresora y conllevaría a un caso particular de endogeneidad[16]. En el uso de una adecuada VI, ¿puede ayudar el uso del coeficiente de correlación parcial o total, entre la variable regresora con problemas de medición y una potencial variable instrumental? Es una buena posibilidad.

(ii) Contraste de Hausman

Es el constraste más usual para evaluar la endogeneidad de las Variables Regresoras. En este caso particular se presume que la endogeneidad de una (o más) variable(s) regresora(s) se debe a un error ocasionado en la medición de la variable regresora, lo que ocasiona finalmente su correlación con el término de error.

El proceso involucra las siguientes pruebas de hipótesis:
Ho: la variable regresora X no es estocástica
Ha: la variable regresora X es estocástica y tendría correlación con el error.

El procedimiento de prueba permite comparar los estimadores ***MCO*** de los parámetros del modelo con la (o las) variable dependiente "reportada" contra lo obtenido con la (o las) variables instrumentales. Bajo la hipótesis nula los estimadores son insesgados y consistentes, mientras que bajo rechazo de la hipótesis nula, solo el estimador asociado a la VI mantiene tal propiedad. Entonces el principio es que cuanto mayor es la diferencia entre las estimaciones de parámetros por ambos métodos, mayor evidencia de la presencia de regresoras estocásticas o la presencia de error de medida en las variables regresoras.

[16] Según Pyndick (pp. 168), "una fuente de falla en la violación del supuesto de independencia entre regresoras y residual, es el *error de medición* en una o más variables independientes".

En la práctica, ¿cómo es el procedimiento del contraste Hausman? Para probar las hipótesis antes indicadas, se usan regresiones auxiliares. Suponga que el modelo original es:

$$Y = \beta_0 + \beta_1(X_i) + \mu \qquad (4.16)$$

Y queremos evaluar el "carácter estocástico" de la variable explicativa X_i. El procedimiento incluye:

(i) Buscar una Variable Instrumental (VI) apropiada (Z_i^*), pre-establecida por la teoría o por empirismo, que reemplace a X_i.

(ii) Se usa MCO para estimar una nueva ecuación auxiliar (4.17) y obtener residuales $\widehat{w}_i$:

$$X_i = \alpha_0 + \alpha_1(VI_i) + w_i \qquad (4.17)$$

Y se obtiene la serie: $\widehat{w}_i = X_i - \widehat{\alpha_0} + \hat{\alpha}_1(VI_i)$

(iii) Se ejecuta una "nueva" regresión:

$$Y = \beta_0 + \beta_1(X_i) + \delta(\widehat{w}) + \varepsilon \qquad (4.18)$$

La decisión final depende del coeficiente $\hat{\delta}$: si $\hat{\delta}$ es estadísticamente significativo, entonces se rechaza la Ho, concluyendo que hay un problema relacionado con la endogeneidad de la variable regresora X_i, por lo que los estimadores podrían resultar sesgados e inconsistentes.

Frente al resultado de rechazo de la hipótesis nula Ho, en lugar de la ecuación (4.16), una mejor opción de estimación es:

$$Y = \beta_0 + \beta_1(VI_1) + \mu \qquad (4.19)$$

En cambio, si no existiría evidencia suficiente para rechazar la Hipótesis nula, no habría mayor problema en la estimación de la ecuación original (4.14): $Y = \beta_0 + \beta_1(X_i) + \mu$

Ejemplo 4.2. Gasto Público de Gobiernos Estatales (adaptado del ejemplo 7.3, pp. 206-207 de Pyndick y Rubinfeld, 2000). Suponiendo que, en los Estados Unidos de América, los gastos de los 50 gobiernos estatales varían por Estado, en función a las siguientes variables determinantes: subsidios para ayuda económica (*AID*), el ingreso presupuestal que reciben (*INC*) y la población (*POP*) en la región, se ha especificado el siguiente modelo:

$EXP = \beta_1 + \beta_2(AID) + \beta_3(INC) + \beta_4(POP) + \varepsilon_t$

Usando MCO, se obtuvo los siguientes resultados:

$$\widehat{EXP} = \quad -46.8 \ + 0.0032(AID) + 0.0019(INC) \ - 0.597(POP) \qquad (4.20)$$

Est. "t" (- 0.56) (13.64) (8.12) (- 5.710)

R^2=0.993, F=2190

Intuimos que hay una posible fuente de error de medición en la variable *AID, por tanto existiría un problema de endogeneidad.* Decidimos entonces realizar la prueba Hausman a fin de medir la existencia de *error en la medición* de AID. Para ello usamos una Variable Instrumental denominada *Población de niños de escuelas primarias y secundarias (PS).*

Paso 1. Se realiza la regresión de la V. Regresora en cuestión (*AID*) con respecto a la variable instrumental *PS,* y se obtiene el residual estimado:

$$\hat{w} = AID - 77.95 + 0.845(PS) \qquad R2=0.87$$

Paso 2. Se añade el residual estimado en la ecuación original y así podemos establecer la endogeneidad (y, de ser el caso, "corregir" el problema):

$$\widehat{EXP} = \ -138.5 + 0.00174(AID) + 0.0018(INC) - 0.518(POP) + 1.372(\hat{w}) \qquad (4.21)$$

Est."t" (- 1.41) (1.94) (7.55) (- 1.29) (1.73)

R2=0.99

La decisión final depende de la significancia estadística asociada al coeficiente de $\hat{w}$. Según la prueba "t" el coeficiente de $\hat{w}$ se obtuvo en 1.73 < 1.96 (tt), entonces no podemos rechazar la Ho y concluimos que no hay problemas de endogeneidad: se podría usar MCO con la ecuación (4.20) original.

Pero si (por ejemplo) "t " hubiera resultado en 2.2, entonces rechazamos la Ho y aceptaríamos que hay "error de medición"; la última ecuación (4.19) (incluyendo $\hat{w}$) sería una ecuación que habría corregido tal "error de medición".

4.5. Criterios de Elección entre Modelos Alternativos

La "Bondad de Ajuste" puede ser entendida como lo "bien" que los datos muestrales se ajustan a la línea de regresión estimada o ajustada. Un indicador que mayoritariamente es utilizado, en cuanto a la bondad de ajuste, es el coeficiente de determinación[17].

En términos de una correcta especificación del modelo, en ocasiones se requiere seleccionar una ecuación, entre dos o más alternativas, cuya especificación se aproxime bien a los datos muestrales. Esta elección se puede realizar utilizando uno o más indicadores que son también denominados "indicadores de bondad de ajuste". Es lo que se presenta a continuación.

4.5.1. Coeficiente de determinación

El coeficiente $\boldsymbol{R^2}$ es el que mide el nivel de ajuste del modelo que se ha estimado, es decir, evalúa si las variables regresoras explican adecuadamente la variable dependiente. Si, por ejemplo, el coeficiente de determinación fuera 0.90 significa que la variación de la variable dependiente es explicada por las variables regresoras en un 90%; el 10% restante es explicado por efectos agrupados en el residuo. Nótese que en la formula mostrada a continuación, el coeficiente $\boldsymbol{R^2}$ depende de la Suma de cuadrados del Residuo (***SCR***) y la Suma de cuadrados Totales (***SCT***).

$$R^2 = 1 - \frac{SCR}{SCT}$$

Es importante indicar que este estadístico de bondad de ajuste presenta algunas limitaciones: (i) mide bien la bondad de ajuste en cuanto a la predicción dentro de la muestra, pero no ofrece la misma garantía en el caso de predicciones fuera de la muestra. (ii) el R^2 su valor a medida que se añaden más variables explicativas al modelo, independientemente de la significatividad o importancia de las variables regresoras añadidas, por lo que siempre en la comparación de dos valores, se podría estar seleccionado al que tenga más variables explicativas.

4.5.2. Coeficiente de determinación ajustado

Debido a la primera limitación indicada, en la práctica se usa el coeficiente **R^2 ajustado**, que se define de modo tal que en su fórmula "penaliza" la inclusión de nuevas variables explicativas en el modelo: si bien al aumentar el número de regresores aumenta también la Suma de Cuadrados Explicados, la inclusión de nuevas variables explicativas reduce los grados de libertad del modelo, por lo que no siempre resultará adecuado incorporar nuevas variables al mismo, todo dependerá de la significancia que ellas tengan. El $\boldsymbol{R^2}$ ajustado es un indicador que penaliza cuando se agregan más variables regresoras, dependiendo de su importancia en la explicación de

---

[17] Algunos econometristas indican que el coeficiente de determinación R^2 es un indicador de importancia relativa y complementaria en un modelo, lo más importante- sugieren- es la significancia y eficiencia de los indicadores que estiman a los parámetros.

Y (si las variables regresoras se asocian a estadísticos "t" mayores a 1, entonces aportan al modelo, caso contrario, no[18]). representado por $\bar{\boldsymbol{R}}^{\mathbf{2}}$) es explicado mediante la siguiente fórmula:

$$\bar{R}^2 = 1 - \frac{n-1}{n-k}(1 - R^2) \qquad (4.22)$$

Es importante indicar que, el intento de comparar dos modelos, en base a los respectivos coeficientes de determinación, requiere que el tamaño de la muestra y la variable dependiente deban ser los mismos; las variables regresoras pueden adoptar cualquier forma. Por ejemplo, en el caso de un modelo cuadrático y otro doble-logaritmico no sería posible la comparación de "mejor ajuste" de un modelo sobre otro, en cuanto al uso del R^2, dado que la variable dependiente es diferente. Lo mismo sucedería en el caso de los siguientes dos modelos:

$$Y_i = \beta_0 + \beta_1 X_i + \beta_2 X_i^2 + \beta_3 X_i^2 + \mu_i$$

$$Ln\, Y_i = \beta_0 + \beta_1 X_i + \beta_2 X_i^2$$

La comparación sería posible solo después de haber realizado la transformación correspondiente en la variable dependiente, de modo tal que las variables dependientes de ambos modelos sean equivalentes y, por tanto, el valor R^2 sea posible de comparar. Los pasos son básicamente dos:

(i) Llevar la variable dependiente logarítmica (2ª ecuación) a una transformación lineal, mediante el uso de antilogaritmos que permita construir una 2ª opción serial con el antilogaritmo:
$Z_i = EXP\ (Ln\ Y_i)$

(ii) Utilizar la fórmula del coeficiente de determinación, utilizando la nueva serie Z_i y el valor real de la ecuación 2ª en la fórmula: $R^2 = \frac{(\sum \hat{Y}_i Y_i)^2}{\sum \hat{Y}_i^2 \sum Y_i^2}$

(iii) Utilizar la ecuación (4.20) para obtener el coeficiente ajustado. Entonces recién comparar los valores de R^2 de ambas ecuaciones, a fin de determinar el modelo con forma funcional adecuada.

4.5.3. Criterio Akaike (AIC)

Es un estadístico que mide el buen ajuste de la data a la regresión estimada, permitiendo la selección entre dos modelos de ajuste alternativos. También penaliza la inclusión de nuevas variables regresores en el modelo, seleccionando como modelo

[18] Nota 5 a pie de página de Pyndick y Rubinfeld, pp. 93.

más adecuado aquel que presenta un menor valor de dicho coeficiente, en el marco de una serie de números reales. Su fórmula de cálculo responde a la siguiente expresión:

$$AIC = LN\left(\frac{SCR}{n}\right) + \frac{2k}{n}$$

Su valor también aparece calculado automáticamente en la estimación facilitada por cualquier programa de estimación econométrica (software). El coeficiente Akaike es útil para predicción dentro y fuera de la muestra.

4.5.4. Criterio Schwarz (BIC)

Es una alternativa que penaliza en un grado mayor (que AIC y R2) la inclusión de nuevas variables regresoras en el modelo. Al igual que en el caso anterior, se considera mejor modelo aquel que presenta un menor valor del coeficiente, en el marco de evaluación en una serie de números reales. Su valor aparece también calculado en la estimación facilitada por cualquier software econométrico utilizado. El criterio Schwarz sirve para comparaciones dentro y fuera de la muestra, y se obtiene a partir de la siguiente expresión.

$$BIC = \frac{k}{n} ln(n) + LN\left(\frac{SCR}{n}\right)$$

Comparación de modelos alternativos

Para poder comparar modelos según los criterios AIC y BIC, es requisito obligatorio que las estimaciones a comparar tengan la misma variable dependiente y que adicionalmente tengan el mismo tamaño de muestra. Un modelo es mejor especificado que otro si es que ambos criterios (AKAIKE y SCHWARZ) presentan los "menores" valores comparativos. Sin embargo, como fue presentado previamente, el coeficiente de determinación R^2 permite la comparación entre modelos, después de haber "igualado" la variable dependiente.

En resumen:

Criterio	Indicador	Comentario
R^2 ajustada[19]	$\bar{R}^2$=1- $\frac{SCE/(n-k)}{SCT/(n-1)}$	Penaliza cuando se agregan más variables regresoras.

[19] Se sabe además que: $R^2 = \frac{SCE}{SCT} = 1 - \frac{SCR}{SCT}$

Akaike (CA)	$LnCA = (\frac{2k}{n}) + Ln(\frac{SCR}{n})$	Impone mayor "penalización" que R2, por añadir regresoras.
Schwarz (CS)	$LnCS = \left(\frac{k}{n}\right) Ln(n) + Ln(\frac{SCR}{n})$	Impone mayor "penalización" que CS, por añadir regresoras.

Donde:
$SCR = \sum \hat{\mu}_i^2$
n = tamaño de muestra
k = número de parámetros

4.5.5. Comparación de coeficientes de determinación (R^2)

En cuanto a la bondad de ajuste, en general la comparación entre modelos requiere que el tamaño de la muestra (n) y la variable dependiente deban ser los mismos. En tal sentido, dos modelos que no podrían ser comparados serían los siguientes:

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \beta_3 X_{3i} + \mu_i \qquad (4.23)$$

$$LnY_i = \alpha_0 + \alpha_1 LnX_{1i} + \alpha_2 X_{2i} + \alpha_3 X_{3i} + \mu_i \qquad (4.24)$$

Desde el punto de vista del coeficiente de determinación, la no-comparación de los modelos (4.23) y (4.24) se fundamenta en el hecho que por definición el coeficiente R^2 mide la proporción de variación en la variable dependiente explicada por las variables regresoras. Por tanto, en la ecuación (4.23) se tiene:

$$1 - R^2 = \frac{\Sigma(\mu_i)^2}{(Y_i - \bar{Y})^2} \qquad (i)$$

Mientras que en la ecuación (4.24), similar expresión sería:

$$1 - R^2 = \frac{\Sigma(\mu_i)^2}{(LnY_i - \overline{LnY})^2} \qquad (ii)$$

El resultado casi obvio es que los resultados de (i) y (ii) no son comparables. Como vimos antes, en la ecuación (4.23) una modificación en $\boldsymbol{Y_i}$ indica un cambio absoluto en la variable dependiente, mientras que en (4.24) una modificación en $\boldsymbol{Y_i}$ indica un cambio relativo en la variable dependiente.

Sin embargo, Gujarati y Porter (2010) proponen una forma de comparación entre modelos, usando el coeficiente de determinación $\mathbf{R^2}$, después de una transformación en la variable dependiente, procedimiento que se puede ilustrar del siguiente modo: suponga que se tienen dos modelos simples ajustados, no comparables:

$$\widehat{Y_i} = \hat{\beta}_0 + \hat{\beta}_1 X_{1i} \qquad (4.25)$$

$$\widehat{LnY_i} = \hat{\alpha}_0 + \hat{\alpha}_1 LnX_{1i} \qquad (4.26)$$

Entonces el procedimiento es el siguiente:
(i) De la ecuación (2) tomar el antilogaritmo de la serie $\widehat{LnY_i}$ y calcular el coeficiente de correlación:

$$r = \frac{(\sum Y_i \hat{Y}_i)}{\sqrt{(\sum Y_i^2)(\sum \hat{Y}_i^2)}}$$

(ii) Estimar el coeficiente de determinación "comparable" de la ecuación (4.26), a partir del coeficiente de correlación calculado en (*i*). Es decir, r^2 (el coeficiente de correlación elevado al cuadrado) es el coeficiente de determinación "comparable" con el coeficiente de determinación calculado a partir de la ecuación (4.25).

Ejemplo 4.3. Caso adaptado de Gujarati y Porter (2010, pp. 204), en que se evaluó el consumo de café de los Estados Unidos de América (USA) mediante el uso de la siguiente relación (período 1970-1980): $Y_i = f(X_i)$

Donde:
Y_i: consumo de tazas de café al día
X_i: precio unitario de la taza de café.

Los resultados MCO con dos tipos de ecuaciones (lineal y doble logarítmica), son los siguientes

$$\hat{Y}_i = 2.691 - 0.480\,X_i \qquad (4.27)$$
$$(EE)\ (0.122) \quad (0.114) \qquad r_1^2 = 0.663$$

$$\widehat{LnY_i} = 0.777 - 0.253LnX_i \qquad (4.28)$$
$$(EE) \quad (0.0152) \ (0.114) \qquad r_2^2 = 0.678$$

En primera instancia uno estaría tentado a elegir el modelo (4.28), dado el mayor valor del coeficiente r^2, entonces según el procedimiento antes indicado, suponiendo la obtención de un valor "comparable" ($r_2^2 = 0.658$), entonces de la comparación:

Ecuación (1): $r_1^2 = 0.663$
Ecuación (2): $r_2^2 = 0.658$

Hay una diferencia que podría resultar significativamente importante. Cuando comparamos el valor de 0.658 con 0.745, resulta un valor "comparable" menor, que comparado con 0.663, nos llevaría a concluir que el modelo doble lineal tendría mejor ajuste que el doble-logarítmico.

4.6. Función Box-Cox

Es una función no-lineal que en Econometría se usa, entre otros, para hallar una mejor especificación de un modelo para un problema de regresión específico dado. Para efectos ilustrativos (con dos variables) está dado por una ecuación simple general:

$$\frac{Y_i^{\lambda}-1}{\lambda} = \alpha + \beta\left[\frac{Y_i^{\lambda}-1}{\lambda}\right] + \mu_i$$

Cuando λ=1, todo se reduce al modelo de regresión lineal:

$$Y_i - 1 = \alpha + \beta(X_i - 1) + \mu_i$$

Pero cuando λ=0, el análisis es más complejo pues el lado de la Variable dependiente, sería indeterminado. Entonces usando una expansión de Series de Taylor[20]:

$$Y_i^{\lambda} = \exp(\lambda LogY_i) = 1 + \lambda LogY_i + \left(\frac{1}{2}\right)(\lambda LogY_i)^2 + \cdots$$

Es demostrable la obtención final de una función semilogarítmica:

$\frac{Y_i^{\lambda}-1}{\lambda} = LogY_i$, entonces:

$$LogY_i = \alpha + \beta[X_i] + \mu_i$$

Otros valores de λ, dan lugar a otras formas funcionales diferentes. Por ejemplo, si λ es igual a –1, entonces la ecuación será una reciproca:

$$Y_i^{-1} = \beta_0 + \beta_1 X_i + \beta_2 X_i$$

La función de Box-Cox es útil pues incluye de los modelos no lineales en las variables (intrínsecamente lineales). Si λ es conocido simplemente reemplazamos. Pero excepto para los casos extremos -1, 0, 1 es difícil dar un valor de "λ". Entonces, mediante un tratamiento de λ como un parámetro adicional desconocido en la ecuación, se obtiene gran flexibilidad mediante estimación por máxima verosimilitud.

En resumen, contaremos con dos opciones para la estimación: uno muy sencillo, de ir calculando valores para λ y luego escoger la ecuación con el menor valor del coeficiente de determinación (R2) o la suma de cuadrados residual (SCR). La segunda opción, recomendable, es obtener simultáneamente los 3 (o más) estimadores de los parámetros incluyendo λ, utilizando el método de máxima verosimilitud.

[20] Es en realidad el uso de la regla de L´Hospital para series matemáticas.

Ejemplo 4.3. Si se tiene el siguiente modelo:

Y= valor de las residencias unifamiliares universitarias, determinado por un conjunto de variables expresadas como variables regresoras en la Tabla 4.1.

Se usó el método de Máxima Verosimilitud para determinar mediante Box-Cox, la especificación funcional más apropiada. La muestra fue de 5900 departamentos de 25 determinadas ciudades de los Estados Unidos (USA).

El valor estimado de λ es -0.1 (muy pequeño, casi cero). Aunque no es significativa la prueba *t* (0.05 margen de error), una especificación Log-lin es sugerida. Los resultados son mostrados en la Tabla 4.1.

