Econometría

Maestría en Economía - Maestría en Econometría

Lecture 2

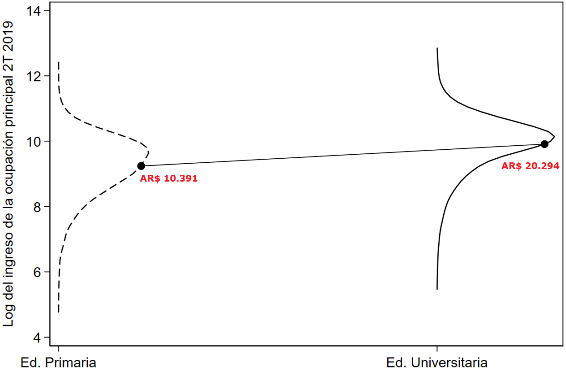
Agenda

1. Variables Omitidas: Motivación
   - La Ecuación del Salario con Habilidad no Observada
   - Consecuencias de Ignorar la Presencia de Variables Omitidas
2. Soluciones al Problema de las Variables Omitidas
   - Variables "Proxy"
   - Variables Instrumentales
   - Mínimos Cuadrados en dos Etapas
3. El enfoque de la función de control para la endogeneidad
4. Errores no Esféricos
   - Efectos de Reducir el Tamaño de las Clases
   - Propiedades de MCC en Presencia de Errores no Esféricos
   - Inferencia en Presencia de Errores no Esféricos
   - Mínimos Cuadrados Generalizados

Agenda

Sesgo por variables omitidas

- ¿Cuál es el retorno monetario a la educación?
- En promedio, personas con mayor educación ganan más que personas menos educadas.
- Imagine, para empezar, que la educación es una decisión binaria: "obtengo educación primaria o universitaria"
- Si la educación se asigna aleatoriamente, entonces podríamos comparar el ingreso promedio de quienes reciben educación universitaria con aquellos que reciben educación primaria.



Sesgo por variables omitidas

- En ausencia de aleatorización de la educación surgen dos potenciales sesgos al comparar el ingreso laboral promedio de ambos grupos de educación.
- Sesgos provocados por diferencias en características observables entre ambos grupos (se corrigen controlando por esas características)
- Sesgos provocados por diferencias en características no observables entre ambos grupos.
- Como ejemplo piense que un sesgo positivo surgiría si personas con mayor capacidad de ingresos (más productivos, más hábiles, etc.) obtuvieran más educación.
- En este caso la comparación del ingreso promedio de ambos grupos capturaría no solo el retorno a la educación sino también las diferencias en habilidad.

Sesgo por variables omitidas

- La ecuación estructural típica para el salario es como sigue,

$$log(\text{wage}) = \beta_0 + \beta_1 x + \beta_2 x^2 + \beta_3 s + \gamma h + v$$

  donde x representa años de experiencia en el mercado de trabajo, s representa años de educación formal y h es la habilidad natural del individuo. v es el error estructural que satisface el supuesto de exogeneidad estricta: $E(v|x, s, h) = 0$.
- La teoría económica establece un perfil salarial creciente y cóncavo en experiencia, sugiriendo que $\beta_1 > 0$ y $\beta_2 < 0$ en la ecuación anterior. Además, la teoría del capital humano sugiere una relación directa entre salario y educación ($\beta_3 > 0$) y entre salario y habilidad ($\gamma > 0$).

Sesgo por variables omitidas

- Empíricamente el problema para estimar una ecuación salarial como la anterior es que la habilidad de una persona no es observable. Por lo tanto, la ecuación estimable es,

$$log(\text{wage}) = \beta_0 + \beta_1 x + \beta_2 x^2 + \beta_3 s + u$$

  donde $u = \gamma h + v$. En este modelo, en general $E(u|x, s) \neq 0$ debido a la probable correlación entre los años de educación y la habilidad de una persona.
- En econometría una variable explicativa x_j se dice que es endógena si está correlacionada con el error de la ecuación.
- Empíricamente, la endogeneidad aparece frecuentemente cuando se presenta el problema descripto para la ecuación del salario. Es decir cuando tenemos el denominado problema de las variables omitidas.

Agenda

- Considere el siguiente modelo estructural,

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_k x_k + \gamma q + v \tag{1}$$

donde $E(v|x_1, x_2, \ldots, x_k, q) = 0$ y q es la variable no observada.

- Nuestro interés es estimar correctamente los β's que son los efectos parciales de las variables observadas manteniendo constantes el resto de las variables explicativas, incluyendo a q.
- El modelo que se podría estimar es,

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_k x_k + u \tag{2}$$

donde $u = \gamma q + v$.

- Sin pérdida de generalidad se puede asumir que $E(q) = 0$ de forma tal que $E(u) = 0$.

Mínimos Cuadrados Clásicos

- En este modelo si q está correlacionada con alguna de las variables explicativas, entonces u estará correlacionado también y tenemos el problema de la endogeneidad.
- Sabemos que si no se satisface (al menos) el supuesto de exogeneidad contemporánea MCC no dará estimaciones consistentes de los parámetros.
- Escribamos la proyección lineal de q en las k variables explicativas observadas como,

$$q = \delta_0 + \delta_1 x_1 + \delta_2 x_2 + \cdots + \delta_k x_k + r \tag{3}$$

  donde, por definición de proyección lineal, $E(r) = 0$ y $Cov(x_j, r) = 0$, $j = 1, 2, \ldots, k$.
- Sustituyendo la ecuación (3) en la (1) podemos ver que estimaría MCC aplicado sobre la ecuación (2).

- $$y = (\beta_0 + \gamma\delta_0) + (\beta_1 + \gamma\delta_1)x_1 + (\beta_2 + \gamma\delta_2)x_2 + \cdots + (\beta_k + \gamma\delta_k)x_k + v + \gamma r$$

  donde el error $v + \gamma r$ cumple con el supuesto de exogeneidad estricta.
- De la ecuación anterior surge claramente que MCC aplicado en (2) dará estimadores consistentes, $\hat{\beta}_j$, de los parámetros $\beta_j + \gamma\delta_j$.
- Esta especificación es la más general que se puede tener. Muchas veces en la práctica la variable omitida solo está relacionada con alguna de las variables explicativas observadas. Esta especificación puede obtenerse haciendo ceros a los δ_j correspondientes en la ecuación (3) arriba.

La Ecuación del Salario con Habilidad no Observada

- Volviendo al ejemplo del salario,

$$log(\text{wage}) = \beta_0 + \beta_1 x + \beta_2 x^2 + \beta_3 s + \gamma h + v$$

  la teoría económica sugiere correlación entre s y h. Las personas con mayor habilidad alcanzan una educación más alta. Supongamos que la relación entre la habilidad y la educación es $h = \pi_0 + \pi_1 s$.

- Si omitimos h en la estimación,

$$log(\text{wage}) = \beta_0 + \beta_1 x + \beta_2 x^2 + \beta_3 s + u$$

  MCC dará un estimador de β_3 sesgado y no consistente.

La Ecuación del Salario con Habilidad no Observada

- Reemplazando h en la ecuación del salario tenemos,

$$log(\text{wage}) = (\beta_0 + \pi_0\gamma) + \beta_1 x + \beta_2 x^2 + (\beta_3 + \pi_1\gamma)s + v$$

- En este ejemplo particular, MCC da estimaciones insesgadas y consistentes de β_1 y β_2 pero no asi de β_3. Si $\pi_1 > 0$ como sugiere la teoría económica, entonces MCC sobre-estimará el coeficiente asociado con los retornos a la educación.

