INTRODUCCIÓN A LA ECONOMETRÍA

3er curso LE y LADE

Tema 2c

Dpto. de Econometría y Estadística (EA3)

UPV—EHU

2.7a Omisión de variables relevantes.

Omisión de variables relevantes

- relación verdadera:

$$Y = X\beta + u = \left[X_I \;\middle|\; X_{II} \right] \begin{pmatrix} \beta_I \\ \hline \beta_{II} \end{pmatrix} + u$$

$$X = \left[\begin{array}{cccc|cccc} 1 & x_{11} & \cdots & x_{K_1,1} & x_{K_1+1,1} & \cdots & x_{K1} \\ 1 & x_{12} & \cdots & x_{K_1,2} & x_{K_1+1,2} & \cdots & x_{K2} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 1 & x_{1T} & \cdots & x_{K_1,T} & x_{K_1+1,T} & \cdots & x_{KT} \end{array} \right], \quad \beta = \begin{pmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_{K_1} \\ \hline \beta_{K_1+1} \\ \vdots \\ \beta_K \end{pmatrix}.$$

$$Y = X_I\beta_I + X_{II}\beta_{II} + u$$

- relación estimada :

$$Y = X_I\beta_I + v \qquad \text{donde} \quad v = X_{II}\beta_{II} + u,$$

entonces $\mathsf{E}(v) \neq 0 \quad \rightsquigarrow \quad \mathsf{E}(\widehat{\beta}) \neq \beta.$

es decir $\widehat{\beta}$ es sesgado.

Omisión de variables relevantes: consecuencias

Resumen:

- Estimador MCO de los coeficientes es *sesgado*

  (excepto si $x_I' x_{II} = 0$).

- Estimador MCO del intercepto es *siempre sesgado*.
- Estimador de varianza del error es *siempre sesgado*.

2.7b Multicolinealidad

Multicolinealidad Perfecta

Caso extremo:

- combinación lineal exacta:
  - $\sum_{k=0}^{K} \lambda_k X_{kt} = 0, \quad \lambda \neq 0, \quad X_{0t} = 1,$
  - $\exists X_i \mid X_i = \lambda_0^* + \sum_{\substack{k=1 \\ k \neq i}}^{K} \lambda_k^* X_{kt},$
  - $\exists X_i, X_j \mid$ Corr$(X_i, X_j) = 1,$
  - $\exists X_i \mid$ regres aux X_i sobre $\{X_k\}_{\substack{k=1 \\ k \neq i}}^{K} \quad \rightsquigarrow \quad \mathrm{R}_i^2 = 1.$
- Problema:
  - $\mathrm{rk}\, X < K{+}1$, (X no es de rango completo)
  - $\rightsquigarrow \quad \det(X) = 0$
  - $\rightsquigarrow \quad \nexists (X'X)^{-1}$
  - $\rightsquigarrow$ ¿ $\widehat{\beta}$?

Multicolinealidad perfecta: ejemplo

- Sea $X_{4t} = 2X_{1t} \quad \forall t$:

$$X_{4t} = 0 + 2X_{1t} + 0 \cdot X_{2t} + 0 \cdot X_{3t} + 0 \cdot X_{5t} + \cdots + 0 \cdot X_{Kt},$$

- ¿no error? $\Rightarrow$ regres aux X_4 sobre $\{X_k\}_{\substack{k=1 \\ k \neq 4}}^{K} \quad \rightsquigarrow$!‘$\mathrm{R}_4^2 = 1$!!
- Especificación del modelo:

$$Y_t = \beta_0 + \beta_1 X_{1t} + \beta_2 X_{2t} + \beta_3 X_{3t} + \beta_4 X_{4t} + \cdots + u_t, t = 1, 2 \ldots, T,$$
$$X_{4t} = 2X_{1t},$$

- y sustituyendo en el modelo:

$$Y_t = \beta_0 + \beta_1 X_{1t} + \beta_2 X_{2t} + \beta_3 X_{3t} + \beta_4 (2X_{1t}) + \cdots + u_t,$$
$$= \beta_0 + \underbrace{(\beta_1 + 2\beta_4)}_{\beta_1^\star} X_{1t} + \beta_2 X_{2t} + \beta_3 X_{3t} + \cdots + u_t$$

- ahora tenemos un parámetro menos para estimar.