Tabla 4.1. Estimaciones de máxima verosimilitud del ejemplo 4.3

Variable	**Coeficiente**	**Razón "t"**
Intercepto	6.007	6.14
λ (lambda)	– 0.10	2.62
Tamaño y atributos de calidad		
Año de construcción	0.115	2.61
Número de baños	0.124	2.36
Número de habitaciones	0.194	2.71
Cochera (1: si, 0: no)	0.047	2.11
Plagas(1: si, 0: no)	0.028	1.90
Vecindario (categorías varias)	0.058	2.27
Atributos de ambientación		
Sótano (1: si, 0: no)	0.027	2.21
Aire acondicionado (1: si, 0: no)	0.059	2.25
Calefacción (1: si, 0: no)	0.045	1.77
Atributos del clima		
Cálido (grados y N° días)	0.048	1.42
Frío (grados y N° días)	0.042	1.52
R^2= 0.48		

Fuente: ejemplo 10.2 de Pyndick & Rubenfeld (2001), pp. 292

Ejercicios del Capítulo IV: Especificación y Diagnóstico de Modelos

4.1. Fundamente la veracidad de los siguientes enunciados.

(a) Si en el modelo: $Y_i = \beta_0 + \beta_1 X_i + \beta_2 X_i^2 + \mu_i$, el estimador $\hat{\beta}_2$ resultara estadísticamente significativo, se debería conservar el término lineal aunque $\hat{\beta}_1$ resulte no significativo?.

(b) ¿Será idéntica línea de regresión si estimamos los siguientes modelos:

$Y_i = \beta_0 + \beta_1 X_i + \beta_2 X_2 + \mu_i$ (1)
$y_i = \alpha_0 + \alpha_1 x_i + \alpha_2 x_2 + \mu_i$ (2)

La ecuación (2) es igual a (1), pero con las variables en desviaciones respecto a sus medias.

> ***Sugerencias:***
> *(a) Cierto, pues en una ecuación cuadrática se requieren ambos coeficientes, necesariamente.*
>
> *(b) Cierto. En este caso el modelo (1) en desviaciones es:* $y_i = \beta_1 x_i + \beta_2 x_2 + (\mu_i - \bar{\mu})$.
> *En el modelo (2) se espera que el coeficiente* α_0 *resulte estadísticamente igual a cero. Por tanto, ambas ecuaciones tienen la misma pendiente.*

4.2. Considere los dos siguientes casos y modelos:

Caso	Teórico y verdadero	Empírico
“A”	$Y_i = \beta_1 X_i + \mu_i$ (1)	$Y_i = \alpha_0 + \alpha_1 X_i + v_i$ (2)
“B”	$Y_i = \beta_0 + \beta_1 X_i + v_i$ (1)	$Y_i = \alpha_1 X_i + v_i$ (2)

(a) En el caso “A” evalúe error de especificación e sus implicancias

(b) En el caso “B” también evalúe error de especificación e sus implicancias

> ***Sugerencias:***
> *(a) En el caso “A” existe un error de especificación y las implicancias son las siguientes:*
> *- Siendo* $E(\hat{\alpha}_1) = \beta_1$*; es decir el estimador de la pendiente es insesgado. Sin embargo, es comprobable que:* $V(\hat{\alpha}_1) \neq V(\hat{\beta}_1)$*. El estimador es ineficiente.*

(b) Caso "B": siendo $E(\hat{\beta}_1) \neq \alpha$: el estimador de la pendiente es sesgado. También las varianzas son diferentes, $V(\hat{\alpha}_1) \neq V(\hat{\beta}_1)$, por tanto el estimador es ineficiente.

4.3. Considere nuevamente el caso de un modelo sub-especificado.

El modelo verdadero es: $Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \mu_i$ (1)

El modelo empírico es: $Y_i = \alpha_0 + \alpha_1 X_{1i} + v_i$ (2)

(a) ¿Cuándo es el sesgo de error positivo y cuándo es negativo?

(b) ¿Son válidos los contrastes de significación habituales?

Sugerencias:
(a) Dado que es demostrable que $E(\hat{\alpha}_1) \neq \beta_1$y las El sesgo de omisión de la variable relevante X_{2i} depende de varios elementos: β_2 la COV muestral entre X_{1i} y X_{2i}, así como la varianza muestral de X_{1i}. Por ser la varianza positiva, a menudo el sesgo es positivo.

(b) Dado que el estimador es sesgado y también lo es la varianza de las perturbaciones, las pruebas de hipótesis son automáticamente inválidas.

4.4. Suponga que el verdadero modelo es:

$Y_i = \beta_0 + \beta_1 X_{1i} + v_i$ (1)

Y se utiliza en la práctica el modelo

$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \mu_i$ (2)

(a) Serán los 2 coeficientes de determinación (R2) diferentes en las ecuaciones (1) y (2)

(b) Serán los estimadores $\hat{\beta}_0$ y $\hat{\beta}_1$, de la ecuación (2), insesgados? ¿Por qué?

(c) En que afecta la inclusión de la variable X_{2i}.

Sugerencias:
(a) En la ecuación (2) la variable X_{2i} es irrelevante. Los coeficientes de determinación serán casi siempre mayores que los de la ecuación (1).

(b) Si, en el caso de variables irrelevantes y modelo verdadero sobreidentificado, entonces los estimadores serán insesgados. Note que β_3 será cero.

(c) La inclusión de la variable X_{2i} afecta en cuanto a la eficiencia.

$V(\hat{\beta}_1)=\frac{\sigma_\mu^2}{\sum X_i^2}$ *en el modelo (1) verdadero*

$V(\hat{\beta}_1)=\frac{\sigma_\mu^2}{\sum X_i^2(1-r_{12}^2)}$ *en el modelo (2) empírico.*

4.5. Considere la ecuación Cobb-Douglas, verdadera:

$$Y_i = \alpha_0 + \alpha_1 LnL_i + \alpha_2 LnG_i + \alpha_3 LnK_i + \mu_i \qquad (1)$$

Donde:
Y= producción; *L*i= trabajo; *G*i= gerencia; *K*i= capital.

Pero la realidad ha permitido la estimación de la ecuación (2):

$$Y_i = \beta_0 + \beta_1 LnL_i + \beta_3 LnK_i + \mu_i \qquad (2)$$

Evalúe las implicancias de este problema de especificación, en cuanto a lo siguiente:

(a) Serán $E(\hat{\beta}_1)=\alpha_1$? y $E(\hat{\beta}_2)=\alpha_2$. *Explique.*

(b) Suponiendo *G* es una variable irrelevante, ¿cómo cambia la respuesta previa? Explique.

Sugerencias:
(a) La omisión de una variable relevante conlleva estimaciones viciadas de los estimadores del modelo empírico. Por tanto, ambos valores esperados son viciados.

(b) Si G fuera una variable irrelevante, entonces los estimadores serán no viciados pero con dificultades en las varianzas.

4.6. Se desea analizar el gasto familiar semanal en alimentación, *G*i, para lo que se regresa esta variable en función de la renta familiar (R_i). Con un *n*=100 familias, se obtiene la siguiente estimación:

$G_i = 28.5 + 0.8R_i + \hat{\mu}_i$ (1)

(EE) (1.2) (0.2) R2=0.61

(a) ¿Qué opina y concluye al respecto de la propensión marginal a consumir?

(b) Analizado lo anterior (i), se plantea introducir otra variable potencialmente explicativa del Gasto familiar: es F_i (número de miembros de la familia) obteniéndose la siguiente ecuación estimada:

$G_i = \ 13.6 + \ 0.32R_i + 0.12F_i + \hat{v}_i$ (2)
(EE) (2.2) (0.08) (0.02) R2=0.89

Comentar y comparar los resultados de las ecuaciones (1) y (2)

Sugerencias:
(a) El valor de "t" calculado es: $t_c = \frac{0.8}{0.2}$*=4. >* t_t *(1.98), entonces la PMC resulta ser negativa, lo que contrasta con la teoría. Es posible que el modelo (1) esté mal especificado.*

(b) Los resultados "t" del modelo (2) indican también que las dos variables explicativas son significativas. Ambos coeficientes estimados tienen el signo positivo esperado: a mayor renta mayor gasto. El modelo (1) se ha omitido una variable relevante, lo que implica que en (1) los contrastes son sesgados y las pruebas de hipótesis no válidas.
El sesgo del estimador de la PMC en (1) depende del coeficiente de la variable F_i *y de la correlación que exista entre la variable incluida (renta) y la omitida* F_i *(miembros de la familia). Cuanto mayor es la correlación muestral entre ambas variables, mayor es el sesgo.*

4.7. Considere el siguiente modelo**:**

$R_i = \alpha_0 + \alpha_1 V_i + \mu_i$ (1)

Donde:
R = tasa de rendimiento promedio de activos de capital
V = volatilidad de activos

La verdadera volatilidad no es directamente observable, por tanto, se mide del siguiente modo: $r_{it} = \varphi_0 + W(r_{mt}) + e_t$

Donde:
r_{it} = tasa de rendimiento del valor "*i*"
r_{mt} = tasa de rendimiento del mercado durante el tiempo "*t*"

En este caso de evaluación de activos de capital, el coeficiente W de la ecuación (2) es considerado el mejor instrumento (VI) para la medición de la volatilidad de activos. Por tanto, en lugar de estimar (1) se estima la siguiente ecuación:

$$R_i = \alpha_0 + \alpha_1 W_i + \mu_i \qquad (2)$$

W_i es el coeficiente de esta ecuación (2), pero como las W_i son estimadas (no dadas), la relación entre la verdadera V_i y W_i es:

$$W_i = V_i + \tau_i$$

Donde τ_i es el error de corrección.
Respecto al parámetro α_1, ¿será insesgado el estimador $\hat{\alpha}_1$ de la ecuación (2)?. ¿Qué tipo de acción tomar?

Sugerencias:
Si $\tau_i > 0$ *entonces el estimador* $\hat{\alpha}_1$*resulta un estimador sesgado.*
Si $\tau_i = 0$, W_i *resulta ser misión de una variable instrumental relevante. Su relevancia podría evaluarse mediante uso del Contraste Hausman.*

4.8. Considere el modelo adecuado:

$$Y_i = \beta_0 + \beta_1 X_i + \mu_i \qquad (1)$$

Pero que se estima mediante uso de: X_i^* en lugar de X_i

¿Cuál será el efecto de los errores de medición expresados en c/u de los siguientes ítems?:

(a) $X_i^* = X_i + 5$

(b) $X_i^* = 3X_i$

(c) $X_i^* = X_i + \varepsilon_i$; donde ε_i es un elemento aleatorio "apropiado".

Sugerencias:
En la ecuación correcta (1) se reemplazan los valores de X_i, obteniéndose:
(a) $\hat{\beta}_{0c} = \hat{\beta}_0 + 5\hat{\beta}_1$; $\quad \hat{\beta}_{1c} = \hat{\beta}_1$
(b) $\hat{\beta}_{0c} = \hat{\beta}_0$; $\quad \hat{\beta}_{1c} = 3\hat{\beta}_1$
(c) *El estimador del intercepto será insesgado, pero el estimador de la pendiente será sesgado e inconsistente.*

4.9. Considere la ecuación verdadera:

$Y_i = \beta_0 + \beta_1 L_i + \beta_2 K_i + \mu_i$ (1)

Pero la realidad ha permitido la estimación de:

$Y_i = \alpha_0 + \alpha_1 L_i + \mu_i$ (2)

Es demostrable que $\hat{\alpha}_1$ es un estimador viciado, $E(\hat{\alpha}_1) = \beta_1 + \beta_2(b_{21})$
Donde b_{21} es la pendiente de la regresión de la variable excluida K_i en relación a la variable incluida L_i

Demuestre que $\hat{\alpha}_0$ es también un estimador sesgado.

> ***Sugerencias:***
> *De la ecuación empírica* (2) *obtenemo*s:
> $\hat{\alpha}_0 = \bar{Y} - \hat{\alpha}_1 \bar{L}$
>
> *Siendo luego demostrable que* $E(\hat{\alpha}_0) = \beta_0 + \beta_2(\bar{K} - b_{21}\bar{L})$: b_{21} *es la pendiente de la regresión de la variable excluida* K_i *en relación a la variable incluida* L_i.

4.10. Para saber si la variable X_i^2 pertenece al modelo (1) siguiente:

$Y_i = \beta_0 + \beta_1 X_i + \mu_i$ (1)

La prueba RESET de Ramsey involucra los siguientes pasos:
(i) Estimar el modelo lineal (1) para obtener la estimación de los valores:

$\hat{Y}_i = \hat{\beta}_0 + \hat{\beta}_1 X_i$

(ii) Luego se estima la ecuación auxiliar (2):

$Y_i = \alpha_0 + \alpha_1 X_i + \alpha_2 \hat{Y}_i^2 + v_i$ (2)

(iii) Se prueba la significancia de $\hat{\alpha}_2$ en (2).

Pregunta: demuestre que el uso de tal ecuación RAMSET (2) equivale a estimar de manera directa el modelo (3):

$Y_i = \beta_0 + \beta_1 X_i + \beta_2 X_i^2 + u_i$ (3)

Sugerencias:

De la estimación de la ecuación empírica (1), *obtenemos*: $\hat{Y}^2 = \hat{\beta}_0^2 + 2\hat{\beta}_0\hat{\beta}_1 X_i + \hat{\beta}_1^2 X_i^2$

Luego sustituimos este valor en la ecuación (2), para obtener finalmente la ecuación (3) que es la requerida para la demostración.

4.11. Dos investigadores estiman, con la misma muestra de 26 datos anuales, las siguientes dos ecuaciones:

Investigador 1: $\hat{C}_t = 17.3 + 0.56Y_t + 0.19R_t + 0.07S_{t-1}$ (1)
(EE) (3.7) (0.03) (0.15) (0.05) R2=0.93

$\widehat{W}_t = 1.37 + 0.71P_t + 0.40I_t + 0.32PR_{t-1}$ (2)
(EE) (0.04) (0.14) (0.07) (0.03) R2=0.94

Investigador 2: $\hat{C}_t = 16.9 + 0.57Y_t$ (3)
(EE) (1.3) (0.14) R2=0.91

$\widehat{W}_t = 2.71 + 1.35P_t$ (4)
(EE) (0.09) (0.16) R2=0.47

Donde:
C_t = consumo agregado
Y_t = ingreso disponible de las familias, tiempo "*t*"
R_t = nivel de riqueza familiar
S_t = nivel de ahorro
W_t = incremento en el salario promedio
P_t = incremento en el índice de precios consumidor IPC
I_t = incremento en el nivel de actividad económica
PR_t = nivel de productividad

(a) Probar la hipótesis que R_t y S_{t-1} no son conjuntamente significativos en el modelo (1) ni tampoco lo son I_t y PR_{t-1} en el modelo (2).

(b) Cambiaría su interpretación de las dos estimaciones de la función de consumo si supiera los siguientes valores de correlación: (i) entre $(Y_t, R_t) = 0.87$ y (ii) entre $(Y_t, S_{t-1}) = 0.93$.

Sugerencia:
(a) Para el modelo de determinación del consumo, el estadístico es:

$\boldsymbol{F} = \frac{(R_{nr}^2 - R_r^2)/(q)}{(1-R_r^2)/(n-k)} = \frac{(0.93-0.47)/(q)}{(1-0.93)/(25-4)} = 3 < 3.47(Ft, 2,21)$*, entonces no rechazamos la Ho.*

Igualmente, en el caso del modelo de determinación del salario, el estadístico es:
$\boldsymbol{F} = \frac{(R_{nr}^2 - R_r^2)/(q)}{(1-R_r^2)/(n-k)} = \frac{(0.94-0.47)/(q)}{(1-0.94)/(25-4)} = 82.8 > 3.47(Ft, 2,21)$*, entonces rechazamos la Ho.*
Se pueden realizar también las pruebas "t" para corroborar.

(b) Los resultados obtenidos en el modelo (1) se pueden deber a un alto grado de colinealidad entre Y_t y el resto de las regresoras. Por otro lado sabemos que esto implica que las varianzas de los estimadores son muy grandes, lo que lleva a no rechazar la Ho con más facilidad. Las variables R_t y S_{t-1} pueden ser relevantes para explicar el consumo, pero los datos que tenemos no permiten recoger sus efectos por separado una vez incluida la renta. En presencia de multicolinealidad se debería especificar y estimar el modelo (1).

4.12. Se tiene el siguiente modelo de regresión que determina las ventas de un producto Y_t en función de gastos de publicidad X_t

$$Y_t = \beta_0 + \beta_1 X_t + \beta_2 X_t^2 + \beta_3 X_t^3 + u_t \qquad (1)$$

... que se ha estimado con 28 observaciones, obteniéndose lo siguiente:

$$\hat{Y}_t = -73.4 + 0.005X_t + 0.000005X_t^2 - 0.000002X_t^3$$
(*EE*) (594.7) (0.12) (0.000008) (0.0000015) R2=0.834

(a) Pruebe si los datos sugieren que una función cuadrática sería suficiente, en base a:
$$\hat{Y}_t = -721.8 + 0.139X_t - 0.0005X_t^2 \qquad (2)$$
(*EE*) (70.9) (0.0096) (0.0004) R2=0.836

(b) Pruebe si los datos sugieren que la función suficiente sería lineal
$$\hat{Y}_t = -1753.1 + 0.016X_t \qquad (3)$$
(*EE*) (17.21) (0.0096) R2=0.834

(c) Si la estimación de la siguiente función doble-logarítmica proporciona los siguientes resultados:
$$\widehat{LnY_i} = -5.10 + 0.63LnX_i \qquad (4)$$

(EE) (0.15) (0.036) R2=0.703

¿Qué conclusiones puede obtener sobre la forma funcional más adecuada entre X y Y?

(d) Aplicando el contraste de Ramsey al modelo (4), incluyendo dos términos no lineales, se ha obtenido F=83,762. ¿Qué conclusión se puede obtener sobre la forma funcional más adecuada?

Sugerencias:

*(a) Para respuesta a este caso, se requiere contrastar la hipótesis simple sobre el parámetro β_3. El valor absoluto d*e ***t***c=1.33 < 1.98 (***t*** tabular con 24 gl). *Entonces NO se rechaza la hipótesis, o sea existe suficiente evidencia para la validez de una función cuadrática.*

*(b) Para dar respuesta a este caso lineal, se requiere contrastar la hipótesis simple sobre el parámetro β_2. El valor absoluto d*e ***t***c=12.56 > 1.98 (***t*** *tabular con* 24 gl). *Entonces rechazamos la hipótesis, no existe suficiente evidencia para la validez de una función lineal.*

(c) Dados los resultados indicados, no se puede concluir qué forma funcional es preferible, la cuadrática o la doble-logarítmica. No es posible concluir sobre qué forma funcional es mejor (de las dos indicadas). No es posible comparar los $\boldsymbol{R^2}$ *porque la variable dependiente no es la misma.*

(d) El modelo (4) incluyendo los dos términos no lineales tiene la forma:

$LnY_i = \beta_0 + \beta_1 LnX_{1i} + \gamma_1 \hat{Y}_1^2 + \gamma_2 \hat{Y}_2^3 + error$ *(modelo no restringido)*

Ho: el modelo doble-logarítmico es apropiado.

La forma funcional del modelo (4) será válido si los dos términos añadidos en el modelo previo no mejoran el ajuste del mismo. Se debe usar un estadístico F (Wald) para evaluar si el Modelo logarítmico efectivamente sería el mejor.

4.13. El gerente de una agroexportadora de aceitunas desea contar con un modelo para predecir su demanda con el fin de planificar su producción y las necesidades de materias primas. La empresa cree que la demanda de su producto (en miles de cajas), Y_t, viene influenciada básicamente por las siguientes variables:

X_{1t}= precio de venta en planta
X_{2t}= precio de venta de productos substitutos
X_{3t}= gastos en publicidad

Además, el gerente cree que sus ventas crecen en tasa constante en el tiempo, por lo que se propone el siguiente modelo:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4(t) + \mu_t \qquad (1)$$

Con 30 datos mensuales (de enero 1997 a junio 1999) se obtiene la siguiente regresión muestral:

$$\hat{Y}_t = 35.48 - 9.74X_{1t} + 2.94X_{2t} - 0.87X_{3t} + 0.31(t) \qquad \text{SCT=298.85}$$

(EE) (14.4) (3.76) (1.73) (0.75) (0.03) R2=0.834

Siendo algunas pruebas previas, que muestran no-aleatoriedad en los residuos en el tiempo, previendo posible error de especificación, conteste las siguientes preguntas:

(a) Interprete los coeficientes del modelo (1) y comente.

(b) El estadístico de Ramsey, añadiendo dos términos no lineales, toma el valor *F* =59.46 para el modelo (1) ¿Cree usted que el modelo es plausible? De no ser así, formule otro alternativo.