Agenda

Solución usando variables proxy

- El problema de las variables omitidas puede ser solucionado si existe una variable proxy para la variable no observada.
- Los requerimientos formales para que una variable pueda ser considerada proxy de otra son dos.
  1. La variable proxy debe ser redundante en la ecuación estructural. Si w es una variable proxy para q, el requerimiento de redundancia establece que $E(y|x, q, w) = E(y|x, q)$.
  2. La correlación entre la variable omitida q y cada x_j debe ser cero una vez que tomamos en consideración w. En términos de una proyección lineal este supuesto establece que $L(q|1, x_1, \ldots, x_k, w) = L(q|1, w)$.
- Es útil escribir el segundo punto en términos de una ecuación con error,

$$q = \theta_0 + \theta_1 w + r \tag{4}$$

donde por definición $E(r) = 0$ y $Cov(w, r) = 0$. Si w es una variable proxy razonable de q, entonces $\theta_1 \neq 0$. La condición 2 de arriba establece además que $Cov(x_j, r) = 0 \;\; \forall j$.

- Para obtener una ecuación estimable podemos reemplazar (4) en (1),

$$y = (\beta_0 + \gamma\theta_0) + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_k x_k + \gamma\theta_1 w + (\gamma r + v)$$

  donde, bajo los supuestos realizados, el error de la ecuación, $\gamma r + v$, no está correlacionado con x_j, $\forall j$; redundancia establece que w no está correlacionado con v y por definición w no está correlacionado con r.
- Por lo tanto el error satisface el supuesto de exogeneidad contemporánea y MCC aplicados en la ecuación anterior da estimaciones consistentes de $(\beta_0 + \gamma\theta_0)$, $\beta_1, \ldots, \beta_k$ y $\gamma\theta_1$.
- Entonces, bajo los supuestos de variables proxy, MCC estima en forma consistente el efecto parcial de las variables explicativas observadas (x_j). En particular, en el ejemplo del salario empíricamente se utilizan los resultados de tests de inteligencia como proxy de habilidad.

Agenda

Solución usando variables instrumentales

- Las variables instrumentales son una forma de solucionar el problema de endogeneidad de las variables explicativas. En este sentido, es una corrección más general que la de las variables proxy porque no solo puede aplicarse en el caso de variables omitidas, sino en cualquier caso en el que exista endogeneidad de algún regresor (i.e. error de medición en variables explicativas, causalidad simultánea, etc.).
- Considere el siguiente modelo,

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_k x_k + u \tag{5}$$

  donde $E(u) = 0$ y $Cov(u, x_j) = 0$, $j = 1, 2, \ldots, k-1$. En palabras, x_k es potencialmente endógena en (5).
- Por lo tanto, MCC aplicados a (5) nos dará estimadores inconsistentes.

Solución usando variables instrumentales

- Para usar el enfoque de IV con x_k endógena, necesitamos una variable observable z_1, que no esté en (5) y que satisfaga dos condiciones:
    1. z_1 es una variable exógena en (5), es decir $Cov(z_1, u) = 0$.
    2. $\theta_1 \neq 0$ en la proyección lineal de la variable endógena, x_k, sobre todas las variables exógenas,

$$x_k = \delta_0 + \delta_1 x_1 + \cdots + \delta_{k-1} x_{k-1} + \theta_1 z_1 + r_k \tag{6}$$

    donde, por definición, $E(r_k) = 0$ y r_k no está correlacionado con $x_1, x_2, \ldots, x_{k-1}, z_1$.
- En palabras, z_1 está parcialmente correlacionada con x_k una vez que el resto de las variables exógenas han sido tomadas en cuenta.
- Cuando z_1 satisface estas dos condiciones se dice una variable instrumental para x_k. La proyección lineal (6) se denomina ecuación de forma reducida para la variable endógena x_k.

Solución usando variables instrumentales

- Reemplazando (6) en (5) tenemos,

$$y = \alpha_0 + \alpha_1 x_1 + \alpha_2 x_2 + \cdots + \alpha_{k-1} x_{k-1} + \lambda_1 z_1 + v \tag{7}$$

  donde $v = u + \beta_k r_k$, $\alpha_j = \beta_j + \beta_k \delta_j$ y $\lambda_1 = \beta_k \theta_1$.
- Por nuestros supuestos, v no está correlacionado con ninguna de las variables explicativas de (7) y por lo tanto MCC estima consistentemente los parámetros de la ecuación reducida de y.
- Algunas veces, estimar los parámetros de la ecuación reducida (7) tiene interés en si mismo pero en general se trata de estimar en forma consistente los parámetros de (5). Los supuestos hechos para IV también lo permiten.

Solución usando variables instrumentales

- Para ver esto formalmente, escribamos (5) como,

$$y = x\beta + u$$

  donde $x = (1, x_1, \ldots, x_k)$ y $\beta' = (\beta_0, \beta_1, \ldots, \beta_k)$ son de dimensión $1 \times k + 1$.
- Definamos el vector de variables exógenas $z = (1, x_1, \ldots, x_{k-1}, z_1)$ y el vector de parámetros de la ecuación reducida $\delta' = (\delta_0, \delta_1, \ldots, \delta_{k-1}, \theta_1)$.
- Bajo los supuestos de (5) y el supuesto de que la variable instrumental z_1 es exógena se cumplen las siguientes condiciones de ortogonalidad:

$$E(z'u) = 0.$$

Solución usando variables instrumentales

- Por lo tanto,

$$\begin{aligned} z'y &= z'x\beta + z'u \Rightarrow E(z'y) = E(z'x)\beta + E(z'u) \\ &\Rightarrow E(z'y) = E(z'x)\beta \end{aligned} \tag{8}$$

  Donde $E(z'x)$ es de dimensión $k+1 \times k+1$, y $E(z'y)$ es de dimensión $k+1 \times 1$.

- La última expresión representa un sistema de $k+1$ ecuaciones lineales con $k+1$ incógnitas. El sistema tiene una solución única si y solo sí la matriz $E(z'x)$ tiene rango completo (i.e. rango $E(z'x) = k+1$).

Solución usando variables instrumentales

- Note que $x = z\pi + r$ con

$$\pi = \begin{bmatrix} 1 & 0 & \cdots & 0 & \delta_0 \\ 0 & 1 & \cdots & 0 & \delta_1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & \delta_{k-1} \\ 0 & 0 & \cdots & 0 & \theta_1 \end{bmatrix}$$

  y $r = (0, \ldots, 0, r_k)$.
- Entonces $E(z'x) = E(z'(z\pi + r)) = E(z'z)\pi$ de forma tal que para que $E(z'x)$ tenga rango completo necesitamos que $E(z'z)$ tenga rango $k+1$, que es un supuesto estándar, y que π tenga rango $k+1$ que está garantizado por el supuesto 2 de IV ($\theta_1 \neq 0$).
- En este caso, la solución del sistema de ecuaciones está dada por:

$$\beta = [E(z'x)]^{-1}E(z'y)$$

Solución usando variables instrumentales

- Utilizando los análogos muestrales se obtiene el estimador de variables instrumentales,

$$\begin{aligned}\hat{\beta} &= \left(\frac{1}{n}\sum_{i=1}^{n} z_i' x_i\right)^{-1} \frac{1}{n}\sum_{i=1}^{n} z_i' y_i \\ &= (z'x)^{-1} z'y = (z'x)^{-1} z'(x\beta + u) \\ &= \beta + (z'x)^{-1} z'u \end{aligned} \tag{9}$$

- Usando la WLLN,

$$\left(\frac{1}{n}\sum_{i=1}^{n} z_i' x_i\right) \xrightarrow{p} E(z'x)$$

que bajo los supuestos de IV tiene rango completo.