Multicolinealidad: contra-ejemplo

$$Y_t = \beta_0 + \beta_1^\star X_{1t} + \beta_2 X_{2t} + \beta_3 X_{3t} + \cdots + u_t$$

- Sólo quedan K parámetros a estimar,
  pero β_1 y β_4 no se pueden estimar por separado:
  - únicamente podemos estimar una combinación lineal de ellos:
    $\beta_1^\star = \beta_1 + 2\beta_4,$
  - *es decir* ¡efecto combinado de X_{1t} y X_{4t} sobre Y_t!!
- (Ejercicio: Pruébese con $X_{2t} - 3X_{3t} = 10, \quad \forall t$.)
- multicolinealidad $=$ *relaciones lineales*
  pero. . . ¿qué ocurre si la relación *no es lineal*? *p.ej.:*

$$Y_t = \beta_0 + \beta_1 X_{1t} + \beta_2 X_{1t}^2 + u_t$$

  - X tiene rango completo por columnas $\rightsquigarrow$ no hay problema.

Multicolinealidad perfecta : consecuencias

- algunos parámetros no pueden estimarse por separado.
- algunas estimaciones son sólo c.l. de los parámetros.
- R^2 es correcto:
  recoge correctamente la proporción de (la varianza de) Y_t explicada por la regre
- Las predicciones de Y son todavía validas.

2.7c Multicolinealidad imperfecta

Multicolinealidad imperfecta

- Problema:

$$Y_t = \beta_0 + \beta_1 X_{1t} + \beta_2 X_{2t} + \beta_3 X_{3t} + \beta_4 X_{4t} + \cdots + u_t, t = 1, 2 \ldots, T,$$
$$X_{4t} = 2X_{1t} + v_t,$$

$v_t =$ discrepancia entre X_{4t} y $2\,X_{1t}$,

- relación aproximada :
- regresión auxiliar X_{4t} sobre el resto $\rightsquigarrow$ $\mathrm{R}^2 \approx 1$.
- es una cuestión de grado ($x'x$ no diagonal $\rightsquigarrow$ variables correlacionadas)
- Nota: si no se especifica perfecta/imperfecta quiere decir mc imperfecta.

Multicolinealidad: Síntomas

- Síntomas típicos :
  - alto R^2 (grupo relevante de regresores)
  - pero variables parecen no relevantes individualmente (incapacidad de separar efectos de regresores).
- más formalmente:

$$\text{Var}(\widehat{\beta}^\star) = \sigma^2 (x'x)^{-1} = \frac{\sigma^2}{T} \text{Var}(X^\star)^{-1}$$

$$\Rightarrow \quad \text{Var}(\widehat{\beta}_k) = \frac{\sigma^2}{T\text{Var}(X_k)(1 - \mathrm{R}_k^2)},$$

- de forma que en el ejemplo anterior $X_{4t} \approx 2X_{1t}$:
  - Corr$(X_4, X_1) \uparrow$
  - R_4^2 y R_1^2 $\upuparrows$
  - denominador $\downarrow$
  - varianzas $\upuparrows$

Multicolinealidad: Consecuencias

- Algunos coeficientes no son significativos, a pesar de que sus variables tienen un efecto importante sobre la variable dependiente.
- De todas maneras, Gauss-Markov $\Rightarrow$ estimadores lineales, insesgados y de varianza mínima, entonces *no es posible encontrar un ELI (más)Optimo.*
- R^2 es correcto: recoge correctamente la proporción de (la varianza de) Y_t explicado por la regresión.
- Predicciones de Y son todavía validas.

Multicolinealidad: Como detectar

- Pequeños cambios en los datos
  ⇒ importantes cambios en las estimaciones
  (pueden afectar hasta sus signos).
- Estimaciones de los coeficientes
  no son significativas de forma individual...
- ...pero sí lo son de forma conjunta.
- Alto coeficiente de determinación R^2.
- Regresiones auxiliares entre regresores
  ⇒ alto R^2_k.