(c) Posteriormente, con la misma muestra se ha estimado (2):

$$\hat{Y}_t = 6.92 - 2.15X_{1t} + 1.55X_{2t} - 0.55X_{3t} - 0.32(t) + 0.31(t^2) \qquad (2)$$

(EE) (2.7) (0.69) (0.29) (0.14) (0.02) (0.0007)

- Comente los resultados de este modelo (2).

- ¿Cuál es el crecimiento mensual estimado en las ventas, *ceteris paribus*?

- ¿Qué puedes decir de los problemas y resultados del modelo (1)?

Sugerencia:
(a) El valor F de 32.8 permite rechazar la Ho de significancia conjunta de los estimadores; o sea al menos una o más variables regresoras son significativamente importantes; con la prueba "t" las variables precio y tiempo son significativas. Algunos de los resultados obtenidos no fueron los esperados. Ejemplo la variable precio de los bienes substitutos no es significativa.

(b) Para respuesta a este caso, se requiere usar RAMSEY: Se añaden dos términos no lineales a (1)

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \beta_3 X_{3i} + \beta_4(t) + \gamma_1 \hat{Y}_t^2 + \gamma_2 \hat{Y}_t^3 + error$$

(modelo no restringido)

Donde los valores $\hat{Y}_t^2$ y $\hat{Y}_t^3$son ajustados de (1).

El estadístico F de contraste arroja "rechazar" la Ho al 5% de error.

(c) El gráfico de los residuos del modelo (2) parece aleatorio: oscilan aleatoriamente en torno a cero. Esta función resulta de mayor utilidad para este problema.
- Utilizando (2), el crecimiento temporal (mensual) de la demanda, manteniendo el resto de variables regresoras constantes, es: $\frac{\partial E(Y)}{\partial t} = \beta_4 + 2\beta_5(t)$*; con los datos de la ecuación (2) este valor será variable. Se estima que el crecimiento de la demanda aumenta con el tiempo.*
- El modelo inicial (1) presentaba un problema de mala especificación de la forma funcional, lo que produce sesgos en la estimación MCO y hace que los contrastes de hipótesis no sean válidos. De allí los resultados que se obtenían: signos no esperados para alguna variable explicativa y variables explicativas no significativas.

4.14. Considere la ecuación

$$Ln(wage) = \beta_0 + \beta_1 Educ + \beta_2 Exper + \beta_3 (Exper^2) + \mu \qquad (1)$$

Donde:
$Wage$= salario de Economistas de Lima
$Educ$= nivel educativo de Economistas de Lima
$Exper$= experiencia laboral

Además, suponga que se piensa que la variable "educación" tiene problemas de medición; pero también "educación del padre" y "educación de la madre" pueden ser Variables Instrumentales (VI) que no están correlacionadas con el error.

A fin de evaluar "error de medición" en la variable "*Educ*" se procedió con la prueba específica de Hausman:

(a) Se especifica y estima la "forma reducida" para *Educ*:

$$Educ = \pi_0 + \pi_1 Exper + \pi_2 (Exper^2) + \pi_3 Meduc + \pi_4 Peduc + v$$
Obteniéndose la serie de residuales $\hat{v}$

(b) Se añade $\hat{v}$ como si fuera una variable regresora más en la ecuación original (1)

$$Ln(wage) = \beta_0 + \beta_1 Educ + \beta_2 Exper + \beta_3 (Exper^2) + \delta(\hat{v}) + \varepsilon \qquad (2)$$

Hecha la regresión resulta: $\hat{\delta}$=0.058

Con la información indicada, concluya sobre el "error de medición" en tres niveles de significancia: 0.01, 0.05 y 0.10.

Sugerencias: *revise el procedimiento de aplicación de la prueba de Hausman.*

4.15. Para una misma variable dependiente, se estiman tres modelos:

	Modelo 1	Modelo 2	Modelo 3
n	300	300	300
k	10	20	30
R2	0.89	0.95	0.99
SCR	182	180	178

(a) ¿Qué modelo elegiría según el valor del R^2 ajustado?

(b) ¿Qué modelo elegiría según el valor del criterio AIC (Akaike)?

Sugerencias: (a) Usar la fórmula de R^2 ajustado, en función del valor de R^2, *n* y *k* (b) Usar la fórmula de AIC, en función del valor de *SCR*, *n* y *k*

Ejercicios para uso de software

4.16. Para este ejercicio, con la misma variable dependiente, utilice los siguientes datos hipotéticos generados para dos modelos, uno simple y el otro múltiple**:**

N°	**Modelo simple (1)**			**Modelo múltiple (2)**		
	Y	*X1*		*Y*	*X1*	*X2*
1	20	55		20	55	5
2	22	58		22	58	8
3	23	59		23	59	9
4	27	60		27	60	10
5	28	62		28	62	12
6	31	70		31	70	14
7	33	75		33	75	18
8	35	80		35	80	20
9	36	90		36	90	22
10	40	95		40	95	25

(a) Reporte los resultados de estimación de ambos modelos, estimados en forma lineal. Comente sobre la significancia de los estimadores y sobre los valores de R^2 y R^2 ajustado.

(b) Respecto a la estimación de ambos modelos lineales, en (a), hallar los valores AIC y BIC y comente el modelo que elegiría, en base a tales criterios.

(c) Utilizando los datos del modelo múltiple (2), contemplando una ecuación lineal y otra logarítmica doble, estime los resultados y comente.

(d) De los modelos previos, en (d), hallar los dos valores R^2 ajustado, luego utilícelos para comparar y seleccionar el mejor modelo (lineal contra logarítmico). Comente.

Sugerencias:
Utilice las fórmulas para la comparación R^2 ajustada con AIC y BIC. En la última parte (d), utilice la transformación requerida en Y, para hacer los modelos comparables con este criterio.

4.17. Suponga el siguiente modelo verdadero de demanda[21]:

$$LnY_t = \beta_1 + \beta_2 LnX_{2t} + \beta_3 LnX_{3t} + \beta_6 LnX_{6t} + u_t \qquad (1)$$

Donde**:**
Y_t = consumo per cápita de pollos (Kg.)
X_{2t} = ingreso real disponible per cápita (S/.)
X_{3t} = precio al consumidor de pollos (S/.)
X_{6t}= nivel de renta promedio
μ_t= error.

Pero se ha estimado empíricamente el modelo:

$$LnY_t = \beta_1 + \beta_2 LnX_{2t} + \beta_3 LnX_{3t} + v_t \qquad (2)$$

(a) Realice las pruebas RESET y ML, considerada que (1) es la verdadera ecuación.

(b) Suponga que utilizándose (1) se obtuviera que el estimador de β_6 no fuera significativo. ¿Sería esto un indicativo que no existe error de especificación si se ajustara la ecuación (2)?.
Resuelva utilizando el fichero *Table7_9* (Gujarati & Porter, 2010)

[21] Ejercicio adaptado del capítulo 13 del libro de Gujarati y Porter (2010).

> ***Sugerencias:***
> *(a) La ecuación (1) corresponde al modelo no-restringido y la ecuación (2) al modelo restringido. Aplicando el procedimiento de MCR (mínimos cuadrados restringidos) podemos probar si la restricción (que la variable X6 no pertenece al modelo) es válida.*
> *Las pruebas RESET y el ML (Multiplicador de Lagrange) deberían indicar lo mismo.*
>
> *(b) La variable $Ln(X_6)$ no es estadísticamente significativa en este caso, pero en otra muestra podría serlo.*

4.18. Suponga el siguiente modelo verdadero de demanda[18]:

$$LnY_t = \beta_1 + \beta_2 LnX_{2t} + \beta_3 LnX_{3t} + v_t \qquad (1)$$

Donde**:**
Y_t = consumo per cápita de pollos (Kg.)
X_{2t} = ingreso real disponible per cápita (S/.)
X_{3t} = precio al consumidor de pollos (S/.)
v_t= error.

Pero se ha estimado empíricamente el modelo:

$$LnY_t = \beta_1 + \beta_2 LnX_{2t} + \beta_3 LnX_{3t} + \beta_6 LnX_{6t} + u_t \qquad (2)$$

Siendo X_{6t} el nivel de ingreso promedio.

(a) ¿Qué tipo de error de especificación se estaría cometiendo?

(b) Explique las consecuencias del error de especificación incurrido.

Resuelva utilizando el fichero *Table7_9* (Gujarati & Porter, 2010)

> ***Sugerencias:***
> *(a) Este sería un caso de inclusión de variables innecesarias (modelo sobreestimado).*
>
> *(b) Modelo sobre-especificado: los estimadores serían insesgados y consistentes. Sin embargo, las varianzas estarían sobredimensionadas*

4.19. Si el verdadero modelo fuera: : $Y_i = \beta_1 + \beta_2 X_2 + u_i$ [18]

Pero en la práctica se estima el modelo: $W_i = \beta_1 + \beta_2 Z_2 + v_i$, que tiene problemas de medición en las variables, por lo que:
$W_i = Y_i + \kappa_i$
$Z_i = X_2 + \omega_i$

Resuelva utilizando el fichero *Table13_2* (Gujarati & Porter, 2010)

***Sugerencias*:**
Los resultados del modelo "incorrecto", son:
$\hat{Y}_i = 28.30 + 0.584X_i$
(EE) (12.7) (0.07) R2=0.894
Estos resultados están cerca de los obtenidos con el modelo correcto (ver ecuación 5.11), pero de todos modos las magnitudes de los coeficientes reflejan el sesgo de medición.

4.20. En un modelo que explica los determinantes del sueldo de altos funcionarios (CEO), se ha especificado:

$$Ln(slary) = \beta_0 + \beta_1 Ln(sales) + \beta_2 Ln(roe) + \beta_3 (rosneg) + u_t \qquad (1)$$

Donde**:**
Y_t = salario (S/.)
ros = rendimiento de las acciones de la empresa (S/.)
$rosneg$ = (1: si *ros* < 0 y 0 si *ros* ≥ 0).

Aplique la prueba RESET utilizando la ecuación instrumental[22]:

$$Y_i = \beta_0 + \beta_1 X_1 + \cdots . + \beta_k X_k + \gamma_1 \hat{Y}_t^2 + \gamma_2 \hat{Y}_t^3 + \varepsilon_i \qquad (2)$$

(a) ¿Existe en esta ecuación (1) alguna evidencia de incorrecta forma funcional?

(b) Calcule una forma RESET ue sea robusta a heteroscedasticidad y evaluar si ello modifica la conclusión hallada en (a).

Utilice el fichero *ceosal1* para resolver este ejercicio (Wooldridge, 2010)

[22] Ejercicio adaptado del capítulo 13 del libro de Wooldridge (2010).

Sugerencias: *Realice la estimación de la ecuación del ejercicio de computador C7.5, de Wooldridge (2010) y luego obtenga* $\widehat{Lsalary}$. *Para usar la versión de RESET, añadimos* $\widehat{Lsalary}^2$ y $\widehat{Lsalary}^3$ *y luego se debe obtener el estadístico F-RESET para probar la significancia de estas "nuevas" variables. Resulta que no es posible rechazar la Hipótesis nula.*

4.21. Utilice el fichero *twoyear* para resolver este ejercicio relacionado con la educación universitaria (J. Wooldridge, 2010)[19]

(a) La variable *STOTAL* que puede funcionar como "proxi"[23] para la "capacidad" no observada. Encuentre la media y desviación estándar muestrales de *stotal*.

(b) Utilice las regresiones simples entre $JC = f(stotal)$ y $UNIV = f(stotal)$. Explique la relación entre estas dos variables dependiente, de la educación universitaria, con respecto a la variable *stotal.*

(c) Agregue *stotal* a la ecuación: $Ln(wage) = \beta_0 + \beta_1 JC + \beta_2 univ + \beta_3 exper + u_t$

Donde**:**
JC = años en una carrera corta universitaria
$univ$ = años en una carrera corta universitaria
$exper$ = meses de experiencia

Contraste la hipótesis que el rendimiento de educación "corta" y "larga" es igual, frente al caso que la educación "larga" es superior.

(d) Agregue el término cuadrático de la variable *stotal* y evalué su importancia en el modelo.

(e) Agregue a la ecuación en (b) las interacciones *stotal*∗*jc* y entre *stotal*∗*univ*. Indique si tales términos son significativos.

(f) ¿Cuál podría ser el modelo correcto en que se controle la capacidad mediante el uso de la variable *stotal*? Justifique.

[23] Una variable "proxi" es una variable observada relacionada, pero no idéntica, a una variable explicativa inobservable, en un modelo de regresión multivariado.

Sugerencias:
(a) La media de la variable STOTAL es 0.047, la desviación estándar es 0.011.

(b) En la regresión simple de JC sobre STOTAL el coeficiente de la pendiente es 0.11 (D.E. 0.11); la correlación entre ambas variables es débil. Por otro lado, en la regresión simple de UNIV sobre STOTAL el coeficiente de la pendiente es 1.170 (Desv. Estándar=0.029); ambas variables están positivamente correlacionadas.

(c) Cuando añadimos STOTAL a (4.17) obtenemos:

$$Ln(\widehat{wage}) = 1.495 + 0.063(INC) + 0.686(UNIV) + 0.00488(EXPER) + 0.494(STOTAL);$$

(EE) (0.021) (0.0068) (0.0026) (0.0016) (0.0068)

R2=0.228; n=6758.

(d) Cuando añadimos el término cuadrático STOTAL^2 *a la ecuación estimada en (iii), su coeficiente resulta en 0.0019 (tc=0.4).Concluya.*

(e) El estadístico F para comprobar la exclusión de los términos de interacción es 1.9 (p-valor de 0.141).

(f) Sería recomendable el modelo de (v) donde STOTAL aparece solo a nivel.

CAPÍTULO V. MODELOS CON DATOS PANEL

Hasta ahora los modelos que hemos revisado han sido multivariados, uniecuacionales y asumiendo diferentes formas funcionales; aunque todos han tenido un elemento en común: han usado una sola dimensión de análisis; corte transversal o series de tiempo. Pero todos los modelos vistos pueden ser también especificados y resueltos con base en la utilización de una base de datos *panel*, es decir una base de datos que utiliza las dos dimensiones posibles de la data: espacio y tiempo, o corte transversal y series de tiempo. Un conjunto de datos panel puede ser de mucha utilidad en la medida que permite "clasificar efectos económicos" que no podrían ser distinguidos sólo con el uso de datos de corte transversal o de series de tiempo.

En economía aplicada, los modelos que utilizan bases de datos *panel* son también conocidos como "Modelos Panel" y la base de datos toman también el nombre de "base de datos bi-dimensionales", "data longitudinal", "data de historia de sucesos", etc. En rigor, sin embargo, la denominación "*panel*" se refiere a los datos propiamente y se denomina "modelos panel" a los modelos que utilizan tales datos.

Aunque los datos *panel* son aún escasos, por ser costosos, son más frecuentes que en el pasado. Los modelos con datos *panel* contienen matemática compleja; sin embargo, existe software que facilita mucho el trabajo con ellos. En este capítulo se presenta todo en forma sencilla y sin la profundidad matemática que podría ser encontrada en otros libros de texto.

Si en una estructura de datos *panel* se tiene que el número de sujetos de corte transversal, "n", es mayor que el número de períodos de tiempo "t ", entonces se refiere a una base de datos de "*panel corto*"; viceversa, un "*panel largo*" es aquel en que el número de períodos de tiempo es mayor que los de corte transversal "n". Asimismo, un panel "balanceado" es aquel en que cada entidad o sujeto (empresas, países, individuos, etc.) tiene la misma cantidad de observaciones; un panel "desbalanceado" no tiene esa característica.

5.1. Características de la Base de Datos

- Los modelos *panel* contienen una mayor cantidad de información con más variabilidad, menos colinealidad entre regresoras, mayor grado de libertad y mayor eficiencia en la estimación de parámetros.

- La dimensión temporal le otorga mayor posibilidad de estudiar la dinámica causal de cambios entre variables económicas, sobretodo.

- Los modelos *panel* permiten estudiar mejor algunos fenómenos relativamente complejos en economía. Ejemplo: cambio tecnológico, economías de escala, cambio climático, etc.

- Una desventaja tal vez sería la complejidad matemática de sus fundamentos en la estimación de modelos, pero es subsanable en tiempos actuales de "digitalización".

5.2. Opciones de Estimación

¿Cómo estimamos los parámetros de un modelo que utiliza datos de panel?. Al menos contamos con 3 posibilidades:

(a) Mediante mínimos cuadrados ordinarios (MCO) con datos totalmente agrupados, desatendiendo las dimensiones de corte transversal y de la serie de tiempo (en inglés esta opción es denominada "*pooled*"). En este caso frecuentemente el modelo toma el nombre de *Modelo de Datos Agrupados*.

(b) Mediante MCO para datos de panel, pero atendiendo la naturaleza del corte transversal y/o de la serie de tiempo. En el primer caso, por ejemplo, se permite que los individuos de corte transversal tengan diferentes interceptos (construidos con variables ficticias), a la vez que se controlan efectos inobservables constantes en el tiempo. Al modelo resuelto mediante este procedimiento, se suele denominar *Modelo Panel de efectos fijos* (MEF).

(c) Mediante el uso del método de mínimos cuadrados generalizados (MCG) y atendiendo la naturaleza del corte transversal y/o de la serie de tiempo. Suponiendo la atención es fijada en la dimensión del corte transversal, se propone que los individuos tengan diferentes interceptos, pero obtenidos en forma aleatoria de una población mayor de instancias de Corte Transversal. Del mismo modo se puede aplicar al caso de las unidades temporales. Al modelo resuelto mediante este procedimiento, se suele denominar *Modelo Panel de efectos aleatorios* (MEAL).

Las opciones de estimación (b) y (c) expresan la idea de reconocer que un conjunto de efectos inobservables (variables omitidas) pueden conducir a cambios en los interceptos del corte transversal y de la serie de tiempo. Estos cambios pueden ser "fijos" o pueden ser "aleatorios", dependiendo del método de estimación que se use. A menudo la consideración de las dos dimensiones juntas conlleva problemas de pérdida de grados de libertad y complejidad en el análisis. En este capítulo, para efectos de simplicidad, consideramos solo una dimensión del modelo panel: la del corte transversal, que es el más usado en el ámbito de la economía aplicada.

Mínimos cuadrados generalizados (MCG)

MCG es un método de estimación de parámetros que consiste en utilizar MCO en una ecuación transformada (2), a partir de una ecuación original (1) que tiene dificultades o que sabemos que, con MCO, no podremos conseguir la obtención de estimadores insesgados y óptimos (ELIO). La esencia es llegar a un modelo alternativo, transformado, que finalmente cumple con los supuestos de Gauss-Markov más el de normalidad del residuo.

$$Y_i = \beta_0 + \beta_1 X_{1i} + \cdots + \beta_k X_{ki} + \mu_i \qquad (1)$$

$$\alpha_0 + \alpha_1 X_{1i}^* + \cdots + \alpha_k X_{ki}^* + U_t^* \qquad (2)$$

El método de MCG es ampliamente utilizado, entre otros, para corregir problemas de heteroscedasticidad y correlación serial.

Un ejemplo de un estudio con base de datos panel en consideración de la heterogeneidad de las unidades del corte transversal, es el siguiente[24]:

Se pretende evaluar los determinantes de la pobreza en cinco países durante un período de ***28*** años, período 1983 - 2010 (140 observaciones totales del panel), y se desea averiguar la dependencia de la pobreza respecto a la contaminación por emisión de CO_2 per cápita, la población económicamente activa, el desempleo, producto per cápita (*PBI_PER*) y la población total.

$$Y_{it} = \beta_{0i} + \beta_1 X_{1it} + \beta_2 X_{2it} + \beta_3 X_{3it} + \beta_4 X_{4it} + \beta_5 X_{5it} + U_{it} \qquad (5.1)$$

Donde:
$\boldsymbol{Y_{it}}$ = tasa de pobreza (%)
$\boldsymbol{X_{1it}}$= emisión promedio de carbono (CO2 en Kg x hab)
$\boldsymbol{X_{2it}}$= población económicamente activa / población total (%)
$\boldsymbol{X_{3it}}$= desempleo (%)
$\boldsymbol{X_{4it}}$= PBI per cápita (US$ del año 2000).
$\boldsymbol{X_{5it}}$= Población total (miles de personas).
$\boldsymbol{U_{it}}$= error idiosincrático

[24] Ejemplo adaptado de Gujarati y Porter (2010), sección 16.3, pp. 594.

5.2.1. Modelo con datos agrupados (MDA)[25]

Como se indicó antes, en este caso de datos de panel "agrupado", la base constituye un conjunto de observaciones agrupadas sin atender la naturaleza de corte transversal ni series de tiempo de los datos. Los supuestos manejados en este caso son los siguientes:

Los coeficientes obtenidos de la regresión serán semejantes para todos los sujetos u observaciones. MCO requiere que $U_{it} \sim iid\, \mathrm{N}(0,\sigma_{\mu}^{2})$; es decir no existe autocorrelación ni heteroscedasticidad ni colinealidad. Las variables explicativas no son estocásticas, y si lo fueran, no están correlacionadas con el término de error general.