Solución usando variables instrumentales

- Además,

$$\frac{1}{n}\sum_{i=1}^{n} z_i' u_i \xrightarrow{p} E(z'u) = 0$$

  y el estimador de IV es consistente.
- Volviendo al ejemplo de la ecuación de salarios, la omisión de la habilidad (h) provoca que la variable que mide educación (s) sea endógena en el modelo. Para obtener estimaciones consistentes en la ecuación salarial necesitamos un instrumento para s.
- Card(1995), por ejemplo, utiliza una variable binaria que indica si una persona creció en el vecindario de una universidad como variable instrumental de años de educación.

Agenda

Mínimos Cuadrados en dos Etapas

- Considere nuevamente la ecuación (5) con todos sus supuestos. Es decir, suponga que existe endogeneidad potencial de x_k.
- Supongamos que tenemos más de una variable instrumental para x_k. En particular, supongamos que z_1, z_2, $\ldots$, z_M son variables tal que,

$$Cov(z_h, u) = 0, \quad h = 1, 2, \ldots, M. \tag{10}$$

  cada z_h es exógena en la ecuación (5).
- Si cada una de estas variables tiene alguna correlación parcial con x_k, tenemos M potenciales instrumentos.
- En realidad, hay muchos más que M porque cualquier combinación lineal de $x_1, x_2, \ldots, x_{k-1}, z_1, z_2, \ldots, z_M$ no tiene correlación con u. Qué instrumento deberíamos utilizar?
- Bajo ciertos supuestos Mínimos cuadrados en dos etapas (2SLS) es el estimador de IV más eficiente.

- Definamos el vector de variables exógenas como antes, $z = (1, x_1, x_2, \dots, x_{k-1}, z_1, z_2, \dots, z_M)$ un vector de dimensión $1 \times L$ con $L = k + M$.
- Definamos la proyección lineal de la variable endógena sobre todas las variables exógenas,

$$x_k = \delta_0 + \delta_1 x_1 + \cdots + \delta_{k-1} x_{k-1} + \theta_1 z_1 + \cdots + \theta_M z_M + r_k \tag{11}$$

  donde, por definición, $E(r_k) = 0$ y r_k no está correlacionado con ninguna de las variables en el lado derecho de la ecuación.
- Como ninguna combinación lineal de z está correlacionada con u

$$x_k^* = \delta_0 + \delta_1 x_1 + \cdots + \delta_{k-1} x_{k-1} + \theta_1 z_1 + \cdots + \theta_M z_M \tag{12}$$

  tampoco lo estará.
- Si observáramos x_k^* podríamos utilizarla como instrumento para x_k en (5).

Mínimos Cuadrados en dos Etapas

- Sin embargo, si no existen dependencias lineales exactas entre las variables exógenas se podrían estimar en forma consistente por MCC los parámetros de (11) y definir para cada observación i,

$$\hat{x}_{i,k} = \hat{\delta}_0 + \hat{\delta}_1 x_{i,1} + \cdots + \hat{\delta}_{k-1} x_{i,k-1} + \hat{\theta}_1 z_{i,1} + \cdots + \hat{\theta}_M z_{i,M} \tag{13}$$

- Ahora para cada observación i definamos el vector $\hat{x}_i \equiv (1,\, x_{i,1},\, \ldots,\, x_{i,k-1}, \hat{x}_{i,k})$ y estimemos por IV,

$$\hat{\beta} = \left(\sum_{i=1}^{n} \hat{x}_i' x \right)^{-1} \sum_{i=1}^{n} \hat{x}_i' y = (\hat{x}' x)^{-1} \hat{x}' y. \tag{14}$$

- Este estimador IV es también un estimador de MCC.

$$\hat{x} = z\hat{\delta} = z(z'z)^{-1} z' x = P_z x$$

con P_z una matriz idempotente y simétrica.

Mínimos Cuadrados en dos Etapas

- Por lo tanto, $\hat{x}'x = x'P_z x = (P_z x)'P_z x = \hat{x}'\hat{x}$. Reemplazando esta última expresión en (14) se obtiene,
$$\hat{\beta} = (\hat{x}'\hat{x})^{-1}\hat{x}'y. \tag{15}$$
- El término, mínimos cuadrados en dos etapas viene de este procedimiento.
- Entonces $\hat{\beta}$ se puede obtener con los siguientes pasos
    1. Obtenga $\hat{x}_k$ de la regresión de x_k sobre $x_1, \ldots, x_{k-1}, z_1, \ldots, z_M$. Este paso se denomina regresión de la primera etapa.
    2. Estime por MCC una regresión de y sobre $x_1, \ldots, x_{k-1}, \hat{x}_k$. Esta es la regresión de la segunda etapa.
- El estimador de 2SLS y el de IV son idénticos si solo existe un instrumento para x_k.

Mínimos Cuadrados en dos Etapas

- En términos generales podemos resumir los resultados de 2SLS como sigue. Considere el modelo,

$$y = x\beta + u$$

  donde x es de dimensión $1 \times k$ y varios elementos de x pueden estar potencialmente correlacionados con u.
- Supuesto 1: Para algún vector $1 \times L$, z, $E(z'u) = 0$.
  Note que el supuesto de exogeneidad estricta $E(u|z) = 0$ implica el supuesto 1.
- Supuesto 2: (a) rango $E(z'z) = L$; (b) rango $E(z'x) = k$.
  Técnicamente, la parte (a) del supuesto 2 es necesaria pero no especialmente importante. La parte (b) del supuesto es la realmente importante porque es la condición de rango que permite la identificación de los parámetros del modelo.

Mínimos Cuadrados en dos Etapas

- Definamos el vector de variables exógenas como antes, $z = (1, x_1, x_2, \ldots, x_{k-1}, z_1, z_2, \ldots, z_M)$ un vector de dimensión $1 \times L$ con $L = k + M$.
- Definamos la proyección lineal de x sobre z como $x = x^* + r$ con $x^* = z\pi = z[E(z'z)]^{-1}E(z'x)$ y $r = (0, \ldots, 0, r_k)$.
- Entonces, multiplicando el modelo por $x^{*'}$ y tomando esperanzas tenemos,

$$E(x^{*'}y) = E(x^{*'}x)\beta + E(x^{*'}u) = E(x^{*'}x)\beta$$

  y β está identificado por $\beta = [E(x^{*'}x)]^{-1}E(x^{*'}y)$ si $E(x^{*'}x)$ no es singular.
- Ahora $E(x^{*'}x) = \pi' E(z'x) = E(z'x)'[E(z'z)]^{-1}E(z'x)$ y esta matriz no es singular si $E(z'x)$ tiene rango k, lo que está garantizado si se cumple el supuesto 2 (b).