Multicolinealidad: Algunas soluciones

La multicolinealidad no es un problema fácil de solucionar.
De todas formas, de

$$\mathrm{Var}(\widehat{\beta}_k) = \frac{\sigma^2}{T\mathrm{Var}(X_k)(1-\mathrm{R}^2_k)},$$

tenemos que para reducir la varianza podríamos:

$\mathbf{T}\uparrow$: Incrementar el número de observaciones T .
También, las diferencias entre regresores pueden incrementar.

$\mathrm{Var}(\mathbf{X})\uparrow$: Incrementar dispersión de los datos; *p.ej.* estudio sobre la función del consumo:
muestra de familias ⇝ todas las rentas posibles.

$\mathrm{Var}(\mathbf{X})\uparrow$: Incluir información adicional.
p.ej. imponer restricciones sugeridas por T^a. Ec.

$\sigma^2\downarrow$: Añadir un nuevo regresor relevante todavía no incluido.
También evitaría serios problemas de sesgo.

$\mathbf{R^2_k}\downarrow$: Eliminar variables que pueden ser causa de la multicolinealidad.
(Aunque tener cuidado de no omitir algún regresor relevante).

2.8 Estimador MCO bajo Restricciones.

MRLG bajo restricciones lineales (1)

- objetivos de capítulos previos:
  - Modelo Econométrico (MRLG), características y supuestos básicos...
  - pero... no hay conocimiento sobre los parámetros del modelo.
  - El Método Mínimo-Cuadrático de estimación de parámetros (MCO).
  - Propiedades de los estimadores resultantes .
- objetivos del presente capítulo:
  - Información *a priori* sobre los valores del parámetro (o c.l.) ...
  - dada por
    - teoría económica,
    - otros trabajos empíricos,
    - experiencia propia, etc.
  - Modelo No Restringido ⇒ MC Ordinarios.
  - Modelo Restringido ⇒ MC Restringidos.
  - Comprobar, dado el modelo estimado, si la información es compatible con los datos disponibles.

MRLG bajo restricciones lineales: ejemplos

- función de producción con rendimientos de escala constantes: $\beta_K + \beta_L = 1$.
- demanda del producto como función del precio: $\beta = -1$ (por ejemplo).
- en el MRLG: supongamos que $\beta_2 = 0$ y $2\beta_3 = \beta_4 - 1$:
  - Modelo completo:

$$Y_t = \beta_0 + \beta_1 X_{1t} + \cdots + \beta_{Kt} X_{Kt} + u_t, \text{ con } \beta_2 = 0 \text{ y } 2\beta_3 + 1 = \beta_4;$$

  - Modelo transformado alternativo:

$$Y_t = \beta_0 + \beta_1 X_{1t} + 0 X_{2t} + \beta_3 X_{3t} + (2\beta_3 + 1) X_{4t} + \cdots + \beta_{Kt} X_{Kt} + u_t$$
$$Y_t - X_{4t} = \beta_0 + \beta_1 X_{1t} + \beta_3 (X_{3t} + 2X_{4t}) + \cdots + \beta_K X_{Kt} + u_t$$
$$Y_t^* = \beta_0 + \beta_1 X_{1t} + \beta_3 Z_t + \cdots + \beta_K X_{Kt} + u_t$$

  donde $Y_t^* = Y_t - X_{4t}$ y $Z_t = X_{3t} + 2X_{4t}$.
  - Este modelo transformado:
    - se puede estimar por MCO:
      $\widehat{\beta}_0, \widehat{\beta}_1, \widehat{\beta}_3, \widehat{\beta}_5, \ldots, \widehat{\beta}_K$, junto con $\widehat{\beta}_2 = 0$ y $\widehat{\beta}_4 = 2\widehat{\beta}_3 + 1$.
    - tiene nueva variable endógena Y_t^* (no siempre es así: p.ej. si $\beta_2 = 0$ sólo) y nueva variable explicativa Z_t.