El mismo ejemplo previo expresado a través de la especificación (5.1), con β_{0i} como término constante, permite obtener los resultados mostrados en la Tabla 5.1.

Tabla 5.1. Resultados modelo de datos agrupados

Method: Panel Least Squares
Date: 03/13/12 Time: 21:28
Sample: 1983 2010
Periods included: 28
Cross-sections included: 5
Total panel (balanced) observations: 140

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	30.59479	3.568072	8.574602	0.0000
CO2_KT	-0.000108	3.15E-05	-3.439461	0.0008
DESEMP	0.626466	0.255964	2.447475	0.0157
PBI_PER	0.006681	0.001126	5.934454	0.0000
PEA	5.40E-08	9.45E-08	0.571460	0.5686

R-squared	0.489560	Mean dependent var	46.07755
Adjusted R-squared	0.474435	S.D. dependent var	11.27530
S.E. of regression	8.174120	Akaike info criterion	7.074884
Sum squared resid	9020.192	Schwarz criterion	7.179943
Log likelihood	-490.2419	Hannan-Quinn criter.	7.117577
F-statistic	32.36937	Durbin-Watson stat	0.158952
Prob(F-statistic)	0.000000		

[25] En la literatura normalmente los modelos toman su denominación fundamentalmente en base a la especificación; por ejemplo, modelos cuadráticos, modelos logarítmicos, modelos dinámicos, etc. En algunas ocasiones los modelos también toman su nombre en base a la especificación y el método de estimación; por ejemplo, un modelo de ecuaciones simultáneas o un modelo recursivo. En el caso de modelos de regresión con datos panel, los nombres son asignados básicamente a la estructura bi-dimensional de los datos.

Particularmente, en este caso es notable la existencia de un potencial bajo coeficiente D-W (Durbin Watson), que además de implicar la presencia de autocorrelación (AC), también ha mostrado ser un indicador de alerta de un error de especificación.

Dificultad del modelo MDA

El principal problema del modelo de datos agrupados (*MDA*) es que no distingue entre instancias transversales (países en el ejemplo) u "oculta" la heterogeneidad o individualidad entre países. Tal heterogeneidad se incluye en el término de error general (también denominado *error idiosincrático*). Por tanto, es bastante probable que se produzca un problema de endogeneidad permitiendo que los estimadores finalmente resulten sesgados e inconsistentes.

Alternativas al modelo con datos no agrupados

Como alternativa al modelo panel agrupado, considere el siguiente replanteamiento del modelo:

$$Y_{it} = \beta_0 + \beta_1 X_{1it} + \beta_2 X_{2it} + \cdots + \varepsilon_i + w_t + \nu_{it} \qquad (5.2)$$

Donde:
i = 1, 2, 3,… i-ésima observación de corte transversal
j = 1, 2, 3,… j-esimo período temporal.

El error llamado "idiosincrático" $(\boldsymbol{U_{it}})$ aparece descompuesto como:
$\boldsymbol{U_{it}} = \boldsymbol{\varepsilon_i} + \boldsymbol{w_t} + \boldsymbol{\nu_{it}}$

En la ecuación (15.2) el componente $\boldsymbol{\beta_0}$ recoge factores o "shocks" no observables y que no varían por individuos (ni en el tiempo en PANEL).

El componente $\boldsymbol{\varepsilon_i}$ recoge factores no observables que no varían en el tiempo pero varían en corte transversal, representando (en este ejemplo) la heterogeneidad de los países. Un factor no observado y que no varía en el tiempo puede ser, por ejemplo, "gobernabilidad". Del mismo modo, en la ecuación (5.2) se ha separado uno o más efectos no observables que varían en el tiempo más no en el corte transversal (que es w_t en este caso). El término $\boldsymbol{\nu_{it}}$ debería representar solo eventos puramente aleatorios.

Entonces, la idea de un modelo panel es hacer explícitos los efectos no observados que no varían en alguna dimensión (o en ambas) y que permitan controlar tales efectos (no observados) que son los que ocasionan heterogeneidad entre instancias. Los modelos PANEL pueden ser considerados un "salvataje" del modelo de datos agrupados.

Es importante indicar que en el caso que consideremos, en forma conjunta, heterogeneidad tanto en observaciones de corte transversal como en observaciones temporales, el modelo se vuelve complejo y se perderían muchos grados de libertad, por lo que es más práctico considerar la heterogeneidad o individualidad solo en una de las dos dimensiones.

Nota importante

En el presente capítulo, para efectos de simplicidad, en la presentación de las ecuaciones o modelos que usan datos de panel, la heterogeneidad es especificada en solo una dimensión del modelo (de las dos posibles). La presentación se hace específicamente para un modelo panel con heterogeneidad del corte transversal, pensando en un panel "largo". Dejando en claro que con un panel corto la dimensión en la que se especificaría mejor la heterogeneidad se presenta con la dimensión temporal.

Frecuentemente los modelos panel que consideran una sola dimensión son aquellos en los cuales se tiene evaluada la heterogeneidad de las observaciones de corte transversal, tal como por ejemplo empresas, países, personas, familias, etc. Los modelos panel que consideran la doble dimensión en el análisis, se denominan, a menudo, modelos dobles o bidimensionales.

En el caso de un modelo panel con consideración de la heterogeneidad en el corte transversal, tendremos:
$Y_{it} = \beta_0 + \beta_1 X_{1it} + \beta_2 X_{2it} + \cdots + \varepsilon_i + \nu_{it}$

Donde $\varepsilon_i \sim N(0, \sigma_\varepsilon^2)$

¿Cómo manejar el intercepto de este modelo? $\boldsymbol{\beta_{0i} = \beta_0 + \varepsilon_i}$.

Existen hasta 3 opciones para el manejo del "error no observado" $\boldsymbol{\varepsilon_i}$

(1) $\boldsymbol{\beta_{0i} = \beta_0}$	(2) $\boldsymbol{\beta_{0i} = \beta_0 + \alpha_i}$	(3) $\boldsymbol{\beta_{0i} = \beta_0 + \tau_i}$

La primera opción implicaría el uso de un modelo de datos agrupados, la segunda a un modelo panel de efectos fijos y la tercera a un modelo panel de efectos aleatorios. Es lo que se presenta a continuación.

5.2.2. Modelo panel de efectos fijos inter-grupos (MEF)

Una forma de resolver un modelo panel es mediante el uso de variables ficticias (dummy) que ayuden a representar la heterogeneidad de las unidades, ya sea las observaciones del corte transversal o las del tiempo. La técnica de uso de variables ficticias es realmente la más sencilla y más usada en el trabajo de economía aplicada, se le conoce también como modelo de efectos fijos inter-grupos.

Si el modelo panel es:

$$Y_{it} = \beta_{0i} + \beta_1 X_{1it} + \beta_2 X_{2it} + \cdots + U_{it} \qquad (5.3)$$

Y requerimos manejar diferencias entre las unidades de corte transversal, entonces:

$$Y_{it} = (\alpha_1 + \alpha_2 D_{2i} + \alpha_3 D_{3i} + \alpha_4 D_{4i} + \alpha_5 D_{5i}) + \beta_1 X_{1it} \ldots + \ \beta_2 X_{2it} + \cdots + U_{it}$$

La idea central es que las variables ficticias representen las diferencias o heterogeneidad entre las entidades del Corte Transversal: en este caso se considera que la heterogeneidad entre unidades transversales no varía con el tiempo- tendrían que usar variables ficticias interactivas. La Figura 5.1 muestra diferencias entre el ajuste que puede lograrse entre el modelo de datos agrupados y el modelo panel de efectos fijos.

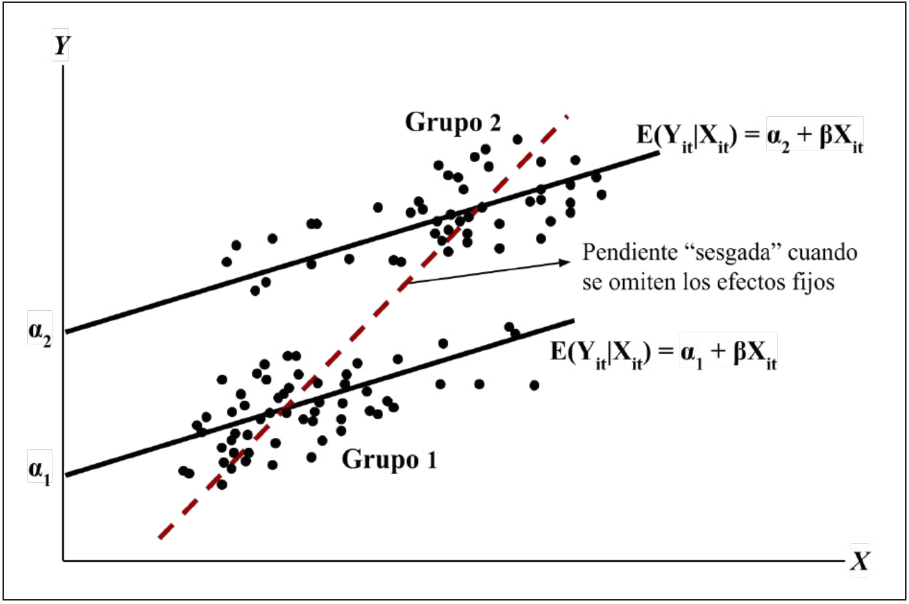


Figura 5.1. Ajustes de un modelo de datos agrupados y otro de efectos fijos

Limitaciones del MEF

- Si se tuvieran que incorporar las dos dimensiones posibles del panel (corte transversal y serie de tiempo), el uso de variables dicotómicas puede crear el problema de la pérdida de grados de libertad, al punto extremo que no habría suficientes datos para un análisis estadístico significativo (esto puede suceder incluso aún con una sola dimensión, pero con demasiadas observaciones transversales en, por ejemplo, un panel corto).

- También cuando hubiera demasiadas variables ficticias en el modelo incrementaría la posibilidad de multicolinealidad, que puede finalmente conllevar imprecisión e ineficiencia de los estimadores.

- El término de error podría generar dificultades debido a que el sub-índice *i* se refiere a las observaciones de corte transversal y el subíndice *t* a las observaciones de series de tiempo (al haber diversas posibilidades de autocorrelación y/o heteroscedasticidad).

Especificación de un modelo de efectos fijos (MEF)

Un modelo de Efectos Fijos con VD temporales sería especificado así:

$$Y_{it} = (\lambda_0 + \lambda_1 T_{84} + \lambda_2 T_{85} + \cdots + \lambda_{27} T_{10}) + \beta_1 X_{1it} + \beta_2 X_{2it} + \cdots + U_{it}$$

Ejemplo 5.1. Retomando el ejemplo previo que busca evaluar los determinantes de la pobreza en 5 países durante un período de ***28 años,*** período 1983 - 2010 (140 observaciones totales del panel), y se desea averiguar los determinantes de la pobreza.

$$Y_{it} = \beta_{0i} + \beta_1 X_{1it} + \beta_2 X_{2it} + \beta_3 X_{3it} + \beta_4 X_{4it} + \beta_5 X_{5it} + U_{it} \qquad (5.4)$$

Donde:
$\boldsymbol{Y_{it}}$ = tasa de pobreza (%)
$\boldsymbol{X_{1it}}$= emisión promedio de carbono (CO2 en Kg x hab)
$\boldsymbol{X_{2it}}$= población económicamente activa / población total (%)
$\boldsymbol{X_{3it}}$= desempleo (%)
$\boldsymbol{X_{4it}}$= PBI per cápita (US$ del año 2000).
$\boldsymbol{X_{5it}}$= Población total (miles de personas).
$\boldsymbol{U_{it}}$= error idiosincrático
i = 1, 2, 3, 4, 5 países en el corte transversal
j = 1, 2, 3,…, 28 años.

En este caso: $\beta_{0i} = (\alpha_1 + \alpha_2 D_{2i} + \alpha_3 D_{3i} + \alpha_4 D_{4i} + \alpha_5 D_{5i})$

$\boldsymbol{D_{1i}}$ es la variable ficticia de referencia, habiéndose omitido el intercepto, a fin de evitar caer en la denominada "trampa de las variables ficticias". Las demás variables ficticias aparecen explícitamente en el modelo y toman los valores 0 y 1 dependiendo del país al que representa. Ejemplo $\boldsymbol{D_{2i}}$=1 es la observación que corresponde al país 2 y "0" en otros casos. Los resultados se presentan en la Tabla 5.2 debajo.

Tabla 5.2. Resultados de un modelo panel de efectos fijos (Ejm. 5.1)

Dependent Variable: POBREZA
Method: Panel EGLS (Cross-section weights)
Date: 03/13/12 Time: 21:29
Sample: 1983 2010
Periods included: 28
Cross-sections included: 5
Total panel (balanced) observations: 140
Linear estimation after one-step weighting matrix
White cross-section standard errors & covariance (d.f. corrected)

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	88.98524	5.431768	16.38237	0.0000
CO2_KT	-0.000266	5.03E-05	-5.297772	0.0000
DESEMP	-0.175069	0.184139	-0.950745	0.3435
PBI_PER	-0.011685	0.002101	-5.561024	0.0000
PEA	9.43E-07	2.56E-07	3.682659	0.0003

Effects Specification

Cross-section fixed (dummy variables)

Weighted Statistics

R-squared	0.694209	Mean dependent var	47.77718
Adjusted R-squared	0.675535	S.D. dependent var	10.87171
S.E. of regression	5.778778	Sum squared resid	4374.650
F-statistic	37.17466	Durbin-Watson stat	0.332872
Prob(F-statistic)	0.000000		

CROSS SECTION FIXED EFECT

Brazil	-6.511957
Colombia	-5.101157
Ecuador	-25.40861
Peru	-21.55026
Venezuela	58.57199

En comparación con los resultados obtenidos con el modelo agrupado (Tabla 5.1), en este caso los resultados muestran que tenemos varios interceptos que expresan heterogeneidad entre los resultados de contaminación en cada país, lo que pone en duda la pertinencia del modelo agrupado de datos (*pooled*). También la significancia estadística de los coeficientes de las variables son diferentes (evaluar).

En adición, podemos también el estadístico ***F*** siguiente (prueba de Wald):

$$\boldsymbol{F} = \frac{(R_{nr}^2 - R_r^2)/(N° \; de \; nuevas \; regresoras)}{(1 - R_{nr}^2)/(n - parámetros \; en \; el \; modelo \; no \; restringido)}$$

Siendo la ecuación (5.4) "no restringida", mientras que la restringida pues impone restricciones de que los interceptos son idénticos para todas las entidades.

Si el valor ***F*** calculado es significativo, 5%, podemos decidir:

Ho: La ecuación (1) está mal especificada
(es decir ambos parámetros $\gamma_1 \; y \; \gamma_2$=0 en la EQ (2))

Ha: No hay evidencia de mala especificación en (1).

En este caso:

$$F = \frac{(0.694 - 0.49)/(5)}{(1 - 0.694)/(140 - 8)}$$

= 17.74 ⟼ se rechaza la Ho de que todos los interceptos son semejantes. Valoramos mejor el EPF con VD.

Es importante, sin embargo, considerar (en el modelo MEF) lo siguiente:

(i) Si permitimos demasiado V. ficticias, se vuelve compleja la interpretación, se pierden GL y se incrementa la posibilidad de multicolinealidad.

(ii) Debemos reflexionar con cuidado sobre $\boldsymbol{v_{it}}$**,** en relación a que cumpla con los supuestos Gauss-Markov requerimientos de MCO.

5.2.3. Modelo de efectos fijos dentro del grupo (MDG)

Este modelo es también un modelo panel de efectos fijos que utiliza un método diferente en la resolución. El método es también denominado "método de transformación intra-grupos"; consiste en utilizar un modelo alternativo, que es transformado mediante el uso de variables diferenciadas respecto a los valores promedio.

Consideremos un modelo con una sola variable explicativa**:**

$$Y_{it} = \beta_0 + \beta_1 X_{it} + \varepsilon_i + U_{it} \qquad (5.5)$$

Ahora, para cada $\boldsymbol{i}$, escribimos una ecuación para los valores medios de las variables, a través del tiempo, obteniendo**:**

$$\bar{Y}_i = \beta_0 + \beta_1 \bar{X}_i + \varepsilon_i + \bar{U}_i$$

$\boldsymbol{t}$ = *1,2,…,T* (o sea los valores están en promedio a través del tiempo)

$\Longrightarrow$ el modelo transformado será**:**

$$(Y_{it} - \bar{Y}_i) = \beta_1 (X_{it} - \bar{X}_i) + (U_{it} - \bar{U}_i) \qquad (5.6)$$

O bien, en valores diferenciales: $y_{it} = \beta_1 x_{it} + u_{it}$

En la ecuación transformada (5.6) se ha eliminado el intercepto fijo $\boldsymbol{\beta_{0i}}$**,** atribuyendo al propio modelo transformado la heterogeneidad requerida. La técnica permite obtener estimadores consistentes de los coeficientes de las pendientes β_i, aunque la eficiencia no estaría garantizada. Con respecto al estimador del *MDG,* se obtienen estimaciones consistentes de los coeficientes de pendiente, mientras que la regresión agrupada ordinaria tal vez no.

Una potencial dificultad sería que esta estimación podría ser que al diferenciar las variables, se podrían distorsionar los valores de los parámetros y eliminar los posibles efectos de largo plazo. El modelo sería solo útil para el corto plazo.

Ejemplo 5.2. En el ejemplo utilizado previamente, referido a la pobreza en cinco países, período entre 1983 y 2010, y propósito de averiguar los determinantes de la pobreza.

$$Y_{it} = \beta_{0i} + \beta_1 X_{1it} + \beta_2 X_{2it} + \beta_3 X_{3it} + \beta_4 X_{4it} + \beta_5 X_{5it} + U_{it} \qquad (5.7)$$

Donde:

$\boldsymbol{Y_{it}}$ = tasa de pobreza (%)

$\boldsymbol{X_{1it}}$= emisión promedio de carbono (CO2 en Kg x hab)

$\boldsymbol{X_{2it}}$= población económicamente activa / población total (%)

$\boldsymbol{X_{3it}}$= desempleo (%)

$\boldsymbol{X_{4it}}$= PBI per cápita (US$ del año 2000).

$\boldsymbol{X_{5it}}$= Población total (miles de personas).

$\boldsymbol{U_{it}}$= error idiosincrático

$\boldsymbol{i}$ = 1,2,3,… 5 países en el corte transversal

$\boldsymbol{j}$ = 1,2,3,… 28 años.

La ecuación panel de efectos fijos con variables ficticias (intergrupos) se representa como:

$$Pobreza_{it} = \alpha_0 + \alpha_1 D_{1i} + \alpha_2 D_{2i} + \alpha_3 D_{3i} + \alpha_4 D_{4i} + \beta_1(co\_KT) + \beta_2(desempleo) + \beta_3(pbi_{per}) + \beta_4(pea) + \beta_5(pop\_mil) + U_{it} \quad (5.8)$$

Mientras que la ecuación panel ***MDG*** (intragrupos) es:

$$(pobreza)_{it} = \beta_1(co2_{kt}) + \beta_2(desempleo) + \beta_3(pbi_{per}) + \beta_4(pea) + \beta_5(pop\_mil) + u_{it} \quad (5.9)$$

Matemáticamente los modelos son iguales (Gujarati y Porter, 2010); los estimadores de pendiente de las variables regresoras, son idénticos en ambas ecuaciones, mientras que los interceptos del segundo modelo (intra-grupos) pueden obtenerse a partir de la fórmula siguiente:

$$\widehat{\alpha}_i = \overline{pobreza}_i - \hat{\beta}_1\overline{c02\_kt}_i - \hat{\beta}_2\overline{desempleo}_i - \hat{\beta}_3\overline{pbi\_per}_i - \hat{\beta}_4\overline{pea\_pop}_i - \hat{\beta}_5\,\overline{pop\_mil}_i$$

En la Tabla 5.3 observe que en el modelo panel intergrupos, el intercepto estimado para cada país representa las características específicas de cada país, aunque no podemos identificar estas características individualmente (pueden ser por ejemplo: tipo de gobierno, composición del gabinete, personalidad y género del Presidente, etc). En resumen, los resultados son los siguientes:

$$\widehat{POBR}_{it} = 25.64D_{0i} + 94.85D_{1i} + 101.86D_{2i} + 99.15D_{3i} + 161.7D_{4i} - 0.00028(CO2_{KT}) - 0.0444(DESEMPLEO) - 0.0057(PBI_{PER}) - 1.495(PEA) + 0.0011(POP).$$

Los resultados del modelo panel ***MDG*** (intra-grupos) se muestran en la Tabla 5.3.