- El estimador de 2SLS se puede escribir con la ecuación (14) o con,

$$\begin{aligned}
\hat{\beta} &= \left[\left(\sum_{i=1}^{n} x_i' z_i\right)\left(\sum_{i=1}^{n} z_i' z_i\right)^{-1}\left(\sum_{i=1}^{n} z_i' x_i\right)\right]^{-1} \\
&\times \left(\sum_{i=1}^{n} x_i' z_i\right)\left(\sum_{i=1}^{n} z_i' z_i\right)^{-1}\left(\sum_{i=1}^{n} z_i' y_i\right) \\
&= \beta + \left[\left(n^{-1}\sum_{i=1}^{n} x_i' z_i\right)\left(n^{-1}\sum_{i=1}^{n} z_i' z_i\right)^{-1}\left(n^{-1}\sum_{i=1}^{n} z_i' x_i\right)\right]^{-1} \\
&\times \left(n^{-1}\sum_{i=1}^{n} x_i' z_i\right)\left(n^{-1}\sum_{i=1}^{n} z_i' z_i\right)^{-1}\left(n^{-1}\sum_{i=1}^{n} z_i' u_i\right) \qquad (16)
\end{aligned}$$

- El estimador de 2SLS es consistente aplicando la WLLN a la ecuación anterior. Además apropiadamente re-escalado es asintóticamente normal.
- La normalidad asintótica de $\sqrt{n}(\hat{\beta} - \beta)$ se sigue de la normalidad asintótica de $n^{-1/2} \sum_{i=1}^{n} z_i' u_i$, que sigue del CLT bajo el supuesto 1,

$$n^{-1/2} \sum_{i=1}^{n} z_i' u_i \sim \text{Normal}(0, \sigma^2 E(z'z))$$

- Si agregamos el supuesto de varianza de los errores esférica podemos derivar la matriz de varianzas y covarianzas de los estimadores de 2SLS.
- Supuesto 3: $E(uu'|z) = \sigma^2$.
-

$$\sqrt{n}(\hat{\beta} - \beta) \xrightarrow{d} \text{Normal}(0, \sigma^2 \{E(x'z)[E(z'z)]^{-1} E(z'x)\}^{-1}) \tag{17}$$

Mínimos Cuadrados en dos Etapas

- La estimación de la matriz de varianzas y covarianzas asintótica de los estimadores de 2SLS se obtiene usando los análogos muestrales de las esperanzas y estimando consistentemente σ^2.
- Definiendo los residuos de la estimación de 2SLS como,

$$\hat{u}_i = y_i - x_i\hat{\beta}, \quad i = 1, 2, \ldots, n.$$

- La estimación consistente de σ^2 viene dada por,

$$\hat{\sigma}^2 = (n-k)^{-1}\sum_{i=1}^{n}\hat{u}_i^2$$

- Por lo tanto $\hat{\sigma}^2(\sum_{i=1}^{n}\hat{x}_i'\hat{x}_i)^{-1} = \hat{\sigma}^2(\hat{x}'\hat{x})^{-1}$ es un estimador válido de la matriz de varianzas y covarianzas de los estimadores de 2SLS.

Mínimos Cuadrados en dos Etapas

- Bajo los supuestos 1, 2 y 3, los estimadores de 2SLS son eficientes dentro de la clase de todos los estimadores de IV que usan instrumentos lineales en z.
- Es posible detectar la presencia de variables explicativas endógenas?
- Test de Hausman (1978)
  Idea: bajo la hipótesis nula de no existencia de endogeneidad, el estimador de MCC, $\hat{\beta}_{MCC}$, y el estimador de variables instrumentales, $\hat{\beta}_{2SLS}$, son estimadores consistentes de β, y el estimador de MCC es el más eficiente.
- Si la hipótesis nula es falsa, el estimador de variables instrumentales, $\hat{\beta}_{2SLS}$, es el único consistente.
- Entonces, bajo la hipótesis nula ambos estimadores deberían diferir solo por error muestral. Es decir, aceptar la hipótesis nula del test es evidencia en favor de exogeneidad.

Mínimos Cuadrados en dos Etapas

- Hausman sugiere utilizar un test de Wald. Supongamos que V_{2SLS} es la matriz de varianzas y covarianzas asintótica del estimador de variables instrumentales y V_{MCC} es la correspondiente al estimador de MCC. Entonces,

$$H = (\hat{\beta}_{MCC} - \hat{\beta}_{2SLS})'[V_{2SLS} - V_{MCC}]^{-1}(\hat{\beta}_{MCC} - \hat{\beta}_{2SLS}) \sim \chi^2(q)$$

  donde q es la dimensión del vector $\hat{\beta}_{MCC}$.

- Si hay una sola variable potencialmente endógena, x_k, el estadístico de Hausman se reduce a,

$$t_H = \frac{(\hat{\beta}_{k,MCC} - \hat{\beta}_{k,2SLS})}{\sqrt{V_{k,2SLS} - V_{k,MCC}}} \sim \text{Normal}(0,\, 1)$$

Función de Control

- En general el método de la función de control se utiliza para manejar la endogeneidad en modelos no lineales.
- El enfoque de la función de control utiliza regresores adicionales para romper la correlación entre las variables explicativas endógenas y los errores noobservables que afectan la variable dependiente.
- Como en el método de VI (o 2SLS) este enfoque también descansa en la existencia de variables exógenas que no aparecen en la ecuación estructural.
- Supongamos que y_1 es la variable dependiente, y_2 es una variable explicativa endógena y $\mathbf{z}$ es un vector de dimensión $1 \times L$ de variables exógenas (incluye un 1 en el primer elemento para la constante).
- Considere el siguiente modelo:

$$y_1 = \mathbf{z}_1 \boldsymbol{\delta}_1 + \alpha_1 y_2 + u_1 \tag{18}$$

con $\mathbf{z}_1$ un subconjunto $1 \times L_1$ de $\mathbf{z}$.

Función de Control

- Las variables en $\mathbf{z}$ son exógenas en el mismo sentido que con VI, es decir $E(\mathbf{z}'u_1) = 0$.
- Por lo tanto se puede estimar $(\boldsymbol{\delta}_1, \alpha_1)$ en forma consistente usando 2SLS (agregando la condición de rango estándar).
- En el método de la función de control la forma reducida de la variable endógena juega un rol fundamental:

$$
\begin{aligned}
y_2 &= \mathbf{z}\pi_2 + v_2 && (19) \\
\mathrm{E}\left(\mathbf{z}'v_2\right) &= \mathbf{0} && (20)
\end{aligned}
$$

donde π_2 es $L \times 1$

- En este modelo la endogenidad de y_2 aparece si y solo si u_1 está correlacionado con v_2.
- Consideremos la proyección lineal de u_1 sobre v_2:

$$
u_1 = \rho_1 v_2 + e_1 \tag{21}
$$

donde $\rho_1 = \mathrm{E}\left(v_2 u_1\right) / \mathrm{E}\left(v_2^2\right)$ es el coeficiente poblacional.