MRLG bajo restricciones lineales (2)

- El método de “transformación” es bueno solamente para casos sencillos.
- En general, q restricciones lineales (no redundantes) entre parámetros:

$$\begin{array}{c} 1 \\ \vdots \\ q \end{array} \begin{pmatrix} \diamond & \diamond & \diamond & \ldots & \diamond \\ \vdots & & & & \\ \diamond & \diamond & \diamond & \ldots & \diamond \end{pmatrix} \begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \ldots \\ \beta_K \end{pmatrix} = \begin{pmatrix} \diamond \\ \vdots \\ \diamond \end{pmatrix}$$

  - para una matriz R y vector r dados,

$$\underset{(q \times K+1)}{R} \quad \beta = \underset{(q \times 1)}{r}$$

  - <1>ejemplo de caso no valido (¿por qué?):

$$\beta_3 = 0, \quad 2\beta_2 + 3\beta_4 = 1, \quad \beta_1 - 2\beta_4 = 3, \quad 6\beta_4 = 2 - 4\beta_2 + \beta_3$$

MRLG bajo restricciones lineales (2cont)

- Escribir el ejemplo anterior $\beta_2 = 0$ y $2\beta_3 = \beta_4 - 1$ ($q = 2$ restricciones) como en la formula general:

$$\underset{(2 \times K+1)}{\begin{pmatrix} 0 & 0 & 1 & 0 & 0 & 0 & \ldots & 0 \\ 0 & 0 & 0 & 2 & -1 & 0 & \ldots & 0 \end{pmatrix}} \underset{(K+1 \times 1)}{\begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \ldots \\ \beta_K \end{pmatrix}} = \underset{(2 \times 1)}{\begin{pmatrix} 0 \\ -1 \end{pmatrix}} .$$

$$R \qquad \beta \qquad r$$

- En general, escribimos el MRLG sujeto a q restricciones lineales como:

$$\underset{(T \times 1)}{Y} = \underset{(T \times K+1)}{X} \underset{(K+1 \times 1)}{\beta} + \underset{(T \times 1)}{u} ,$$
$$\underset{(q \times K+1)}{R} \underset{(K+1 \times 1)}{\beta} = \underset{(q \times 1)}{r} .$$

Estimación:MC Restringidos.

- Típico ejercicio de optimización:

$$\min_{\beta} (u'u) \quad \text{donde} \quad u = Y - X\beta,$$
$$\text{sujeto a } R\beta = r.$$

- Lagrangiano:

$$L(\beta, \lambda) = u'u - 2\lambda'(R\beta - r)$$
$$\min_{\beta, \lambda} L(\beta, \lambda).$$

- Primeras derivadas:

$$\frac{\partial L(\beta, \lambda)}{\partial \beta} = -2\, X'u - 2\, R'\lambda,$$
$$\frac{\partial L(\beta, \lambda)}{\partial \lambda} = -2\,(R\,\beta - r),$$

Estimación: MC Restringidos (cont).

- c.1^{er}o. $\rightsquigarrow$ ecuaciones normales:

$$X'\,\widehat{u}_R + R'\,\widehat{\lambda} = 0, \qquad (4)$$

$$R\,\widehat{\beta}_R = r, \qquad (5)$$

donde $\widehat{\beta}_R$ y $\widehat{\lambda}$ son valores de β, λ que satisfacen las c.1^{er}o. y los residuos

$$\widehat{u}_R = Y - X\,\widehat{\beta}_R. \qquad (6)$$

- Despejando $\widehat{\beta}_R$ y $\widehat{\lambda}$:

$$\widehat{\lambda} = [R\,(X'X)^{-1}R']^{-1}(r - R\widehat{\beta}),$$

$$\widehat{\beta}_R = \widehat{\beta} + (X'X)^{-1}R'[R\,(X'X)^{-1}R']^{-1}(r - R\widehat{\beta})$$

$$= \widehat{\beta} + A(r - R\widehat{\beta}) = (I - AR)\,\widehat{\beta} + Ar \qquad (7)$$

donde $A = (X'X)^{-1}R'[R(X'X)^{-1}R']^{-1}$.