Tabla 5.3. Resultados del modelo panel con desviaciones, ejemplo 5.2

Dependent Variable: DD_POBREZA
Method: Least Squares
Date: 03/13/12 Time: 21:39
Sample: 1983 2010
Periods included: 28
Cross-sections included: 5
Total panel (balanced) observations: 140

Variable	Coefficient	Std. Error	t-Statistic	Prob.
DD_CO2_KT	-0.000266	5.05E-05	-5.247372	0.0000
DD_DESEMP	-0.175069	0.189329	-0.924685	0.3135
DD_PBI_PER	-0.011685	0.003201	-3.661021	0.0000
DD_PEA	9.43E-07	2.58E-07	3.655049	0.0004
R-squared	0.677209	Mean dependent var		4.721218
Adjusted R-squared	0.655035	S.D. dependent var		1.087101
S.E. of regression	0.601078	Sum squared resid		374.6500

F-statistic	35.17512	Durbin-Watson stat	0.383272
Prob(F-statistic)	0.000000		

En resumen:

$$\widehat{pobr}_{it} = -0.00028(co2\_kt) - 0.0444(desempleo) - 0.0057(pbi\_per) - 1.495(pea) + 0.0011(pob)$$

Luego los interceptos pueden ser también calculados del siguiente modo:

$$\hat{\alpha}_i = \overline{pobreza_i} - \hat{\beta}_1\overline{c02\_kt_i} - \hat{\beta}_2\overline{desempleo_i} - \hat{\beta}_3\overline{pbi\_per_i} - \hat{\beta}_4\overline{pea_i} - \hat{\beta}_5\overline{pob_i}$$

5.2.4. Modelo panel de efectos aleatorios (MEAL)

A diferencia del modelo panel de efectos fijos, en que se permite que cada variable explicativa tenga su propio valor de intercepto (fijo), en este caso del MEAL, se supone que los valores del intercepto son una extracción aleatoria de una población mucho mayor de variables explicativas.

La idea básica es comenzar con el modelo:

$$Y_{it} = \beta_{0i} + \beta_1 X_{1it} + \cdots + \beta_k X_{kit} + U_{it}$$

Donde:
i = 1,2,3,… i-ésima observación de corte transversal
j = 1,2,3,… j-esimo período temporal.

En el modelo panel de efectos aleatorios, la opción para la heterogeneidad es:

$$\boxed{\boldsymbol{\beta_{0i} = \beta_0 + \tau_i}}$$

Donde τ_i = error no observado

En vez de considerar a $\boldsymbol{\beta_{0i}}$ como un valor fijo, se supone ahora que es una Variable Aleatoria, pues el efecto de heterogeneidad se expresa a través de la variable aleatoria $\boldsymbol{\tau_i}$**.** El término $\boldsymbol{\beta_{0i}}$ es una VA con valor promedio igual a $\boldsymbol{\beta_0}$ (sin subíndice "*i*") y que se expresa como:

$$\boldsymbol{\beta_{0i} = \beta_0 + \tau_i} \qquad i = 1,2, ..., n$$

Conceptualmente los individuos son "parte" de un universo más amplio de países (en este caso). Además, si se supone que $\boldsymbol{\tau_i}$ NO está correlacionado con las V. regresoras, el modelo adecuado es el que usa panel con efectos aleatorios.

Entonces, si en el modelo
$Y_{it} = \beta_0 + \beta_1 X_{1it} + \cdots + \beta_k X_{kit} + \varepsilon_i + v_{it}$

definimos el término de error compuesto como: $\boldsymbol{U_{it} = \varepsilon_i + v_{it}}$

El término de error compuesto $\boldsymbol{U_{it}}$ consta de dos componentes: $\boldsymbol{\varepsilon_i}$ que representa el error de corte transversal o error específico de individuos, y por $\boldsymbol{v_{it}}$ que expresa el error aleatorio del modelo.

Supuestos claves del Modelo de Efectos Aleatorios

$\varepsilon_i \sim N(0, \sigma_\varepsilon^2)$
$U_{it} \sim N(0, \sigma_U^2)$
$E(U_{it}, \varepsilon_i) = 0$; $E(\varepsilon_i, \varepsilon_j) = 0$ $(i \neq j)$
$E(U_{it}, U_{is}) = E(U_{it}, U_{is}) = E(U_{it}, U_{js}) = 0$ $(i \neq j;\ t \neq s)$
También $\boldsymbol{U_{it}}$ no está correlacionado con ninguna ***X.***

Los componentes de error no están correlacionados entre sí y tampoco autocorrelacionados en las unidades de series de tiempo ni en las de corte transversal.

Estimación en el MEAL

Para la estimación del modelo panel con efectos aleatorios, se utiliza el método de Mínimos Cuadrados Generalizados, consistente en la utilización del término: una EDG (ecuación en diferencias generalizada) que consiste en utilizar el término:

$$\varphi = 1 - \left[\frac{\sigma_U^2}{(\sigma_U^2 + T\sigma_\varepsilon^2)}\right]^{0.5}$$

Que permite finalmente construir una ecuación en diferencias generalizada:
$(Y_{it} - \varphi\bar{Y}_i) = \beta_0(1 - \varphi) + \beta_1(X_{1it} - \varphi\bar{X}_{1i}) + \cdots + \beta_k(X_{kit} - \varphi\bar{X}_{ki}) + (w_{it} - \varphi\bar{w}_i)$ (5.10)

La ecuación (14.11) es una transformación que utiliza cuasi-desviaciones respecto a la media; aunque no es obvio, los términos de error de esta ecuación, no presentan Autocorrelación. Esta transformación (14.11) permite la "inclusión" de variables explicativas que permanecen constantes en el tiempo.

5.3. Selección del modelo más adecuado

Un elemento que es importante destacar es la elección del mejor modelo panel, lo que finalmente se convierte en la selección del mejor método. Afortunadamente existen algunos procedimientos estadísticos que nos permiten tomar una decisión al respecto. El procedimiento consiste en lo siguiente:

(i) aplicar un contraste de redundancia, a fin de elegir la mejor opción entre un modelo de datos agrupados (MDA) frente aa un modelo de efectos fijos.

El contraste consiste en comparar la hipótesis que los términos constantes son todos iguales, frente a la opción que existen efectos fijos diferenciados en el corte transversal. Formalmente, las hipótesis que se contrastan en este caso son:

Ho: todos los interceptos son iguales (modelo agrupado sería válido).
Ha: al menos un intercepto diferente (modelo MEF es una mejor especificación).

Hipótesis	Estadístico de Prueba	Decisión
Ho: todos los interceptos son iguales (modelo MDA es válido). ***Ha:*** al menos un intercepto es diferente (Modelo de efectos fijos).	$F = \dfrac{(SCR_R - SCR_{NR})/m}{(SCR_{NR})/(nT - k - m)}$ *m = Nº restricciones* *n = Nº individuos* *k = Nº parámetros* *T = tiempo*	Si *Fc > Ft* Entonces rechazo la ***Ho.*** y el modelo panel de efectos fijos es el relevante (no restringido- NR). Caso contrario es válido el modelo restringido- R (*pooled*).

El valor R se asocia a: $Y_{it} = \beta_0 + \beta_1 X_{1it} + \beta_2 X_{2it} + \cdots + U_{it}$ (5.11)

NR se asocia al modelo: $Y_{it} = \beta_0 + \beta_1 X_{1it} + \beta_2 X_{2it} + \cdots + \varepsilon_i + v_{it}$ (5.12)

El valor ***F*** es**:**

$$F = \frac{(SCR_R - SCR_{NR})/m}{(SCR_{NR})/(nT - k - m)}$$

En el ejemplo ilustrativo toma el valor:

$$F = \frac{(5069-3227)/4}{(3227)/(140-6-4)} = \frac{460.5}{24.8} = 18.56$$

Alternativamente, puede utilizarse también:

$$F = \frac{(R_{NR}^2 - R_R^2)/m}{(1 - R_{NR}^2)/(nT - k - m)} \quad (5.13)$$

$$F = \frac{(0.817 - 0.713)/4}{(1 - 0.817)/(140 - 6 - 4)} = \frac{0.026}{1.4x10^{-3}} = 18.47$$

El valor R^2 es el restringido, proviene del modelo de datos agrupados; el valor no restringido de R^2 viene del modelo de datos no-agrupados, y donde el número de restricciones es 4, pues el modelo MDA supone que los interceptos son iguales.

Obviamente, ***Fc*** de 15.3 (para 4 gl del numerador y 129 gl del denominador) expresa uma alta significância estadística. En ambos casos, la regresión restringida (16.3.1) no parece ser válida, por lo que optamos por el modelo panel de efectos fijos (modelo de datos no-agrupados).

(ii) aplicar un contraste entre modelos panel, a fin de elegir la mejor opción entre con datos panel de efectos fijos y otro panel de efectos aleatorios; el constaste corresponde al sugerido por Hausman:

El contraste consiste en comparar la hipótesis que el intercepto principal puede ser descompuesto en un conjunto de valores fijos, frente a la opción que el intercepto principal es una variable aleatoria que expresa la heterogeneidad entre individuos en forma aleatora en torno a un valor promedio de todas las opciones posibles. Las hipótesis que se contrastan en este caso son:

Ho: el modelo MEA sería mejor que MEF.
Ha: el modelo MEF sería una mejor opción, antes que el MEAL.

Hipótesis	**Estadístico de Prueba**	**Decisión**
Ho: *MEAL* superior a *MEF* (pues los efectos aleatorios NO estarían correlacionados con regresoras). **Ha:** *MEAL* no apropiado.	***χ2*** asintótica, proveído por software (***GL***=Nº de parámetros)	Si ***p*** (asociado a ***χ2)*** es menor a *0.05* (5%), entonces el modelo *MEAL* no es apropiado por probable correlación entre efectos aleatorios con las variables regresoras.

En adición a lo anterior, con relación al modelo panel más apropiado, es importante tomar en cuenta lo siguiente:

Modelo panel de efectos fijos (MEF)	**Modelo panel de efectos aleatorios (MEAL)**	**Modelo con datos agrupados (*pooled*)**
El MEF siempre permite obtener	El MEAL es consistente, aunque el verdadero	Cumpliéndose el requisito de no endogeneidad, el

estimadores consistentes, independientemente de si el modelo verdadero a ser usado sea un MEA o un modelo de datos agrupados.	modelo que se debe usar sea el de datos agrupados. Pero si el verdadero modelo que debe usarse fuera MEF, entonces los estimadores del MEA serán inconsistentes.	modelo de datos agrupados ofrece estimadores consistentes.

También es importante tomar nota de las siguientes recomendaciones prácticas:

- Es probable que exista poca diferencia entre los estimadores de los modelos panel cuando tenemos un modelo panel largo (grande la serie de tiempo y pequeño el corte transversal).

- Si tenemos un panel corto (muchas observaciones del corte transversal y pocas en la serie de tiempo), las estimaciones de ambos modelos varían significativamente y pequeño el corte transversal). Se requiere escoger la mejor opción en base a las pruebas del caso, la naturaleza del problema de investigación y la propia experiencia.

Supuestos de los modelos PANEL
(tomado de Wooldridge (2010), Apéndice 14-A)

(1) Para **cada *i*,** el modelo es**:**

$$Y_{it} = \beta_0 + \beta_1 X_{1it} + \beta_2 X_{2it} + \varepsilon_i + \mu_{it} \qquad t = 1, \ldots, T$$

Donde $\boldsymbol{\beta_j}$ son los parámetros a estimar.

(2) Tenemos una muestra en la dimensión transversal.

(3) Para cada "***t***", el valor esperado del error idiosincrático es**:**

$$E(\mu_{it} / X_i, \ \varepsilon_i) = 0$$

(4) Cada variable explicativa cambia en el tiempo (al menos para algún ***i***), y no hay multicolinealidad perfecta entre las regresoras.

Bajo los **04** supuestos arriba, el estimador de efectos fijos es insesgado.

(5) $V(\mu_{it} / X_i, \ \varepsilon_i) = \ V(\mu_{it}) = \sigma_{it}^2$ **,** *para todo* ***t*** *= 1, ... T*

(6) Para todo ***t* ≠ *s*,** los errores idiosincráticos no están correlacionados entre sí.

Bajo los supuestos 1-6 el estimador de efectos fijos de los *coeficientes* $\boldsymbol{\beta_j}$ es un estimador lineal insesgado óptimo.

(7) $\mu_{it} \backsim NI(0, \sigma_\mu^2)$, condicionado a $\boldsymbol{Xi}$ y el error no observado $\boldsymbol{\varepsilon_i}$.

Con el supuesto **7,** los estadísticos ***"t"*** y ***"F"*** tienen distribuciones exactas; sin este supuesto se tienen que trabajar con distribuciones asintóticas, aprox.

(8) Si $\boldsymbol{\varepsilon_i}$ NO está correlacionado con las regresoras, el modelo adecuado es el de Efectos Aleatorios. Caso contrario es el de efectos fijos.

5.4. Modelos Panel con datos no balanceados

Por diversos motivos, en conjuntos de datos de panel es común que falten datos; ocurre a menudo en paneles "largos". Las modificaciones, en el processo de estimación, para permitir tamaños de grupos desiguales es relativamente simple, requiere de ponderación en la estimación de los coeficientes. Afortunadamente la prática es más sencilla que la teoria, pero el tratamento depende de si se trata de un modelo de efectos fijos o uno de efectos aleatorios.

Primero que el tamaño de muestra es $\sum_{i=1}^{n} T_i$ (en lugar de *nT*). T_i sería el número de períodos para la unidad "*i*" de CT. Ello implica pequenos cambios en el cálculo de S^2 y V(b), así como em el estadístico "*F*". Segundo, los promedios de los grupos deben basarse en T_i que varía entre los grupos. Las medias totales para los regresores son:

$$\bar{\bar{X}} = \frac{\sum_{i=1}^{n}\sum_{t=1}^{T_i} X_{it}}{\sum_{i=1}^{n} T_i} = \frac{\sum_{i=1}^{n} T_i \bar{X}_i}{\sum_{i=1}^{n} T_i} = \sum_{i=1}^{n} w_i \bar{X}_i; \text{ Donde } w_i = \frac{T_i}{\sum_{i=1}^{n} T_i} \qquad (5.14)$$

Caso 1. Modelo de Efectos Fijos (MEF) no balanceado

En caso del MEF, una modificación para permitir tamaños de grupos desiguales es simple y usa también Mínimos Cuadrados de Var. Ficticias, pero utilizando también la filosofía de la transformación "intra-grupos" (Greene, 2012).

El tamaño muestral completo es $\sum_{i=1}^{n} T_i$ en vez de $nT \rightarrow$ pequeñas modificaciones en los cálculos de S^2, $V(\hat{\beta})$ y el estadístico "*F*".

Las medias de los grupos de *CT* deben basarse en T_i (que varía entre grupos). Entonces, para estimar $\hat{\beta}$ se usa la SC intra-grupos- en lugar de *(X´X)*:

$$(X'M^oX) = \sum_{i=1}^{n} X_i'M_i^oX_i = \sum_{i=1}^{n}\left(\sum_{t=1}^{Ti}(X_{it}-\bar{X}_{i.})(X_{it}-\bar{X}_{i.})'\right)$$

Entonces:

(1°) $\hat{\beta} = (X'M^oX)^{-1}(X'M^oY)$

(2°) $V(\hat{\beta}) = S^2(X'M^oX)^{-1}$.

En la práctica, el método sigue siendo MC con *VF*. En E-views es importante generar una variable "identificadora" (por ejemplo *CELLID* en E-views).

Caso 2. Modelo de Efectos Aleatorios (MEAL) no balanceado

Los paneles incompletos añaden un nuevo nivel de dificultad al Modelo de Efectos Aleatorios:

(1°) Existe Heteroscedasticidad entre grupos originada por el diferente tamaño de los grupos; entonces es necesario utilizar τ_i (específico) en la EQ en diferencias generalizada, es decir:

$(Y_{it}-\tau_i\bar{Y}_i) = \beta_1(1-\tau_i)+\beta_2(X_{2it}-\tau_i\bar{X}_{2i})+\cdots+\beta_k(X_{kit}-\tau_i\bar{X}_{ki})+(w_{it}-\tau_i\bar{w}_i)$ (5.15)

En lugar de:

$(Y_{it}-\tau\bar{Y}_i) = \beta_1(1-\tau)+\beta_2(X_{2it}-\tau\bar{X}_{2i})+\cdots+\beta_k(X_{kit}-\tau\bar{X}_{ki})+(w_{it}-\tau\bar{w}_i)$ …

Matricialmente, se usa: $\hat{\beta} = (X'\Omega X)^{-1}(X'\Omega Y)$

(2°) Aun cuando el estimador MCVF (apropiadamente calculado) ofrece todavía un estimador consistente de σ_ϵ^2, pero necesitamos una estimación adecuada de σ_u^2; entonces utilizamos el estimador de medias de grupos:

O, lo que es lo mismo: $\varepsilon_{*i} = (\mu_i+\bar{\varepsilon}_{i.}) = \mu_i + \frac{\sum_t^{Ti}\varepsilon_{it}}{T_i}$

Pero siendo los errores μ_i heteroscedásticos, entonces:

$V\left[\mu_i + \frac{\sum_t^{Ti}\varepsilon_{it}}{T_i}\right] = \sigma_u^2 + \frac{\sigma_\varepsilon^2}{T_i} = k_i^2$

En este contexto la consistencia es todavía aplicable cuando "n" tiende a infinito, (no Ti); entonces el estimador de la varianza a partir de medias de grupo es un estimador consistente de:

$$\sigma_u^2 + \sigma_\varepsilon^2 \, plim \frac{1}{n} \sum_{i=1}^{n} \frac{1}{T_i}$$

o de: $\sigma_u^2 + \sigma_\varepsilon^2 (plim Q_n) = \bar{k}^2$

Se necesita un supuesto adicional: si Ti se distribuye aleatoriamente entre los individuos alrededor de la media de T, entonces $(plim Q_n) = \dfrac{1}{T}$

Entonces podemos continuar con la estimación, usando Mínimos Cuadrados Generalizados Factibles (MCGF).

Ejercicios del capítulo V: Modelos con datos panel

5.1. ¿Cuándo resultan relativamente inapropiados los datos panel?

> ***Sugerencias:***
> *Revisar la sección 16.1 del libro de Gujarati & Porter.*

5.2. En un conjunto de datos que incluya todos los departamentos del Perú en la dimensión del corte transversal y dos puntos de tiempo ¿Qué tipo de modelo panel sería y cuál podría ser el más apropiado en cuanto a la estimación? ¿Por qué?

> ***Sugerencias:***
> *Es un panel largo. Conceptualmente un modelo de efectos fijos podría ser el más apropiado.*

5.3. (Adaptado del ejercicio 14.4 de Wooldridge, 2009). Para determinar los efectos del conocimiento de la "tasa de graduación" de los universitarios peruanos (***EXT***) en relación al porcentaje de postulación universitaria a sus respectivas universidades; se recopilan datos de las postulaciones a partir de una muestra de 50 universidades peruanas, para los años 1985, 1990 y 1995.

La variable dependiente es la postulación universitaria (*APP*), que se mide como Nº de postulantes al año. Especificar una ecuación semilogaritmica que permita estimar los efectos de la "tasa de graduación" (*EXT*) y la pensión semestral (*COSTO*- en miles de soles) en el ***cambio porcentual*** en las solicitudes de postulación (*APP*).

(a) Especifique una ecuación PANEL, en forma unidireccional y con el tiempo expresado con variables ficticias explícitas.

(b) ¿Cómo serían el sentido del cambio de los coeficientes antes indicados?

> ***Sugerencias***:
> *Un modelo de efectos no observados sería:*
> $Ln(app_{it}) = \delta_0 + \delta_1 d90_t + \delta_2 d95_t + \beta_1 (ext)_{it} + \beta_2 (costo)_{it} + a_i + \mu_{it}$
> ***t**=1,2,3*
>
> *Es interesante que un incremento en 1% en la "graduación" podría llevar a una mejora de la postulación en 100.($\boldsymbol{\beta_1}$). Es bastante probable que $\boldsymbol{a_i}$ tenga correlación con la tasa de graduación y con el costo de los estudios, por lo que no podría ser aconsejable usar un modelo de datos agrupados.*

5.4. Sea una ecuación que permite medir el efecto del precio de los huevos (soles x docena) en la demanda de los mismos (*Y*, en millones).