Función de Control

- Por definición $\mathrm{E}\left(v_2 e_1\right) = 0$, y $\mathrm{E}\left(\mathbf{z}' e_1'\right) = \mathbf{0}$ por que u_1 y v_2 no estan correlacionados con $\mathbf{z}$.
- Reemplazando (21) en (18) tenemos

$$y_1 = \mathbf{z}_1 \boldsymbol{\delta}_1 + \alpha_1 y_2 + \rho_1 v_2 + e_1 \tag{22}$$

  donde v_2 aparece como variable explicativa.
- Note que e_1 no está correlacionado con v_2 ni con $\mathbf{z}$ y como y_2 es una función lineal de $\mathbf{z}$ y v_2, e_1 tampoco está correlacionado con y_2.
- La ecuación (22) sugiere una forma de estimar en forma consistente $(\boldsymbol{\delta}_1, \alpha_1)$: MCC
- Único problema: no observamos v_2
- El procedimiento entonces involucra dos pasos:
  1. Estimar por MCC la ecuación (19) y construir los residuos $\hat{v}_2$
  2. Estimar por MCC la ecuación (22) reemplazando v_2 por $\hat{v}_2$.

- Note que en el procedimiento anterior tenemos

$$\begin{aligned} y_1 &= \mathbf{z}_1\boldsymbol{\delta}_1 + \alpha_1 y_2 + \rho_1 v_2 + e_1 \\ &= \mathbf{z}_1\boldsymbol{\delta}_1 + \alpha_1 y_2 + \rho_1(y_2 - \mathbf{z}\pi_2) + e_1 \\ &= \mathbf{z}_1\boldsymbol{\delta}_1 + \alpha_1 y_2 + \rho_1(y_2 - \mathbf{z}\pi_2 \pm \mathbf{z}\hat{\pi}_2) + e_1 \\ &= \mathbf{z}_1\boldsymbol{\delta}_1 + \alpha_1 y_2 + \rho_1(y_2 - \mathbf{z}\hat{\pi}_2) + e_1 + \rho_1(\mathbf{z}\hat{\pi}_2 - \mathbf{z}\pi_2) \\ &= \mathbf{z}_1\boldsymbol{\delta}_1 + \alpha_1 y_2 + \rho_1 \hat{v}_2 + \text{error} \end{aligned} \tag{23}$$

  donde error $= e_1 + \rho_1(\mathbf{z}\hat{\pi}_2 - \mathbf{z}\pi_2)$ depende del error muestral de $\hat{\pi}_2$, salvo que $\rho_1 = 0$.
- Esto implica que la matriz de varianzas y covarianzas de los coeficientes estimados en el segundo paso del procedimiento deberá tomar en cuenta este error muestral.
- La estimación por MCC de (23) es una ejemplo de **estimador de la función de control.**
- La inclusión de los residuos $\hat{v}_2$ "controla" por la endogeneidad de y_2 en la ecuación original (aunque lo hace con error muestral porque $\pi_2 \neq \hat{\pi}_2$).

Función de Control

- Como comparan los enfoques de 2SLS y función de control?
- Se puede mostrar algebraicamente que los estimadores de $(\boldsymbol{\delta}_1, \alpha_1)$ son exactamente los mismos.
- La ecuación (23) provee otra forma de contrastar por endogeneidad: $H_0 : \rho_1 = 0$ versus $H_1 : \rho_1 \neq 0$ con un estadístico t de significatividad individual.
- Se puede mostrar que este contrate es igual al test de Hausman que describimos para variables instrumentales.

Agenda

La Política

- Recordemos el ejemplo de una política educativa que consiste en reducir el tamaño de las clases en la educación primaria y su efecto sobre el aprendizaje de los alumnos.
- La política es tener clases con menos alumnos por profesor. La variable que mide el aprendizaje son las notas en pruebas estandarizadas de fin de año.
- En la práctica, la implementación más común de esta medida se realiza en dos etapas:
  1. 1ra Etapa: se eligen aleatoriamente algunas escuelas de la población de escuelas.
  2. 2da Etapa: Se asignan aleatoriamente las escuelas elegidas en la primera etapa a los grupos de tratamiento y de control.
- En este contexto, los alumnos de una misma escuela o clase tienden a tener puntajes en las pruebas que están correlacionados ya que están sujetos a algunas de las mismas influencias ambientales y de origen familiar.

La Política

- En términos matemáticos el modelo es

$$y_i = \alpha + \beta s_i + u_i \tag{24}$$

donde y_i es la nota del alumno i y s_i es la asignación aleatoria de la política.

- Hasta ahora, para hacer inferencia estadística, supusimos que u_i tenía varianza constante y covarianzas iguales a cero.
- Lo que sucede con la implementación de la política es que ahora la $cov(u_i, u_j) \neq 0$ si los alumnos i y j pertenecen a la misma escuela.
- ¿Qué sucede con la inferencia en este modelo si pasa esto?

Agenda

Propiedades de MCC

- Considere el siguiente modelo,

$$y = x\beta + u, \;\; E(u|x) = 0, \;\; E(uu'|x) = \sigma^2\Omega. \tag{25}$$

- El estimador de MCC es,

$$\begin{aligned} \hat{\beta} &= (x'x)^{-1}x'y \\ &= \beta + (x'x)^{-1}x'u \end{aligned} \tag{26}$$

- Los estimadores de MCC son insesgados y consistentes,

$$\begin{aligned} E(\hat{\beta}|x) &= \beta + (x'x)^{-1}x'E(u|x) \\ &= \beta \end{aligned} \tag{27}$$

y aplicando la ley de expectativas iteradas, $E(\hat{\beta}) = \beta$.

Propiedades de MCC

- $$\begin{aligned} \hat{\beta} &= \beta + \left(n^{-1} \sum_{i=1}^{n} x_i' x_i \right)^{-1} n^{-1} \sum_{i=1}^{n} x_i' u_i \\ &\xrightarrow{p} \beta + [E(x'x)]^{-1} E(x'u) \\ &\xrightarrow{p} \beta \end{aligned} \tag{28}$$

  donde la primera convergencia se obtiene aplicando la WLLN y la segunda sigue del supuesto de exogeneidad estricta.
- Apropiadamente re-escalado, los estimadores de MCC son asintóticamente normales,

$$\begin{aligned} \sqrt{n}(\hat{\beta} - \beta) &= + \left(n^{-1} \sum_{i=1}^{n} x_i' x_i \right)^{-1} n^{-1/2} \sum_{i=1}^{n} x_i' u_i \\ &\xrightarrow{d} \text{Normal}(0, [E(x'x)]^{-1} E(x'uu'x)[E(x'x)]^{-1}) \end{aligned}$$

Propiedades de MCC

- La matriz de varianzas y covarianzas asintótica de $\hat{\beta}$ puede escribirse como,

$$\begin{aligned} Var[\sqrt{n}(\hat{\beta}-\beta)] &= [E(x'x)]^{-1}E[E(x'uu'x|x)][E(x'x)]^{-1} \\ &= [E(x'x)]^{-1}E[x'E(uu'|x)x][E(x'x)]^{-1} \\ &= \sigma^2[E(x'x)]^{-1}E[x'\Omega x][E(x'x)]^{-1} \end{aligned} \tag{29}$$

- Note que si la varianza de los errores fuera esférica, la ecuación anterior se reduce a la matriz de varianzas y covarianzas asintótica que obtuvimos para MCC: $\sigma^2[E(x'x)]^{-1}$.
- Entonces, una consecuencia de la heterocedasticidad y/o de la correlación serial es que matriz de varianzas y covarianzas asintótica convencional de MCC es incorrecta.