Estimación MCR: características

- Expresión (7): $\widehat{\beta}_R = \widehat{\beta} + A(r - R\,\widehat{\beta}) \rightsquigarrow$
  - la estimación restringida $\widehat{\beta}_R$ puede obtenerse como una función de la estimación ordinaria (no restringida): $\widehat{\beta}$
  - $R\widehat{\beta} \simeq r \quad \Rightarrow \quad \widehat{\beta}_R$ (restringido) $\simeq \widehat{\beta}$ (no restringido) .
- <2>Ecuaciones normales (4): $X'\,\widehat{u}_R + R'\,\widehat{\lambda} = 0 \rightsquigarrow$
  - se satisfacen las restricciones (obvio).
  - $X'\widehat{u}_R \neq 0$, *es decir*:
    - suma de residuos restringidos no es cero,
    - residuos restringidos no ortogonales a las variables explicativas,
    - entonces, residuos restringidos no ortogonales a $\widehat{Y}_R$ ajustado.
  - SCT $\neq$ SCR$_R$ + SCE$_R$
    (compárese con caso ordinario y con ecuación transformada: ¿ R^2 ??).

Propiedades del estimador MCR (1)

Expresión (7) : $\widehat{\beta}_R = (I - AR)\,\widehat{\beta} + Ar \quad \rightsquigarrow$

1. **Lineal:** estimador MCR $\widehat{\beta}_R$ es c.l. del estimador MCO $\widehat{\beta}$, que es lineal, luego, a su vez, $\widehat{\beta}_R$ es también lineal
2. **Sesgo:** estimador MCR $\widehat{\beta}_R$ es $\begin{cases} \text{sesgado,} & \text{si } R\,\beta \neq r \text{ ,} \\ \text{insesgado,} & \text{si } R\,\beta = r \text{ cierto} \end{cases}$

   Demo:

   $$\text{E}(\widehat{\beta}_R) = (I - AR)\,\text{E}(\widehat{\beta}) + Ar = (I - AR)\,\beta + Ar = \beta + A(r - R\beta).$$
3. **Matriz de Covarianzas:** $\text{Var}(\widehat{\beta}_R) = (I - AR)\text{Var}(\widehat{\beta}) = \sigma^2(I - AR)(X'X)^{-1}$
   Demo: (ver apuntes o libro de texto)

Propiedades del estimador MCR (2)

4. **Varianza menor** que los estimadores MCO,
   aunque las restricciones no sean ciertas:

   Demo:

   $$\text{Var}(\widehat{\beta}_R) = \text{Var}(\widehat{\beta}) - AR\,\text{Var}(\widehat{\beta})$$

   $$= \text{Var}(\widehat{\beta}) - (\text{matriz psd}).$$

5. resultado sorprendente (aparentemente):
   - menos “incertidumbre” sobre parámetros
     $\rightsquigarrow$ más precisión en estimación...
   - pero... hacía un resultado erróneo (sesgado)
     si la restricción no es cierta.

Multicolinealidad frente a restricciones

Es importante distinguir claramente entre dos casos diferentes:

- relaciones lineales entre regresores (*es decir* multicolinealidad):

  p.ej. $X_{4t} = 2X_{1t}$

  $\Rightarrow$ falta información para estimaciones individuales.
- relaciones lineales entre coeficientes:

  p.ej. $\beta_4 = 2\,\beta_1$

  $\Rightarrow$ información extra sobre parámetros

  $\rightsquigarrow$ estimadores con varianza más pequeña.
- modelos a estimar respectivamente:

$$Y_t = \beta_0 + (\underbrace{\beta_1 + 2\beta_4}_{\beta_1^*})X_{1t} + \beta_2 X_{2t} + + \cdots + u_t,$$

$\Rightarrow \widehat{\beta}_1^*$ pero ¿ $\widehat{\beta}_1, \widehat{\beta}_4$?

$$Y_t = \beta_0 + \beta_1(\underbrace{X_{1t} + 2X_{4t}}_{X_{1t}^*}) + \beta_2 X_{2t} + + \cdots + u_t,$$

$\Rightarrow \widehat{\beta}_1$ y $\widehat{\beta}_4 = 2\,\widehat{\beta}_1$