Si se agrupan todas las observaciones y se tienen los resultados MCO de la siguiente ecuación:

$\hat{Y} = \quad 3132.25 - 22.895(precio)$ (1)

Estad. "t" (5.32) (-3.29) r2= 0.69 *d*=2.03

Luego se ha utilizado un modelo panel de efectos fijos con v. ficticias (D=1 en 1991) habiéndose obtenido:

$\hat{Y} = \quad 3154.50 - 34.69D_i - 22.94(precio)$ (2)

Estad. "t" (5.32) (-3.29) (3.10) r2= 0.58 *d*=2.04

¿Qué podría concluir de los resultados de las dos ecuaciones?

***Sugerencias**:*
El coeficiente de la variable dummy para 1991 no es estadísticamente significativo; sugiere que los interceptos de ambos períodos son estadísticamente semejantes.

5.5. Conteste lo siguiente:

(i) ¿En qué consiste el modelo panel de efectos fijos?, ¿Cómo los datos de PANEL presentan la dimensión de espacio? ¿Cómo es que el MEF permite el tratamiento de ambas dimensiones?

(ii) ¿Qué se entiende por MEAL? ¿En qué se diferencia del MEF en cuanto al manejo de las unidades de CT? ¿Cuándo resulta + apropiado el MEAL en relación al MEF?

***Sugerencias**:*
(i) Modelo con datos en dos dimensiones: CT y tiempo. En un MEF generalmente se incluye la dimensión de corte transversal, pero se puede permitir ambas dimensiones.

(ii) En el MEAL asumimos que el intercepto de cada microunidad es una VA con media y varianza de la propia ecuación. No se necesitan "n" interceptos dummy para microunidades. Si el término de error y las variables regresoras son independientes, el MEAL sería el modelo más apropiado.

Ejercicios para uso de software.

5.6. Este ejercicio usa datos de la Tabla 1.1 y el fichero *Table_1.1* de Gujarati & Porter (2010), ejercicio16.11. Los datos se refieren a un panel "corto" con solo dos años en el tiempo y 50 estados norteamericanos en el corte transversal (pero no está ordenado acorde a una base de datos panel).

(a) Siendo ***Y***= huevos producidos (en millones) y ***X***= precio de los huevos (centavos por docena). Calcule el modelo $Y_i = \beta_1 + \beta_2 X_i + \mu_i$ para los años 1990 y 1991 pero por separado

Sugerencias:

Año	*Intercepto*	*Pendiente*	R^2	*D*
1990	*3118.48*	*-22.498*	*0.0834*	*1.98*
	(3.572)	*(-2.089)*		
1991	*3149.36*	*-22.3485*	*0.0972*	*1.9*
	(3.884)	*(-2.274)*		

* *Los datos entre paréntesis son los valores calculados del estadístico* ***"t"***.

(b) Arregle los datos en forma de un panel, es decir agrúpense las observaciones para los dos años y estime la regresión agrupada (MDA). ¿Qué supuestos se están haciendo al agrupar los datos?

Sugerencias:
La regresión con datos agrupados (pooled) resulta en lo siguiente:

Eggs = 3123.26 - 22.895Price
Estad "t" (5.33) (- 3.117)
r^2*=0.0903* ***d****=2.0037*

(c) Utilícese el modelo de efectos fijos inter-grupos, haciendo la distinción entre los dos años y presente los resultados de la regresión

Sugerencias:
Haciendo D=0 para 1990 y D=1 para 1991, los resultados de la regresión son:

Eggs = 3153.08 – 34.698D – 22.94Price
Estad."t" (5.08) (-0.109) (-3.103)
r^2*=0.0903 d=2.0047*

El coeficiente "dummy" para 1991 no es estadísticamente significativo.

(d) Si son 50 estados o regiones involucradas en el Modelo ¿Puede utilizarse el modelo de efectos fijos, haciendo la distinción entre los 50 Estados? ¿Por qué?

Sugerencias: *revisar la sección 16.5 del libro de Gujarati & Porter (2010).*

(e) ¿Tendría sentido distinguir el efecto por estado y el efecto por año? De ser así, ¿cuántas variables dicotómicas se tendrían que introducir?

Sugerencias: *Para tal efecto sería complicado el manejo de los grados de libertad.*

5.7. Este Ejercicio ha sido adaptado del libro de Gujarati & Porter, que usa el fichero "Table1_1.2" (ejercicio 16.13) referido a un panel largo. Se refiere a los determinantes de la inversión de cuatro grandes empresas transnacionales: General Electric (GE), US Steel (US), General Motors (GM) y Westinghouse (WEST). Los datos corresponden a un set de varios años (1935- 1954) estudiados por un Economista apellidado Grunfeld, mediante uso de un modelo panel con cuatro unidades de corte transversal y 20 años. Las variables regresoras son el valor real de la empresa (VR) y el capital social (CS), mientras que la variable dependiente es el nivel de inversión (INV).

(a) Estímese la función de inversión Grunfeld analizada para las cuatro grandes empresas transnacionales GE, GM, US y WS, de manera individual, utilizando las dos variables regresoras antes indicadas (VR y CS).

***Sugerencias**:* *La idea es estimar regresiones individuales, MDA, MEF con Var. Ficticias, a fin de comparar y escoger la mejor opción.*

(b) Agrupe los datos de todas las empresas y estime la función de inversión Grunfeld por MCO.

Sugerencias: *Misma sugerencia previa.*

(c) Estime la función de inversión con el modelo de efectos fijos (MEF) y compare los resultados con la regresión agrupada que se estimó en (b) ¿Cuál es la conclusión sobre la selección de un modelo u otro?

Sugerencias: Se puede probar cuál de las dos opciones es mejor (MDA o MEF)- Revisar páginas 597-598 del libro de Gujarati-Porter (2010).

(d) ¿Cómo decidiría entre la regresión MDA y la regresión del MEF? (realice los cálculos).

Sugerencias:
Modelo restringido es el modelo MDA. Modelo no restringido: individuales sumados (se necesitan las SCR).

No es conveniente usar los datos fusionados (es mejor el modelo de Efectos Fijos).

5.8. Use el fichero "Philips", que contiene datos relevantes (pero separados por país) para la evaluación del modelo siguiente:

$$Y_{it} = \beta_1 + \beta_2 X_{2it} + U_{it} \qquad (1)$$

Y: tasa de desempleo civil (%)
X: salarios, por hora, en el sector manufacturero (US$)
i: unidades de corte transversal (tres países: UK, USA y Canadá)
t: serie temporal.

(a) A priori ¿cuál es la relación esperada entre ***Y*** y ***X***? ¿Por qué?

Sugerencia:
A priori se esperaría una relación inversa entre las dos variables, pues si el desempleo es alto habrá menor presión para incrementos salariales, asumiendo el resto de cosas constante.

(b) Presente resultados de la regresión del modelo (1) para cada país y Conjuntamente. Concluya.

(c) Organice la información en formato panel y estime el modelo agregado.

Sugerencias:

País	***Intercepto***	***Slope***	***R^2***	***SCR***	***d***
Canadá	*85.829*	*-0.729*	*0.048*	*5,372*	*0.0088*
(t)	*(3.871)*	*(-0.295)*			
UK	*156.441*	*-9.119*	*0.425*	*7,856*	*0.3591*
(t)	*(6.695)*	*(-3.646)*			
USA	*152.466*	*-9.668*	*0.542*	*3,375*	*0.4910*
(t)	*(10.925)*	*(-4.616)*			
MDA	*132.389*	*-6.341*	*0.3365*	*1,941*	*0.2440*
(t)	*(13.475)*	*(-5.424)*			

(d) Calcule el modelo panel de efectos fijos con variables ficticias y el modelo panel de efectos aleatorios. Compare y concluya.

Sugerencias:
Para saber cuál es el mejor modelo, debemos utilizar el test de Hausman.

5.9. Ejercicio adaptado del libro de Wooldridge (pp. 485), que corresponde al fichero *WAGEPAN.* Consiste en una muestra de 545 hombres que trabajaron cada año, de 1980 a 1987. Algunas variables del conjunto varían con el tiempo (experiencia- *Exper*, afiliación sindical- *Union*, y estado civil- *Married*, son las más importantes) y pueden ser utilizadas para explicar el comportamiento de la variable dependiente (salario- *Wage*). Se solicita lo siguiente:

(a) Cuáles podrían ser dos factores que no cambiarían en el tiempo

Sugerencias:
Factores no cambiantes (en el tiempo) son, ejemplo, raza y educación.

(b) Comparar dos regresiones: MDA y Panel MEF con el siguiente modelo semilogaritmico (Log-Lin)**:**

$Log(Wage) = F\ (Exper, Exper^2, Married, Union)$

Donde
Wage = salarios de varones (US$ por hora)
Exper = años de experiencia
Married = ficticia (casado o no)
Union = ficticia (sindicalizado o no)

En base a los dos modelos anteriores discuta la significancia e importancia del impacto de ambos tipos de modelo (coeficientes de regresión, R^2 ajustado, Durbin-Watson, variabilidad de datos). Elija la mejor opción de estimación.

Sugerencias:
No es conveniente usar los datos fusionados, es mejor el modelo MEF.

(c) Utilice el mismo modelo en (ii) para regresionar un modelo PANEL de efectos aleatorios. Compare los resultados con el modelo de efectos fijos y pruebe la pertinencia de utilizar un modelo PANEL de Efectos fijos frente a otro de efectos aleatorios.

Sugerencias: Leer argumentos expresados en pp. 491-492 del libro de Wooldridge

5.10. El siguiente modelo intenta explicar los determinantes del precio de alquileres de habitaciones en ciudades universitarias (n = 64). El fichero *RENTAL* (de Wooldridge, 2010) incluye el modelo siguiente:

$$Ln(RENT_{it}) = \beta_0 + \delta_0 Y90_t + \beta_1 Ln(POP_{it}) + \beta_2 Ln(AVGINC)_{it} + \beta_3 (PCTSTU)_{it} + a_i + \mu_{it}$$

Donde:
RENT = precio de alquileres de habitaciones en ciudades universitarias
Y90 = dummy año 1990 contra 1980
POP = población de la ciudad
AVGINC = nivel de renta media
PCTSTU = % de la población estudiantil (respecto al total).
a_i = error no observado
μ_{it}= error idiosincrático.

(a) Estimar la ecuación mediante MCO para datos agrupados; presentar los resultados en forma habitual. Interpretar la estimación del parámetro de la variable ficticia de tiempo y la de la variable *PCTSTU*.

Sugerencias: *Los resultados mostrarían que el estimador asociado a D90 es significativo, lo que implica que (todo lo demás constante), las rentas nominales crecen aproximadamente 26% en 10 años. El coeficiente de PCTSTU significa que un incremento (de PCTSTU) en 1% incrementa la renta en un 0.5% aunque poco importante (magnitud).*

(b) ¿Serán válidos los errores estándar encontrados en (a)? Justifique su respuesta.

Sugerencias: *Si a_i> 0, entonces el error que cruza los dos puntos de tiempo, para cada ciudad, están positivamente correlacionados, lo que invalida MCO y también las inferencias estadísticas asociadas.*

(c) Estimar un modelo panel de efectos fijos y comparar con el modelo agrupado previo en (a).

Sugerencias:
$Log(RENT) = -0.386\ (Y90) + 0.072\ Log(POP) + 0.310\ Log(AVGINC) + 0.0112\ PCTSTU$
$(EE)\quad (.037)\qquad (.088)\qquad (.066)\qquad (.0041)$
$T=2,\ n = 64,\ R^2 = 0.72$

(d) En el modelo de la parte (c), reporte los EE robustos de WHITE y compare.

Sugerencias:
Hay que hacerlo con uso de software econométrico. Note que heteroscedasticidad es generalmente un problema en este caso, más no necesariamente la autocorrelación (AC).

5.11. Este ejercicio ha sido adaptado del libro de Wooldridge (2009), ejercicio C13.3. Usa datos del fichero KIELMC**,** con la pretensión de evaluar el efecto que tiene la ubicación de un incinerador de basura, en relación con el precio de las casas alrededor, siendo las variables definidas las siguientes:

PRICE= precio de la casa (US$)- variable dependiente.
NEARINC= variable ficticia (1 cerca del incinerador, 0 lejos)
AGE= antigüedad de la casa (años)
LAND= área total (m2)
ROOMS= Nº de habitaciones
AREA= área construida (m2)
BATHS= Nº de baños
INTST= distancia de la casa a la avenida principal.
DIST= distancia de cada casa a un incinerador (pies).

Considere el modelo:

$$Log(PRICE) = \beta o + \delta o Y81 + \beta 1\ Log(DIST) + \delta 1\ [Y81 * Log(DIST)] + U$$

(a) Si la construcción del incinerador reduce el valor de las casas más cercanas a su ubicación, ¿cuál sería el signo esperado de $\boldsymbol{\delta 1}$**?** **¿**Qué significa si $\beta 1 > 0$**?**

Sugerencias:
Con todo lo demás constante, las casas más alejadas del incinerador deberían tener más valor.

(b) Estime el modelo del inciso (a) e informe los resultados de la manera acostumbrada. Interprete el coeficiente de *Y81* Log(DIST)*. ¿Qué puede concluir?

Sugerencias: *La ecuación estimada es*
$Ln(\widehat{PRICE})$ = *8.06 + 0.011 (Y81) + 0.317 Log(DIST) + 0.048 [Y81*Ln(DIST)]*
(EE) (0.51) (0.81) (0.052) (0.082)
$R^2=0.396$ $R^2a=0.39$ $n=321$

(c) Al modelo previo añada las variables regresoras *AGE, AGE2, ROOMS, BATHS, LOG(INTST), LOG(LAND) y LOG(AREA).* En estas circunstancias, ¿cuál es su conclusión acerca del efecto del incinerador sobre los valores de las casas?

Sugerencias:
Cuando añadimos la lista de las características de la vivienda en el modelo de regresión, el coeficiente asociado a **[***Y81*Log(DIST)***]** *es 0.062 (con EE=0.05).*
Ha: $\widehat{\delta 1} > 0$ *está próximo a 0.108, que es una significancia con α=0.10.*

(d) ¿Por qué el coeficiente de *LOG(DIST)* es positivo y estadísticamente significativo en el inciso (*b*) pero no en el inciso (c)?.

5.12. Este ejercicio ha sido adaptado del Ejercicio C-14.2 del libro de Wooldridge (pp 498). Usar los datos anuales (período 1982 a 1987) del fichero **CRIME4** que corresponde a un Estado de USA, el propósito es explicar la tasa de delincuencia (*CRMRTE*). Las variables claves son:

Variable	**TIPO**
CRMRTE	Índice de delincuencia (delitos cometidos por habitante)
PRBARR	Probabilidad estimada de ser detenido
PRBCONV	Probabilidad estimada de ser condenado después de ser detenido
PRBPRIS	Probabilidad de cumplir una sentencia en prisión
AVGSEN	Duración media en el cumplimiento de una sentencia
POLPC	Nº de policías por habitante

Nota: algunos de los factores que pudieran ubicarse en el error no observado son: las actitudes a la delincuencia, los registros históricos y las normas de denuncias del crimen, etc.

Usar una función doble logarítmica (en todas las variables) para lo siguiente:

(a) Procesar un modelo de efectos fijos (MEF) que contraste sólo diferencias cruzadas frente a otro MEF que contraste diferencias cruzadas más diferencias temporales.

(b) Use el contraste de Hausman para determinar el mejor método o modelo panel.

(c) Con el mejor modelo hallado en (b), use Errores Estandar robustos y concluya.

(d) Con el mejor modelo hallado en (b), interprete el coeficiente de la "probabilidad de ser detenido".

Sugerencias:
(a) Con un MEF en una sola dimensión del panel, para diferencias cruzadas, frente a otro MEF que contrasta diferencias cruzadas más diferencias temporales, básicamente los resultados son los mismos.
(b) Procese la prueba de Hausman para optar entre el MEF y el MEAL.
(c) Cambian algunos Errores Estándar, pero no la significancia de los estimadores.
(d) Probabilidad de ser detenido (PRBARR): el coeficiente es estadísticamente significativo (-0.38), implica que un incremento de 10% en la probabilidad de ser detenido hace disminuir el índice de delincuencia, en un 3.8% (es inelástica).

5.13. Use el archivo *MARACUYA* y estime la siguiente función de producción mediante un modelo panel para la región LAMBAYEQUE, que tiene las siguientes provincias: Lambayeque, Chiclayo y Ferreñafe. Período anual 1997-2009.

$$Pdmara_{it} = \beta_1 + \beta_2 Scmara_{it} + \beta_3 Prmaya_{it} + \beta_4 Tmin_{it} + \beta_5 Tmin2_{it} + \beta_6 pp_{it} + \beta_7 pp2_{it} + AR(1) + u_{it}$$

Donde *Pdmara* es la producción de maracuyá (TM), *Scmara* es la superficie cosechada (Ha.), *Prmara* es el precio real de maracuyá, *Tmin* y *pp* son la temperatura minima y la precipitación respectivamente. También se agregó sus términos cuadráticos de estas Variables regresoras y un proceso autorregresivo de orden 1 [*AR*(1)], para evitar problemas de autocorrelación.

(a) Estime, con la especificado indicada, el Modelo de datos agrupados o de coeficientes constantes.

(b) Especifique y estime todas las formas posibles de representar un Modelo de efectos fijos

(c) Especifique y estime un modelo de efectos aleatorios (MEAL)

(d) Realice las pruebas de Mínimos Cuadrados Restringidos (Prueba *F*), Hausman, PRV y elija el mejor modelo.

Sugerencia:
(c) No se puede estimar efectos aleatorios cuando el número de unidades transversales es menor al número de coeficientes usados en la regresión.

CAPÍTULO VI. MODELOS DE ECUACIONES MÚLTIPLES

6.1. Antecedentes

En términos de la vida real, los modelos uni-ecuacionales tienen dos limitaciones: (i) la relación es uni-causal, o sea no hay retroalimentación de la Variable Dependiente; (ii) no explican un comportamiento causal de las regresoras, en relación con otras variables exógenas.

Los modelos multi-ecuacionales o de ecuaciones múltiples a menudo nos ayudan a superar las limitaciones indicadas, son a menudo resueltas en forma simultánea; el comportamiento de las variables se determina en forma conjunta. Pueden incluir identidades, además de ecuaciones de comportamiento.

A menudo el método de MCO es lo más adecuado en modelos uni-ecuacionales; sin embargo, ya no lo es en el caso de ecuaciones múltiples. En este último caso, los procesos de especificación, identificación y estimación son más complejos.

En el caso de un modelo de ecuaciones múltiples, un ejemplo simple, real y sencillo, usado para efectos didácticos es el de un mercado simple de oferta y demanda:

$$Q_t^s = \alpha_1 + \alpha_2 P_t + \alpha_3 P_{t-1} + \varepsilon_t \qquad (1)$$

$$Q_t^d = \beta_1 + \beta_2 P_t + \beta_3 Y_t + \mu_t \qquad (2)$$

$$Q_t^s = Q_t^d \quad \text{(condición de equilibrio)} \quad (3)$$

Donde:
Variables endógenas: $Q_t^s \; y \; Q_t^d$ y P_t (variables determinadas dentro del sistema).

Variables predeterminadas: 2 (ayudan a causar el movimiento de las V. endógenas).
- Endógenas rezagas: P_{t-1}
- Exógenas: Y_t (determinadas plenamente fuera del sistema).

Supongamos ahora que tenemos el modelo (en desviaciones respecto a la media):

$$q_t^s = \alpha_2 p_t + \varepsilon_t \qquad (1)$$

$$q_t^d = \beta_2 P_t + \beta_3 Y_t + \mu_t \qquad (2)$$

Éste es llamado un modelo en su Forma Estructural (MFE), cuya forma está dada por la teoría subyacente: variables endógenas están en el lado izquierdo y si es simultáneo, también están en como V. predeterminadas en el lado derecho.