Agenda

Inferencia Estadística

- La inferencia estadística convencional de MCC no es válida en presencia de errores no esféricos.
- Para poder hacer inferencia estadística en este modelo necesitamos tests estadísticos robustos ante la presencia de heterocedasticidad y/o correlación serial.
- Para esto necesitamos estimar la matriz de varianzas y covarianzas asintótica de MCC correcta,

$$\widehat{Var(\hat{\beta})} = \left(\sum_{i=1}^{n} x_i' x_i\right)^{-1} \left(\sum_{i=1}^{n} x_i' \hat{u}_i \hat{u}_i' x_i\right) \left(\sum_{i=1}^{n} x_i' x_i\right)^{-1} \tag{30}$$

  donde $\hat{u}$ son los residuos de la estimación por MCC. Esta es la matriz de varianzas y covarianzas robusta ante la presencia de heterocedasticidad y correlación serial de White (1980).

- La raiz cuadrada de los elementos de la diagonal principal de (30) son los errores estándar robustos. Los estadísticos t se calculan de la forma usual con estos errores estándar robustos.
- Para realizar contrastes sobre combinación lineal de coeficientes el estadístico de Wald se construye con la fórmula habitual,

$$W = (R\hat{\beta} - r)'[R\widehat{Var(\hat{\beta})}R']^{-1}(R\hat{\beta} - r)/\#r \tag{31}$$

donde $\widehat{Var(\hat{\beta})}$ está definida por la ecuación (30)

Agenda

- Una alternativa a la estimación por MCC y la inferencia estadística robusta es utilizar el método de mínimos cuadrados generalizados (MCG).
- Considere la estimación del mismo modelo que antes,

$$y = x\beta + u, \;\; E(u|x) = 0, \;\; E(uu'|x) = \sigma^2\Omega.$$

y asumamos por un momento que los elementos de Ω son conocidos.

- Como Ω es definida positiva, su inversa también lo es. Por lo tanto, es posible encontrar una matriz no singular P tal que:

$$\Omega^{-1} = P'P. \tag{32}$$

- Premultiplicando (25) por P se obtiene,

$$y_* = x_*\beta + u_*, \tag{33}$$

donde $y_* = Py$, $x_* = Px$ y $u_* = Pu$.

Mínimos Cuadrados Generalizados

- El modelo transformado cumple con el supuesto de exogeneidad estricta,

$$\begin{aligned} E(u_*|x_*) &= E(u_*|x) = E(Pu|x) \\ &= PE(u|x) = 0 \end{aligned}$$

- Además, de acuerdo con (32), $\Omega = P^{-1}(P')^{-1}$. Por lo tanto,

$$\begin{aligned} Var(u_*) &= E(Puu'P') \\ &= \sigma^2 P\Omega P' \\ &= \sigma^2 PP^{-1}(P')^{-1}P' \\ &= \sigma^2 I_n \end{aligned} \tag{34}$$

  y el modelo transformado cumple con los supuestos del modelo de regresión lineal múltiple y puede ser estimado por MCC.

Mínimos Cuadrados Generalizados

- El método de MCG minimiza,

$$\begin{aligned} RSS(\hat{\beta}) &= (y_* - x_*\hat{\beta})'(y_* - x_*\hat{\beta}) \\ &= (y_*' - \hat{\beta}'x_*')(y_* - x_*\hat{\beta}) \\ &= y_*'y_* - \hat{\beta}'x_*'y_* - y_*'x_*\hat{\beta} + \hat{\beta}'x_*'x_*\hat{\beta} \\ &= y_*'y_* - 2\hat{\beta}'x_*'y_* + \hat{\beta}'x_*'x_*\hat{\beta} \end{aligned} \tag{35}$$

- El estimador de MCG es,

$$\begin{aligned} \hat{\beta} &= (x_*'x_*)^{-1}x_*'y_* \\ &= (x'\Omega^{-1}x)^{-1}x'\Omega^{-1}y \end{aligned} \tag{36}$$

- y usando la teoría desarrollada para MCC,

$$Var(\hat{\beta}) = \sigma^2(x_*'x_*)^{-1} = \sigma^2(x'\Omega^{-1}x)^{-1},$$

- Una estimación consistente de la varianza del estimador de MCG es,

$$\widehat{Var(\hat{\beta})} = s^2(x_*'x_*)^{-1} = s^2(x'\Omega^{-1}x)^{-1}, \tag{37}$$

con,

$$\begin{aligned} s^2 &= (y_* - x_*\hat{\beta})'(y_* - x_*\hat{\beta})/(n-K) \\ &= [P(y - x\hat{\beta})]'[P(y - x\hat{\beta})]/(n-K) \\ &= [(y - x\hat{\beta})]'\Omega^{-1}[(y - x\hat{\beta})]/(n-K). \end{aligned}$$

- La raiz cuadrada de los elementos de la diagonal principal de (37) son los errores estándar de los estimadores de MCG y pueden utilizarse para construir los estadísticos t.
- Restricciones lineales del tipo $H_0 : R\beta = r$ pueden contrastarse utilizando el test de Wald,

$$W = (R\hat{\beta} - r)'[Rs^2(x'\Omega^{-1}x)^{-1}R']^{-1}(R\hat{\beta} - r)/\#r \tag{38}$$

Mínimos Cuadrados Generalizados

- En términos generales, los resultados de MCG pueden resumirse de la siguiente manera.
- Para mostrar consistencia necesitamos reforzar el supuesto 3 de no singularidad.
  Supuesto 3': Ω es positiva definida y $E(x'\Omega^{-1}x)$ no es singular.
- Usando la WLLN,

$$n^{-1}\sum_{i=1}^{n} x_i'\Omega x_i \xrightarrow{p} E(x'\Omega^{-1}x) \equiv A$$

  y por el supuesto 3',

$$\left(n^{-1}\sum_{i=1}^{n} x_i'\Omega x_i\right)^{-1} \xrightarrow{p} A^{-1}$$

Mínimos Cuadrados Generalizados

- Usando la WLLN y el supuesto de exogeneidad estricta,

$$n^{-1}\sum_{i=1}^{n} x_i'\Omega u_i \xrightarrow{p} E(x'\Omega^{-1}u) = 0$$

- Prueba:

$$\begin{aligned} vec[E(x'\Omega^{-1}u)] &= [E(u' \otimes x')]vec(\Omega^{-1}) \\ &= [E(u \otimes x)']vec(\Omega^{-1}) = 0 \end{aligned}$$

- Por lo tanto,

$$\begin{aligned} \hat{\beta} &= \beta + (n^{-1}x'\Omega^{-1}x)^{-1}n^{-1}x'\Omega^{-1}u \\ &\xrightarrow{p} \beta + A^{-1}E(x'\Omega^{-1}u) = \beta \end{aligned} \tag{39}$$

y el estimador de MCG es consistente.