Un reordenamiento que permita solo las variables endógenas como función de solo variables predeterminadas, tendremos el modelo en Forma Reducida (MFR):

$$q_t = \frac{\beta_3\alpha_2}{\alpha_2-\beta_2}y_t + \frac{\beta_4\alpha_2}{\alpha_2-\beta_2}w_t + \frac{\alpha_2\mu_t-\beta_2\varepsilon_t}{\alpha_2-\beta_2} = \pi_{12}y_t + v_{1t} \qquad (3)$$

$$p_t = \frac{\beta_3}{\alpha_2-\beta_2}y_t + \frac{\beta_4}{\alpha_2-\beta_2}w_t + \frac{\mu_t-\varepsilon_t}{\alpha_2-\beta_2} = \pi_{22}y_t + v_{2t} \qquad (4)$$

En forma resumida:

$$q_t = \pi_{12}Y_t + v_{1t} \qquad (5)$$

$$p_t = \pi_{22}Y_t + v_{2t} \qquad (6)$$

Concentrándonos en la ecuación de oferta (Forma Estructural 1) y supongamos uso de MCO, entonces:

$$\hat{\alpha}_2 = \frac{\sum q_t p_t}{\sum p_t^2} \qquad (7)$$

Substituyendo q_t, de la ecuación (1), en (7):

$$\hat{\alpha}_2 = \frac{\sum(\alpha_2 p_t+\varepsilon_t)p_t}{\sum p_t^2} = \alpha_2 + \frac{\sum \varepsilon_t p_t}{\sum p_t^2}$$

No hay seguridad de $\frac{\sum \varepsilon_t p_t}{\sum p_t^2} = 0 \longmapsto$ entonces el estimador $\hat{\alpha}_2$ no sería ni insesgado ni consistente. MCO podría sobreestimar al verdadero parámetro. Entonces en una ecuación con variables endógenas en el lado derecho de otra ecuación, entonces hay un problema de simultaneidad (endogeneidad).

6.2. Especificación del Modelo de Ecuaciones Múltiples

Suponga que conocemos la forma reducida de un sistema de ecuaciones, ¿es suficiente para permitirnos discernir el valor de los parámetros de la forma estructural? NO, esta limitación se llama problema de la Identificación.

Una vez especificado un modelo estructural, debemos verificar si podemos obtener el conocimiento de los parámetros estructurales, una vez que se ha estimado la forma reducida. En el modelo de mercado si es posible las ecuaciones de oferta y demanda si se conoce *P* y *Q*.

En el modelo estructural I:

$$Q_t^s = \alpha_1 + \alpha_2 P_t + \varepsilon_t \qquad (1)$$
$$Q_t^d = \beta_1 + \beta_2 P_t + \mu_t \qquad (2)$$

Es claro que hay un problema de identificación, pues a partir de la forma reducida, no es posible establecer los parámetros en cada ecuación, y por tanto no es posible establecer una única ecuación de oferta y otra de demanda. Solo en casos de uso del modelo para fines de predicción, no se necesita de la identificación; las ecuaciones de forma reducida pueden usarse en forma directa.

En el modelo estructural II:

$$Q_t^s = \alpha_1 + \alpha_2 P_t + \varepsilon_t \qquad (3)$$

$$Q_t^d = \beta_1 + \beta_2 P_t + \beta_3 Y_t + \mu_t \qquad (4)$$

Hay un problema de identificación parcial, pues a partir de la forma reducida solo es posible establecer los parámetros en la ecuación de oferta, mientras no ha sido así en el caso de la demanda; por tanto, no es posible establecer una única ecuación de oferta y otra de demanda. Gráficamente, los cambios en demanda determinan equilibrios que a su vez permiten la determinación de la curva de oferta. Nótese que es el movimiento de Y en el tiempo (a lo largo de observaciones en un análisis de corte transversal) el que es necesario para la identificación de la ecuación de oferta.

En el modelo estructural III:

$$Q_t^s = \alpha_1 + \alpha_2 P_t + \alpha_3 T_t + \varepsilon_t \qquad (5)$$

$$Q_t^d = \beta_1 + \beta_2 P_t + \beta_3 Y_t + \mu_t \qquad (6)$$

Acá, en adición a lo previo (II), la participación de la temperatura (T) en la ecuación de oferta (5) y la demanda no depende de la temperatura, por lo que ello nos permite también identificar la demanda: oferta y demanda identificadas. Acá el movimiento de los valores de equilibrio de P y Q es complejo, pero si se dispone de suficientes datos podemos determinar oferta y demanda en forma única.

En el modelo estructural IV:

$$Q_t^s = \alpha_1 + \alpha_2 P_t + \alpha_3 T_t + \varepsilon_t \qquad (7)$$

$$Q_t^d = \beta_1 + \beta_2 P_t + \beta_3 Y_t + \beta_4 R_t + \mu_t \qquad (8)$$

En este caso la curva de demanda cambia con el tiempo como resultado de cambio en las dos variables. La ecuación de oferta está sobre-identificada debido a que hay dos variables exógenas en el sistema de ecuaciones de Oferta-Demanda excluidas de la ecuación de oferta: o sea hay dos formas en las que se pueden obtener los parámetros estructurales a partir de los parámetros de la forma reducida.

Un procedimiento que ayuda mucho, pero no es tratado aquí: es el del sistema de ecuaciones; se obtiene la MATRIZ y luego las condiciones necesarias (condición de orden) y las condiciones suficientes (condición de rango) de dicha matriz[26]. Se supone el software lo usa de manera muy rápida y sencilla. Pero, aunque el "orden" estricto y las condiciones de "rango" pueden satisfacerse con un proceso así, es improbable que el modelo tenga mucho poder predictivo, dado que algunas o todas las variables agregadas pueden tener poco efecto en las variables endógenas correspondientes.

6.3. Condición de Simultaneidad

Cuando está presente la simultaneidad, una o más de las variables explicativas (en una determinada ecuación) será endógena y estaría correlacionada con el término de error; por lo tanto, el modelo correcto no sería un modelo uni-ecuacional, sino otro de ecuaciones simultáneas.

Si no hay simultaneidad, el método de MCO sería apropiado para la estimación de los parámetros del modelo. Sin embargo, si hubiera simultaneidad, los estimadores MCO serían sesgados e inconsistentes; mientras que el uso de VI se hace necesario pues aportaría consistencia y eficiencia de los estimadores.

Suponga

$q_t^s = \alpha_2 p_t + \varepsilon_t$ (1)
$q_t^d = \beta_2 p_t + \beta_3 y_t + \beta_4 w_t + \mu_t$ (2) Forma Estructural

Por la ecuación (2) sabemos que y_t y w_t son exógenas

$q_t = \pi_{12} y_t + \pi_{13} w_t + v_{1t}$ (1)
$p_t = \pi_{22} y_t + \pi_{23} w_t + v_{2t}$ (2) Forma reducida

Ho: p_t y ε_t no tienen correlación (no simultaneidad).

Ha: p_t y ε_t si tienen correlación, entonces se requiere uso de variables instrumentales y no usar MCO

Procedimiento para realizar el contraste de Hausman:

(i) Uso de la forma reducida (ii) para estimar $p = f(y, w)$ y obtener los residuales v_{2t}

Es decir estimamos:

[26] Detalles de las condiciones matriciales para identificación de un modelo de ecuaciones múltiples puede encontrarse en el Apéndice 12.1 del libro de Pyndick y Rubinfeld (2001).

$\hat{p}_t = \hat{\pi}_{22}y_t + \hat{\pi}_{23}w_t$ (3a) y $p_t = \hat{p}_t + \hat{v}_{2t}$ (3b)
(ii) Realizamos la regresión q=f($\hat{p}_t$,$\hat{v}_{2t}$) y entonces una prueba de "*t*" en el coeficiente de la variable residual $\hat{v}_{2t}$

Reemplazando (3b) en la ecuación (1) estructural:

$$q_t = \alpha_2(\hat{p}_t + \hat{v}_{2t}) + \varepsilon_t$$
$$q_t = \alpha_2(\hat{p}_t) + \alpha_2(\hat{v}_{2t}) + \varepsilon_t$$

Es decir, estimamos la ecuación:

$$q_t = \alpha_2(\hat{p}_t) + \delta(\hat{v}_{2t}) + \varepsilon_t \qquad (4)$$

Para la simultaneidad se requiere la prueba "fácil" mediante el reemplazo de (3b) en (4):

$$q_t = \alpha_2(p_t - \hat{v}_{2t}) + \delta(\hat{v}_{2t}) + \varepsilon_t$$
$$q_t = \alpha_2(p_t) + (\delta - \alpha_2)\hat{v}_{2t} + \varepsilon_t$$

Bajo la hipótesis nula (Ho) de No simultaneidad, la correlación entre p_t y ε_t tiende a cero cuando crece el tamaño de muestra.

Implica que $\delta = \alpha_2$, de modo que el coeficiente de $\hat{v}_{2t}$ debería ser CERO.

Ejemplo de Simultaneidad

Suponga el modelo:

$$EXP = \beta_1 + \beta_2(AID) + \beta_3(INC) + \beta_4(POP) + \varepsilon_t$$

Donde:
EXP= gastos públicos de gobiernos estatales y locales USA
AID= nivel de subsidios de apoyo federal en USA
INC= nivel de ingresos estatales
POP= población estatal.

Suponga también que:

$$AID = \delta_1 + \delta_2(EXP) + \delta_3(PS) + \mu_t$$

Es decir, *AID* está determinada por el nivel de gasto público (*EXP*). La variable *AID* puede estar determinada por la misma *EXP* al igual que por la población de niños de primaria y secundaria.

Para probar la simultaneidad con respecto a la variable *AID* procedemos así:

(i) Regresión de la variable *AID* = *f* (exógenas). Es la forma reducida. Si *PS* es la población de niños en escuelas primarias y secundarias, que funcionaría como instrumento en el caso de endogeneidad:

$\hat{w} = AID - 41.61 + 0.00036(INC) + 0.175(POP) - 0.833(PS)$
R2=0.93

(ii) Se agrega $\hat{w}$ a la regresión original y resulta:

$\widehat{EXP} = -89.4 + 4.50(AID) - 1.39(\hat{w}) + 0.0013(INC) - 0.518(POP)$
Est."t" (-1.04) (5.89) (-1.73) (3.06) (-4.63) R2=0.99

Con $t = -1.73$, no se rechaza la hipótesis nula, al 5%; entonces "se acepta" la Ho: AID_t y ε_t no tienen correlación (no simultaneidad), por tanto se podría aplicar MCO en cada ecuación del modelo estructural.

6.4. Condición de Endogeneidad

La exogenedidad de las variables regresoras es responsabilidad del investigador. Las variables son endógenas ó exógenas, dependiendo del problema, de la teoría y la información a priori. ¿Será posible una prueba estadística de endogenedidad? Gujarati y Porter (2010) proponen una variante del contraste de Hausman.

Suponga un modelo con 3 variables endógenas: *Y*1, *Y*2 y *Y*3 y dos exógenas *X*1 y *X*2

*Y*1=*f* (*Y*2, *Y*3, *X*1, *X*2, ε)

¿Cómo saber si *Y*1, *Y*2 y *Y*3 son verdaderamente endógenas?. Se puede proceder así:

De la forma reducida, para *Y*2 y *Y*3, se obtienen $\hat{Y}_2$ y $\hat{Y}_3$

Entonces se puede estimar la siguiente regresión auxiliar:

$Y_i = \beta_0 + \beta_2 Y_{2i} + \beta_3 Y_{3i} + \alpha_i X_i + \gamma_2 \hat{Y}_2 + \gamma_3 \hat{Y}_3 + error$, a fin de contrastar:

Ho: γ_2=γ_3= 0 (*Y*2 y *Y*3 son exógenas), frente a
Ha: al menos una $\gamma \neq 0$ (*Y*2 y *Y*3 son endógenas)

Nota: podría ser considera una prueba adicional o complementaria a la de Hausman para simultaneidad.

6.5. Modelo de Ecuaciones Recurrentes

Es un caso especial en el que MCO produce estimaciones consistentes.

Se dice que un sistema de ecuaciones es recurrente si c/u de las variables endógenas puede determinarse en forma secuencial y los errores de cada ecuación son independientes entre sí.

Ejemplo: modelo de oferta y demanda

$Q_t^s = \alpha_1 + \alpha_3 P_{t-1} + \varepsilon_t$ (1)
$P_t = \beta_1 + \beta_2 Q_t + \beta_3 W_t + \mu_t$ (2)

En la ecuación de oferta, la cantidad depende del precio del período anterior, como podría ser el caso de productos agrícolas. En la demanda el precio está determinado una vez que se ha suministrado la cantidad de producto. También hacemos el supuesto importante de que la Covarianza (ε_t , μ_t) = 0; es decir que las ecuaciones no están vinculadas por una correlación entre las variables "omitidas". El modelo parece simultáneo, pero no lo es.

Debido a que el precio está rezagado en la ecuación de oferta, no hay retroalimentación directa de la ecuación de demanda la ecuación de oferta. En cualquier modelo recurrente de este tipo, MCO es el procedimiento correcto, ecuación x ecuación.

6.6. Estimación de un Modelo de Ecuaciones Múltiples

Regresando al modelo original de oferta identificada y demanda no:

$q_t^s = \alpha_2 p_t + \varepsilon_t$ (1)
$q_t^d = \beta_2 P_t + \beta_3 Y_t + \mu_t$ (2) Forma Estructural

En la forma reducida:
$q_t = \pi_{12} Y_t + v_{1t}$ (1) Forma reducida
$p_t = \pi_{22} Y_t + v_{2t}$ (2)

La estimación con Variables instrumentales (VI) abre una "posibilidad", pues en el contexto de ecuaciones simultáneas, debido a que las variables predeterminadas en el modelo pueden servir como buenas variables instrumentales, dado el hecho que estén en el modelo, sugiere correlación con las variables endógenas.

6.6.1. Método de mínimos cuadrados indirectos (MCI)

Supongamos ahora que tenemos el modelo (en desviaciones respecto a la media):

$q_t^s = \alpha_2 p_t + \varepsilon_t$ (1)

$q_t^d = \beta_2 P_t + \beta_3 Y_t + \mu_t$ (2) Forma Estructural

La forma reducida será:

$q_t = \pi_{12} Y_t + v_{1t}$ (1)

$p_t = \pi_{22} Y_t + v_{2t}$ (2) Forma Reducida

En la forma reducida es claro que MCO serían insesgados y consistentes.

Nótese que $\frac{\pi_{12}}{\pi_{22}} = \alpha_2$

Ello sugiere que podemos estimar, de manera consistente, mediante: $\hat{\alpha}_2 = \frac{\hat{\pi}_{12}}{\hat{\pi}_{22}}$

Entonces MCO puede usarse para obtener estimaciones consistentes en ecuaciones simultáneas. Desafortunadamente es un procedimiento no generalizable.

En algunos casos no es posible estimar con Mínimos Cuadrados Indirectos (MCI); por ejemplo, no lo es en el caso previo, para la estimación de la ecuación de demanda (esto como ejercicio). En otros casos ofrecen varias estimaciones de pendientes distintas. Ejemplo:

$q_t = \alpha_2 p_t + \varepsilon_t$ (1)

$q_t = \beta_2 p_t + \beta_3 y_t + \beta_4 w_t + \mu_t$ (2) Modelo Estructural

En forma reducida:

$$q_t = \frac{\beta_3\alpha_2}{\alpha_2-\beta_2} y_t + \frac{\beta_4\alpha_2}{\alpha_2-\beta_2} w_t + \frac{\alpha_2\mu_t-\beta_2\varepsilon_t}{\alpha_2-\beta_2} = \pi_{12} y_t + \pi_{13} w_t + v_{1t} \quad (3)$$

$$p_t = \frac{\beta_3}{\alpha_2-\beta_2} y_t + \frac{\beta_4}{\alpha_2-\beta_2} w_t + \frac{\mu_t-\varepsilon_t}{\alpha_2-\beta_2} = \pi_{22} y_t + \pi_{23} w_t + v_{2t} \quad (4)$$

Las ecuaciones de forma reducida pueden ser estimadas en forma consistente usando MCO. Sin embargo, nos enfrentamos con dos opciones de estimadores de pendiente de la oferta cuando intentamos estimar MCI.

$$\frac{\hat{\pi}_{12}}{\hat{\pi}_{22}} = \hat{\alpha}_2 = \frac{\hat{\pi}_{13}}{\hat{\pi}_{23}}$$

En resumen, el procedimiento incluye los siguientes pasos:

Paso 1. Se obtiene las ecuaciones en forma reducida.

Paso 2. Se aplica MCO individualmente a ecuaciones en forma reducida. Esta operación es permisible puesto que las variables explicativas están predeterminadas y por tanto no correlacionadas con las perturbaciones estocásticas. Las estimaciones así obtenidas son insesgadas y consistentes.

Paso 3. Se obtienen estimaciones de los coeficientes estructurales originales a partir de los coeficientes de forma reducida estimados, obtenidos en el paso 2.

El método de mínimos cuadrados indirectos funciona bien en ecuaciones bien identificadas; pero no en ecuaciones sobre-identificadas.

6.6.2. Método de mínimos cuadrados en dos etapas (MC2)

Considere el modelo:

$$Y_{1t} = \beta_{10} + \beta_{11}Y_{2t} + \alpha_{11}X_{1t} + \alpha_{12}X_{2t} + \mu_{1t} \quad (1) \quad \text{Función de Ingreso}$$
$$Y_{2t} = \beta_{20} + \beta_{21}Y_{1t} + \mu_{2t} \quad (2) \quad \text{Oferta monetaria}$$

Donde:
Y_{1t} = ingreso
Y_{2t} = existencia de dinero
X_{1t} = gastos de inversión
X_{2t} = gastos del Gobierno central.
$n = 35$

Claramente las variables X_{1t} y X_{2t} son exógenas, por lo que la forma reducida será:

$$Y_{1t} = \pi_{10} + \pi_{11}X_{1t} + \pi_{12}X_{2t} + \mu_{1t} \quad (3)$$
$$Y_{2t} = \pi_{20} + \pi_{21}X_{1t} + \pi_{22}X_{2t} + \mu_{2t} \quad (4)$$

Al aplicar la condición de orden: la ecuación de ingreso está sub-indentificada y la de oferta está sobre-identificada (no puede estimarse mediante MCI pues hay dos estimaciones de β_{21}(verificar).

Para estimar los parámetros de la ecuación de oferta, usamos MC2E:

Paso 1: Eliminar la correlación probable entre Y_{1t} y μ_{2t} mediante la regresión de Y_{1t} en función de todas las variables predeterminadas en el sistema completo.

Se realiza la regresión de:

$$Y_{1t} = \hat{\pi}_0 + \hat{\pi}_1 X_{1t} + \hat{\pi}_2 X_{2t} + \hat{\mu}_t \quad (1)$$

Entonces luego usar $\hat{Y}_{1t}$ estimado, suponiendo los siguientes resultados:

$\hat{Y}_{1t} = 2689.8 + 1.87(X_{1t}) + 2.034(X_{2t})$
EE = (67.98) (0.1717) (0.1075)
"t"= (39.56) (10.894) (18.92) R2=0.99

Paso 2: Estimar la función de oferta monetaria (2) reemplazando la estimación de la variable endógena $\hat{Y}_{1t}$:

$\hat{Y}_{2t} = -2440.2 + 0.792(\hat{Y}_{1t})$
EE = (126.95) (0.0178)
"t"= (-19.2) (44.524) R2=0.99

Después de una corrección, se obtiene la siguiente información comparativa:

Mínimos Cuadrados en 2 Etapas	**Mínimos Cuadrados Ordinarios**
$\hat{Y}_{2t} = -2440.2 + 0.792(\hat{Y}_{1t})$ EE = (127.37) (0.0212) "t" = (-17.3) (37.304) R2=0.98	$\hat{Y}_{2t} = -2195.5 + 0.7911(\hat{Y}_{1t})$ EE = (126.64) (0.0211) "t" = (-17.3) (37.38) R2=0.9803

Al comparar resultados, se nota que las regresiones son casi iguales, ello no invalida el método MC2 y se explica por el alto valor de R^2 en la primera fase (correlación entre Y y $\hat{Y}$). Después de una corrección, se obtienen los resultados mostrados en el lado izquierdo del recuadro previo.

Ejercicios del Capítulo VI: Ecuaciones Simultáneas

6.1. Considere el siguiente modelo de Oferta y Demanda de Economistas en Lima[27]:

$Y_t = \beta_0 + \beta_1 X_{1t} + \mu_t$ (1)

$Y_t = \beta_0 + \beta_1 X_{1t} + 2\beta_2 X_{2t} + v_t$ (2)

(i) Especifique las variables endógenas y exógenas en el modelo

(ii) En la ecuación (2): ¿$\hat{\beta}_1$ y $\hat{\beta}_2$ son insesgados?

Sugerencias: *(i)* Y_t y X_{1t} son endógenas. *Revisar el capítulo 18 de Gujarati y Porter (2010).*

6.2. Considere el siguiente modelo, basado en la teoría de Phillips:

$W_t = \alpha_0 + \alpha_1 U_t + \alpha_2 P_t + \mu_{1t}$ (1)
$P_t = \beta_0 + \beta_1 W_t + \beta_2 R_t + \beta_3 M_t + \mu_{2t}$ (2)

Donde:
W_t = tasa de cambio de los salarios nominales
U_t = % de desempleo
P_t = tasa de cambios de los precios al consumidor
R_t = tasa de cambio del costo del capital
M_t = tasa de cambio del precio de materias primas
t = tiempo

(i) ¿Sería correctamente aplicable el método de MCO en el sistema?

(ii) Obtenga la identificación de la forma estructural; también la forma reducida del sistema.

Sugerencias: *(i) No sería correctamente aplicable el método de MCO para estimar los parámetros de las dos ecuaciones, dada pues las variables endógenas estarían correlacionadas con el término de error.*   *(ii) Las ecuaciones de la forma reducida serían:*

[27] Algunos ejemplos de esta sección han sido adaptados del capítulo 18 de Gujarati y Porter (2010); otros del capítulo del libro de Wooldridge (2010).

$W_t = f(U, M)$ *y* $P_t = g\ (U, R, M)$ *La ecuación (1) estaría sobreidentificada y la (2) exactamente identificada.*

6.3. Suponga un modelo de regresión verdadero para ventas de un determinado perfume en el distrito de La Molina:

$$Y_{1i} = \alpha_1 + \beta_1 Y_{2i} + \beta_1 Y_{3i} + \beta_3 Y_{4i} + \mu_{1i} \quad (1)$$
$$Y_{2i} = \alpha_2 + \beta_4 Y_{1i} + \beta_5 Y_{5i} + \gamma_1 X_{1i} + \gamma_2 X_{2i} + \mu_{2i} \quad (2)$$
$$Y_{3i} = \alpha_3 + \beta_6 Y_{2i} + \gamma_3 X_{3i} + \mu_{3i} \quad (3)$$
$$Y_{4i} = \alpha_4 + \beta_7 Y_{2i} + \gamma_4 X_{4i} + \mu_{4i} \quad (4)$$
$$Y_{5i} = \alpha_5 + \beta_8 Y_{2i} + \beta_9 Y_{3i} + \beta_{10} Y_{4i} + \mu_{5i} \quad (5)$$

Donde:
Y_{1i} = % de tiendas con existencias del producto
Y_{2i} = unidades vendidas por mes
Y_{3i} = índice de contacto con productor e importador
Y_{4i} = índice de actividad de ventas al mayoreo en el área
Y_{5i} = índice de penetración de la marca del producto en existencia
X_{1i} = población objetivo para el producto
X_{2i} = ingreso per cápita en la población del área de influencia
X_{3i} = distancia del centro de la ciudad a La Molina
X_{4i} = distancia del centro de La Molina al puerto del Callao.

Resuelva lo siguiente:

(i) Identifique las variables endógenas y exógenas del modelo.

(ii) ¿Podría estimar una o más ecuaciones en el modelo mediante MCO? ¿Porqué (si o no)?

(iii) Determine si las ecuaciones del sistema estructural están identificadas, sobre-identificadas o sub-identificadas.

***Sugerencias*:** *(i) Todas las Variables Y son endógenas y las X son exógenas.* *(ii) Si es posible estimar por MCO a nivel individual, pero como son parte del problema, los estimadores tienen los problemas del caso: sesgados e inconsistentes.*

6.4. Este modelo se refiere a la Economía Alemana de la post-guerra. En el modelo estructural:

$Y_t = \beta_0 + \beta_1 Y_{t-1} + \beta_2 I_t + \mu_{1t}$ (1)
$I_t = \beta_3 + \beta_4 Y_t + \beta_5 Q_t + \mu_{2t}$ (2)
$C_t = \beta_6 + \beta_7 Y_t + \beta_8 C_{t-1} + \beta_9 P_t + \mu_{3t}$ (3)
$Q_t = \beta_{10} + \beta_{11} Q_{t-1} + \beta_{12} R_t + \mu_{4t}$ (4)

Donde:
Y_t = ingreso nacional
I_t = formación neta de capital
C_t= nivel de consumo personal
Q_t= utilidades
P = índice del costo de vida
R = productividad industrial
t = tiempo
μ = perturbaciones estocásticas

(a) ¿Qué variables consideraría endógenas y cuáles exógenas?

(b) Determine si las ecuaciones dadas en el modelo están identificadas.

(c) ¿Hay alguna ecuación en el sistema que pueda estimarse mediante el método de mínimos cuadrados uniecuacional?

(d) ¿Cuál sería la razón para incluir la variable *P* en la función de consumo?

***Sugerencias*:**
(a) Las variables endógenas son: Y, C, Q e I. Las predeterminadas serían P, R, Y_{t-1}, C_{t-1} *y* Q_{t-1}.

(b) Todas las ecuaciones están sobre-identificadas.

(d) P es una variable que podría representar la inflación.

6.5. Para el caso de la economía norteamericana postguerra (1948-1957) se estimó el modelo:
$C_t = \beta_1 + \beta_2 YD_{t-1} + \beta_3 M_t + \mu_{1t}$ (1)
$I_t = \beta_4 + \beta_5 (Y_{t-1} - Y_{t-2}) + \beta_6 Z_{t-1} + \mu_{2t}$ (2)
$G_t = \beta_7 + \beta_8 G_{t-1} + \mu_{3t}$ (3)
$Y_t = C_t + I_t + G_t$ (4)

Donde:
Y = producto nacional bruto

C = gasto de consumo personal
I = inversión privada doméstica bruta
G = gasto del gobierno más inversión extranjera neta
YD = ingreso disponible (menos impuestos)
M = oferta monetaria al principio del trimestre
Z = ingreso patrimonial antes de impuestos
t = tiempo
μ = perturbaciones estocásticas

Todas las variables están medidas en forma de primeras diferencias.

Mediante uso del método de MCO a nivel de cada ecuación, se obtuvo los siguientes resultados:

$\hat{C}_t = 0.09 + 0.43\, YD_{t-1} + 0.23\, M_t$ R2=0.23
$\hat{I}_t = 0.08 + 0.43(Y_{t-1} - Y_{t-2}) + 0.48$ R2=040
$G_t = 0.13 + 0.67 G_{t-1}$ R2=0.42

(a) ¿Cómo podría justificarse el uso del método de MCO uni-ecuacional en este caso?

(b) ¿Por qué los valores R^2 de las ecuaciones individuales serán relativamente bajos?

Sugerencias:
(a) Éste no es un sistema de ecuaciones simultáneas. Los estimadores serían insesgados y consistentes.

(b) Cuando las variables están medidas en forma de primeras diferencias, normalmente es removida la tendencia; siendo así suele existir R^2 relativamente bajos.

6.6. En el siguiente modelo siguiente:

$Q_t^s = \alpha_0 + \alpha_1 P_t + \alpha_2 P_{t-1} + \varepsilon_t$ (1)
$Q_t^d = \beta_0 + \beta_1 P_t + \beta_2 Y_t + \mu_t$ (2)
$Q_t^s = Q_t^d$ (3) condición de equilibrio

(i) ¿Se podría decir que la ecuación de oferta está sobre-identificada? ¿Porqué? (si o no).

(ii) Identifique las variables endógenas y predeterminadas.

(iii) ¿Qué restricciones podrían indicarse en los parámetros estructurales de modo tal que la ecuación de oferta quede identificada?

> ***Sugerencias***:
> *(i) Variables endógenas: Q_t^s y Q_t^d y P_t (variables determinadas dentro del sistema). Variables predeterminadas: 02 (ayudan a causar el movimiento de las V. endógenas); una endógena rezagada: P_{t-1} y una exógenas: Y_t (determinadas plenamente fuera del sistema).*
>
> *(ii) La ecuación de oferta está sobre-identificada.*
>
> *(iii) Si α_1=0, o α_2=0 la ecuación de oferta podría estar identificada.*

6.7. Suponga el sistema estructural de ecuaciones (1) y (2).

$$Y_{1t} = \beta_{10} + \beta_{12}Y_{2t} + \gamma_{11}X_{1t} + \mu_{1t} \qquad (1)$$
$$Y_{2t} = \beta_{20} + \beta_{21}Y_{1t} + \gamma_{22}X_{2t} + \mu_{2t} \qquad (2)$$

De los que se obtiene el sistema en su forma reducida:

$$Y_{1t} = \pi_{10}+\pi_{11}X_{1t} + \pi_{12}X_{2t} + w_t \qquad (3)$$
$$Y_{2t} = \pi_{10}+\pi_{11}X_{1t} + \pi_{12}X_{2t} + w_t \qquad (4)$$

(i) ¿Están identificadas las ecuaciones de la forma estructural del sistema?

(ii) ¿Qué sucede con la identificación si se sabe a priori que γ_{11}= 0?.

Suponga ahora que la estimación de las ecuaciones de forma reducida, arroja los siguientes resultados:

$$Y_{1t} = 4 + 3\,X_{1t} + 8\,X_{2t}$$
$$Y_{2t} = 2 + 6\,X_{1t} + 10\,X_{2t}$$

(iii) Obtenga los valores de los parámetros estructurales

(iv) ¿Cómo se probaría la hipótesis nula que γ_{11}= 0?.

> ***Sugerencias***:
> *(i) Las ecuaciones (1) y (2) están exactamente identificadas.*
>
> *(ii)* Si γ_{11}= 0, *la ecuación (1) estaría identificada; la (2) no lo estaría.*
>
> *(iii)* $\hat{\beta}_{10} = -3$; $\hat{\beta}_{12} = 1.25$; $\hat{\beta}_{20} = -6$; $\hat{\beta}_{21} = 2$; $\hat{\gamma}_{11} = 2.25$; $\hat{\gamma}_{22} = -6$
>
> *(iv) Se necesitaría el error estándar de $\hat{\gamma}_{11}$*

6.8. Suponga el siguiente sistema de ecuaciones estructurales:

$$Y_{1t} = \beta_{10} + \beta_{12}Y_{2t} + \gamma_{11}X_{1t} + \mu_{1t} \qquad (1)$$

$Y_{2t} = \beta_{20} + \beta_{21} Y_{1t} + \mu_{2t}$ (2)

De las que ha sido posible plantear y estimar las siguientes ecuaciones de forma reducida:

$Y_{1t} = 4 + 8\, X_{1t}$
$Y_{2t} = 2 + 12\, X_{1t}$

(i) Indique y explique ¿qué coeficientes de la forma estructural podrían ser estimados a partir de los coeficientes de las ecuaciones en forma reducida?

(ii) Suponga se sabría que: (a) β_{10}=0 y (b) β_{12}=0, ¿cómo cambiaría la respuesta en (i)?

> ***Sugerencias***:
> *(i) La ecuación (1) no está identificada. Los coeficientes de la forma estructural no pueden ser identificados.*
>
> *(ii) En este caso ambas ecuaciones estarían identificadas.*

6.9. Considere un modelo de mercado de dinero:

$M_t^d = \beta_0 + \beta_1 Y_t + \beta_2 R_t + \beta_3 P_t + \mu_{1t}$ (1) Demanda de dinero
$M_t^S = \alpha_0 + \alpha_1 Y_t + \mu_{2t}$ (2) Oferta de dinero

Donde las variables son: Dinero (M), Ingreso (Y), Tasa de interés (R), Precio (P).

(i) ¿Están identificada las funciones de demanda y de oferta?

(ii) ¿Qué método utilizaría para estimar los parámetros de las ecuaciones identificada(s)? ¿Por qué?

(iii) Suponga que se modifica la función de oferta agregando las variables explicativas Y_{t-1} y M_{t-1}. ¿Qué sucedería con el tema de la identificación? ¿¡Se utilizaría el método ´previo (ii)? Explique.

(iv) Suponga que se modifica la función de oferta agregando las variables explicativas Y_{t-1} y R_{t-1} ¿Qué sucede con el tema de la identificación? ¿Utilizaría los mismos métodos de estimación que en el caso anterior (ii)? ¿Por qué?

> ***Sugerencias***:
> *(i) La ecuación de demanda no estaría identificada.*
>
> *(ii) La ecuación de oferta estaría sobre-identificada.*

(iii) Se podría usar MC2E para la ecuación sobre-identificada.
(iv) Ambas ecuaciones estarían sobre-identificadas. Podría usarse MC2E.

6.10. Considere el modelo Keynesiano:

$$C_t = \beta_{10} + \beta_{11}Y_t + \mu_{1t} \quad (1)$$
$$I_t = \beta_{20} + \beta_{21}Y_t + \beta_{22}Y_{t-1} + \mu_{2t} \quad (2)$$
$$Y_t = C_t + I_t + G_t \quad (3)$$

Donde:
C: gastos de consumo
I: gastos de inversión
Y: ingreso
G: gastos del gobierno

G y Y_{t-1} se suponen predeterminados

(i) Obtenga las ecuaciones en forma reducida y determine cuáles de las ecuaciones anteriores están identificadas.

(ii) ¿Cuál método puede utilizarse para estimar las ecuaciones sobre-identificada y de la ecuación sobre-identificada? Justifique la respuesta.

Sugerencias:
(i) La ecuación de C_t estaría sobre-identificada y la de I_t identificada.
(ii) Se puede usar MC2E para la función de consumo y MCI para la ecuación de inversión.

6.11. Considere el sistema de ecuaciones de un mercado de joyas[28]:

$$Q_t = \alpha_1 P_t + \beta_1 T_t + \mu_{1t} \quad (1)$$
$$Q_t = \alpha_2 P_t + \beta_2 Z_t + \mu_{2t} \quad (2)$$

(i) Si $\alpha_1 \neq 0$ y $\alpha_2 = 0$, determine la forma reducida de P_t.

(ii) ¿Es probable que la condición $\alpha_1 \neq \alpha_2$ satisfaga un modelo de oferta y demanda? Explique.

[28] Este problema ha sido adaptado del capítulo 16 del libro de Wooldridge (2010)

Sugerencias

(i) Después de varias transformaciones,

$P_t = \pi_{21} + \pi_{22} Z_1 + \pi_{22} Z_2 + v_t$

Donde $\pi_{21} = \frac{\beta_1}{\alpha_1}$; $\pi_{22} = -\frac{\beta_2}{\alpha_1}$ *y que* $v_t = \frac{\mu_{1t} - \mu_{2t}}{\alpha_1}$

(ii) Si la ecuación (1) generalmente esperamos $\alpha_1 > 0$*, y la ecuación (2) es la de demanda entonces* $\alpha_2 < 0$*- Pero en estas condiciones, las formas reducidas de las ecuaciones pueden existir aun en el caso en que la función de oferta no fuera de pendiente positiva y la demanda de pendiente negativa; pero podríamos preguntarnos sobre la utilidad real de tales condiciones de oferta y demanda.*

6.12. En el siguiente modelo de mercado local en que se impone la condición de equilibrio de igualdad entre oferta y demanda de maíz[1]:

$$Q_t = \alpha_1 P_t + \beta_1 Y_t + \mu_{1t} \quad (1)$$
$$Q_t = \alpha_2 P_t + \beta_2 R_t + \gamma_2 R_t^2 + \mu_{2t} \quad (2)$$

Donde:
Q_t: consumo per cápita de maíz (Kg.,)
P_t: precio del maíz (S/. por Kg.)
Y_t: ingreso per cápita gastos del gobierno
R_t: lluvia, en pulgadas.

(i) Explique cómo podría identificar las ecuaciones de oferta y demanda.

(ii) ¿Cómo sería la identificación de cada ecuación?

Sugerencias

(i) Usando conceptos económicos básicos, la 1ª ecuación debe ser la de demanda, pues depende del ingreso, mientras que la segunda sería la de oferta pues depende del clima.

(ii) Revisar "identificación" en el libro de Pyndick y Rubenfeld (2001)

6.13. Desde el punto de vista de la empresa, suponga el siguiente modelo para dos ecuaciones: una de Ganancias y otra de Consumo de cerveza en el área metropolitana de la ciudad.

$$LnG_t = \beta_0 + \beta_1 C_t + \beta_2 E_t + \mu_{1t} \quad (1)$$
$$C_t = \alpha_0 + \alpha_1 LnG_t + \alpha_2 E_t + \alpha_3 LnP_t + \mu_{2t} \quad (2)$$

Donde:
G_t= ganancias anuales mensuales (soles)

Y_t= consumo mensual de cerveza (botellas de litro)
E_t= nivel de educación de la población del área
P_t= índice de precios de la cerveza, incluyendo impuestos.

Suponga también que E_t y P_t son exógenas. Si β_1, β_2, α_1, α_2, α_3 son distintos de cero, entonces: ¿qué ecuación estaría identificada? y ¿cómo se estimaría la ecuación?

Sugerencias
Parece que no existen variables exógenas en la ecuación (1). La ecuación (1) estaría identificada siempre que $\alpha_3 \neq 0$ (y presumiríamos que $\alpha_3 < 0$). Esto permite una variable exógena LnP_t , que puede ser usada como VI para alcohol en la estimación de la ecuación (1) mediante MC2E (que es justo una VI estándar en este caso).

6.14. Considere el modelo siguiente que permite determinar la efectividad de las vacunas COVID para reducir los contagios entre estudiantes de universidades del país.

$$G_t = \beta_0 + \beta_1 C_t + \beta_2 Y_t + \beta_4 D + \mu_1 \qquad (1)$$

Donde:
G_t= estudiantes que se contagiaron (%)
C_t= estudiantes vacunados con al menos 2 dosis (%)
Y_t= nivel de ingreso familiar promedio (S/.)
D variable ficticia que indica si la Universidad está en la capital (1), o no (0).

(i) ¿Porqué G_t y C_t podrían determinarse conjuntamente? ¿Qué signo esperaría de $\hat{\beta}_1$?

Considerando la siguiente ecuación adicional: $C_t = \alpha_0 + \alpha_1 G_t + otros\ factores$ (2)

(ii) Suponga que la vacuna resultara incrementando la tasa de contagios, de manera que $\alpha_1 > 0$ en la ecuación (2): ¿cuál sería el probable sesgo si se estimara β_1 mediante MCO?

Sugerencias:
(i) Se esperaría que $\hat{\beta}_1 < 0$, pues manteniendo constante otras cosas, en promedio una alta tasa de universitarios vacunados debería reducir el % de contagios.

(ii) Si tenemos a la variable G_t en la ecuación (2), entonces en dicha ecuación $\alpha_1 > 0$, pues C_t depende de $\alpha_1 G_t$. Debido a que $\alpha_1 > 0$, entonces Ct estaría positivamente relacionada con μ_1 De hecho, si el error estructural μ_2 en la ecuación (2) no está correlacionada con μ_1, la cov (Ct, μ_1)= $\alpha_1 V(\mu_1) > 0$.

Finalmente, desde que pensamos que $\beta_1 < 0$, entonces MCO es viciado. O sea si usamos MCO en la ecuación (1), es probable que sub-estimamos la importancia de *las vacunas en la reducción de la enfermedad (recuerde que mientras más negativo es β_1, más efectivas son las vacunas).*

Referencias Bibliográficas

Alarcón, J. y J. Nolazco (2014). *Econometría con E-views y Aplicaciones en Economía de Recursos Naturales y Desarrollo Sustentable.* 1ª edición. Fondo Editorial UNALM. Perú.

Greene, W. H. (2012). *Econometric Analysis*. 7th edition. Prentice Hall, New York.

Gujarati, D y W. Porter (2010). *Econometría*. 5ª edición. Editorial Mc Graw Hill.

Hernández, J. y J. Zúñiga (2013). *Modelos econométricos para el análisis económico.* 1ª edición. Editorial ESIC. España.

Kennedy, P. (2011). *A Guide to Econometrics*. 6th edition. Editorial Blackwell Publishing.

Palacios, F., R. García y J. Herrerías (2017). *Econometría: ejercicios resueltos.* Editorial Pirámide- Grupo Anaya, S.A. España.

Pérez, C. (2011). *Econometría Avanzada: técnicas y herramientas*. 2ª edición, Editorial Grupo Garceta, España.

Pichihua, J. (2002). *Econometría: Teoría y Aplicaciones.* Universidad Nacional Agraria La Molina, Perú.

Pindyck, R. & D. Rubinfeld (2001). *Econometría: modelos y pronósticos*. 3a edición, Mc Graw Hill /Interamericana Editores.

Wooldridge, J. (2010) *Introducción a la econometría: un enfoque moderno*. 4ª edición. Editorial CENGAGE Learning.

HOMINEM ET AGRUM COLERE CUPIO

Este libro está pensado para ayudar a los estudiantes en sus cursos y en apoyo a sus investigaciones; también puede ayudar a profesionales, de diversos sectores, en cuanto a la aplicación de técnicas de modelística en sus investigaciones. El objetivo es rescatar aspectos prácticos en Econometría Aplicada, en una presentación sencilla y didáctica.

En cada capítulo el libro contiene ejemplos de uso de los temas tratados, también ejercicios y sugerencias de solución. Los ejercicios están hechos con base en información tomada de la propia experiencia de investigación del autor, así como ejercicios adaptados de otros libros, a menudo con adaptación a circunstancias de nuestra propia universidad, ciudad y país.

ISBN: 978-612-5086-09-9
9 786125 086099

Fondo Editorial
Universidad Nacional Agraria La Molina