- La distribución asintótica se obtiene desde,

$$\begin{aligned}\sqrt{n}(\hat{\beta} - \beta) &= (n^{-1}x'\Omega^{-1}x)^{-1}n^{-1/2}x'\Omega^{-1}u \\ &\xrightarrow{d} A^{-1}\text{Normal}(0,\, \sigma^2 A) \\ &\xrightarrow{d} \text{Normal}(0,\, \sigma^2 A^{-1}) \qquad (40)\end{aligned}$$

donde se utilizó el CLT para,

$$n^{-1/2}\sum_{i=1}^{n} x_i'\Omega u_i \xrightarrow{d} \text{Normal}(0,\, \sigma^2 A)$$

con $A \equiv E(x'\Omega^{-1}x)$.

- Por lo tanto, la matriz de varianzas y covarianzas asintótica del estimador de MCG es,

$$Var(\hat{\beta}) = \sigma^2 A^{-1}/n$$

Mínimos Cuadrados Generalizados Estimados

- Remark: Note que para poder obtener la consistencia del estimador de MCG es necesario asumir que las variables explicativas son estríctamente exógenas. Este supuesto es más fuerte que el necesario para obtener consistencia de MCC que es, como vimos, exogeneidad contemporánea.
- Hasta ahora se asumió que Ω era conocida. En general, en la práctica, esto no es así y se necesita una estimación consistente de la misma.
- El método que utiliza una estimación consistente de Ω es conocido como mínimos cuadrados generalizados estimados (MCGE) ó feasible generalized least squares (FGLS).
- Vamos a mostrar, en lo que sigue, que MCGE es asintóticamente equivalente a MCG.

- El estimador de MCGE es,

$$\begin{aligned}
\tilde{\beta} &= (x'\hat{\Omega}^{-1}x)^{-1}x'\hat{\Omega}^{-1}y \\
&= \beta + (x'\hat{\Omega}^{-1}x)^{-1}x'\hat{\Omega}^{-1}u \\
&= \beta + (n^{-1}\sum_{i=1}^{n} x_i'\hat{\Omega}^{-1}x_i)^{-1}n^{-1}\sum_{i=1}^{n} x_i'\hat{\Omega}^{-1}u_i, \Rightarrow \\
\sqrt{n}(\hat{\beta} - \beta) &= (n^{-1}\sum_{i=1}^{n} x_i'\hat{\Omega}^{-1}x_i)^{-1}n^{-1/2}\sum_{i=1}^{n} x_i'\hat{\Omega}^{-1}u_i \qquad (41)
\end{aligned}$$

- Comparando el segundo término de (41) con el correspondiente a MCG tenemos,

$$n^{-1/2}\sum_{i=1}^{n} x_i'\hat{\Omega}^{-1}u_i - n^{-1/2}\sum_{i=1}^{n} x_i'\Omega^{-1}u_i = n^{-1/2}\sum_{i=1}^{n} x_i'(\hat{\Omega}^{-1} - \Omega^{-1})u_i$$

Mínimos Cuadrados Generalizados Estimados

- Note que la ecuación anterior puede escribirse como,

$$\left[n^{-1/2}\sum_{i=1}^{n}(u_i \otimes x_i)'\right] vec(\hat{\Omega}^{-1} - \Omega^{-1}) \quad = \quad o_p(1) \Rightarrow$$

$$n^{-1/2}\sum_{i=1}^{n} x_i'\hat{\Omega}^{-1}u_i = n^{-1/2}\sum_{i=1}^{n} x_i'\Omega^{-1}u_i + o_p(1) \tag{42}$$

- Usando el mismo argumento,

$$n^{-1}\sum_{i=1}^{n} x_i'\hat{\Omega}^{-1}x_i = n^{-1}\sum_{i=1}^{n} x_i'\Omega^{-1}x_i + o_p(1) \tag{43}$$

- Usando los resultados anteriores tenemos,

$$
\begin{aligned}
\sqrt{n}(\tilde{\beta} - \beta) &= \left(n^{-1} \sum_{i=1}^{n} x_i' \Omega^{-1} x_i \right) \left(n^{-1/2} \sum_{i=1}^{n} x_i' \Omega^{-1} u_i \right) + o_p(1) \\
&= \sqrt{n}(\hat{\beta} - \beta) + o_p(1) \\
&\Rightarrow \sqrt{n}(\tilde{\beta} - \hat{\beta}) = o_p(1) \qquad (44)
\end{aligned}
$$

que los estimadores de MCG y de MCGE son asintóticamente equivalentes ($\sqrt{n}$ asintóticamente equivalentes).

- Empíricamente, la equivalencia asintótica de los estimadores de MCG y MCGE implica que para realizar inferencia estadística sobre β usando MCGE, no hay que preocuparse de que $\hat{\Omega}$ sea un estimador de Ω.
- En otras palabras,

$$\sqrt{n}(\tilde{\beta} - \beta) \xrightarrow{d} \text{Normal}(0, \sigma^2 A^{-1})$$

- Bajo MCGE la estimación de A viene dada por,

$$\tilde{A} \equiv n^{-1} \sum_{i=1}^{n} x_i' \hat{\Omega} x_i$$

- En general la estimación de Ω se hace imponiendo alguna estructura en la matriz (e.g. heterocedasticidad o correlación serial).
- Por ejemplo, si hay heterocedasticidad, una estimación consistente de A es,

$$\tilde{A} = \hat{\Gamma}_0 + \sum_{j=1}^{m} \omega(j, m)(\hat{\Gamma}_j + \hat{\Gamma}_j'),$$

donde,

$$\hat{\Gamma}_j = \frac{1}{n} \sum_{i=j+1}^{n} x_i' \hat{u}_i \hat{u}_{i-j}' x_{i-j}, \;\; j = 0, 1, 2, \ldots, m$$

Mínimos Cuadrados Generalizados Estimados

- $\omega(j,m)$ es una ventana a definir y m es un parámetro de truncamiento.
- Por ejemplo,

$$\omega(j,m) = \begin{cases} 1 & j = 1,\, 2,\, \ldots,\, m. \qquad \text{uniform window} \\ 1 - \frac{j}{m+1}, & j = 1, 2, \ldots, m. \text{ Bartlett (Newey-West) window} \end{cases}$$

- $\hat{u}_i$ $(i = 1, 2, \ldots, n)$ son los residuos de la estimación por MCC.
- Cómo detectar la presencia de heterocedasticidad y/ó correlación serial en el modelo?
- Vamos a desarrollar el test de White para detectar heterocedasticidad y el test de Breusch-Godfrey para detectar correlación serial.

- Test de White

$$H_0 : \text{No existe heterocedasticidad}$$
$$H_1 : \text{Existe heterocedasticidad}$$

  Este contraste asume que la forma funcional de la heterocedasticidad es lineal en todas las variables explicativas del modelo, sus cuadrados y sus productos cruzados.

- Por ejemplo, suponga el siguiente modelo,

$$y_i = \alpha_0 + \alpha_1 x_{i,1} + \alpha_2 x_{i,2} + u_i$$

- La forma funcional de la heterocedasticidad de White es,

$$\sigma_i^2 = \gamma_0 + \gamma_1 x_{i,1} + \gamma_2 x_{i,2} + \gamma_3 x_{i,1}^2 + \gamma_4 x_{i,2}^2 + \gamma_5 x_{i,1} x_{i,2}$$

Mínimos Cuadrados Generalizados Estimados

- En este modelo la hipótesis nula se puede escribir como: $H_0 : \gamma_1 = \gamma_2 = \cdots = \gamma_5 = 0$, y la alternativa como, H_1 : al menos un $\gamma_i \neq 0$, $i = 1, 2, \ldots, 5$.
- El procedimiento de White es como sigue:
  1. Estimar el modelo original por MCC y obtener la serie de residuos y de sus cuadrados.
  2. Estimar la ecuación de la forma funcional de la varianza reemplazando la variable dependiente por la serie de residuos al cuadrado obtenida en el paso anterior.
  3. Construir el estadístico $LM = n \times R^2 \sim \chi^2_q$. El R^2 es el de la regresión del Paso 2 y q es el número de parámetros iguales a cero en la hipótesis nula.

Mínimos Cuadrados Generalizados Estimados

Heterocedasticidad: procedimiento de White

- Suponga el siguiente modelo estructural,

$$y_i = \alpha_0 + \alpha_1 x_{i,1} + \alpha_2 x_{i,2} + u_i \tag{45}$$

- Entonces, la forma funcional de la heterocedasticidad de White es,

$$\sigma_i^2 = \gamma_0 + \gamma_1 x_{i,1} + \gamma_2 x_{i,2} + \gamma_3 x_{i,1}^2 + \gamma_4 x_{i,2}^2 + \gamma_5 x_{i,1} x_{i,2} \tag{46}$$

- FGLS White
  1. Estimar (45) por OLS y obtener las estimaciones de los parámetros del modelo.
  2. Calcular los residuos del modelo y elevarlos al cuadrado, $\hat{u}_i^2$.
  3. Estimar (46) por OLS usando $\hat{u}_i^2$ como proxy de σ_i^2.
  4. Usar las estimaciones de la regresión auxiliar y obtener las variancias ajustadas como:

$$\widehat{\hat{u}_i^2} \equiv \hat{\sigma}_i^2 = \hat{\gamma}_0 + \hat{\gamma}_1 x_{i,1} + \hat{\gamma}_2 x_{i,2} + \hat{\gamma}_3 x_{i,1}^2 + \hat{\gamma}_4 x_{i,2}^2 + \hat{\gamma}_5 x_{i,1} x_{i,2}$$

- FGLS White (Cont.)
  5. Transformar las variables de (46) dividiendolas por $\hat{\sigma}_i^2$ y estimar por OLS,

$$\frac{\hat{u}_i^2}{\hat{\sigma}_i^2} = \gamma_0 \frac{1}{\hat{\sigma}_i^2} + \gamma_1 \frac{x_{i,1}}{\hat{\sigma}_i^2} + \gamma_2 \frac{x_{i,2}}{\hat{\sigma}_i^2} + \gamma_3 \frac{x_{i,1}^2}{\hat{\sigma}_i^2} + \gamma_4 \frac{x_{i,2}^2}{\hat{\sigma}_i^2} + \gamma_5 \frac{x_{i,1}x_{i,2}}{\hat{\sigma}_i^2} + \nu_i$$

  6. Con los $\tilde{\gamma}$'s de la estimación anterior calcular

$$\widehat{\frac{\hat{u}_i^2}{\hat{\sigma}_i^2}} \equiv \tilde{\sigma}_i^2 = \tilde{\gamma}_0 + \tilde{\gamma}_1 x_{i,1} + \tilde{\gamma}_2 x_{i,2} + \tilde{\gamma}_3 x_{i,1}^2 + \tilde{\gamma}_4 x_{i,2}^2 + \tilde{\gamma}_5 x_{i,1}x_{i,2}$$

  7. Usar $\frac{1}{\tilde{\sigma}_i}$ como ponderadores para estimar (45).

Mínimos Cuadrados Generalizados Estimados

- Correlación serial de orden uno
  Supongamos el mismo modelo de antes,

$$y_i = \alpha_0 + \alpha_1 x_{i,1} + \alpha_2 x_{i,2} + u_i$$

En este modelo, la correlación serial de orden uno se especifica como:

$$u_i = \rho u_{i-1} + \epsilon_i, \ |\rho| < 1.$$

- Test de Breusch-Godfrey

$$H_0 : \text{No existe correlación serial de orden uno } (\rho = 0)$$
$$H_1 : \text{Existe correlación serial de orden uno } (\rho \neq 0)$$

- El procedimiento de Breusch-Godfrey es como sigue:
  1. Estimar el modelo original por MCC y obtener la serie de residuos y la serie de residuos rezagada una observación.
  2. Estime una ecuación auxiliar que tenga como variable dependiente a la serie de residuos del paso anterior y como variables independientes a todas las variables explicativas del modelo original más la serie de residuos rezagada una observación.
  3. Construir el estadístico $LM = (n-1) \times R^2 \sim \chi^2_{q=1}$. El R^2 es el de la regresión del Paso 2 y $q = 1$ es el número de parámetros iguales a cero en la hipótesis nula.
- Si se rechazara la hipótesis nula una estimación consistente viene dada por el procedimiento de Cochrane-Orcutt.

Mínimos Cuadrados Generalizados Estimados

- Procedimiento de Cochrane-Orcutt
  Considere el mismo modelo de los ejemplos anteriores,

$$y_i = \alpha_0 + \alpha_1 x_{i,1} + \alpha_2 x_{i,2} + u_i \tag{47}$$

$$u_i = \rho u_{i-1} + \epsilon_i, \ |\rho| < 1. \tag{48}$$

1. Estimar el modelo original (47) por MCC y obtener la serie de residuos, $\hat{u}_i = y_i - \hat{\alpha}_0 - \hat{\alpha}_1 x_{i,1} - \hat{\alpha}_2 x_{i,2}$, y la serie de residuos rezagada una observación, $\hat{u}_{i-1}$.
2. Estimar la ecuación (48) por MCC, reemplazando a u_i y u_{i-1} por $\hat{u}_i$ y $\hat{u}_{i-1}$, respectivamente. Obtener $\hat{\rho}$.
3. Transforme las variables del siguiente modo: $y_i^* = y_i - \hat{\rho} y_{i-1}$, $x_{i,1}^* = x_{i,1} - \hat{\rho} x_{i-1,1}$, $x_{i,2}^* = x_{i,2} - \hat{\rho} x_{i-1,2}$ y $c^* = 1 - \hat{\rho}$.
4. Estimar por MCC el modelo transformado,

$$y_i^* = \alpha_0 c^* + \alpha_1 x_{i,1}^* + \alpha_2 x_{i,2}^* + u_i^*$$

y obtenga nuevos estimadores $\hat{\hat{\alpha}}_0$, $\hat{\hat{\alpha}}_1$, $\hat{\hat{\alpha}}_2$.

Mínimos Cuadrados Generalizados Estimados

- Procedimiento de Cochrane-Orcutt (continuación)

  5. Construir nuevos residuos $\hat{\hat{u}}_i = y_i - \hat{\hat{\alpha}}_0 - \hat{\hat{\alpha}}_1 x_{i,1} - \hat{\hat{\alpha}}_2 x_{i,2}$ y sus rezagos $\hat{\hat{u}}_{i-1}$.
  6. Vuelva al Paso 2 y repita el procedimiento.
  7. El procedimiento termina cuando en dos iteraciones sucesivas el valor estimado para ρ es el mismo.

- Los estimadores obtenidos con este procedimiento son insesgados, consistentes y asintóticamente eficientes.
- Remark: Tanto el test de Breusch-Godfrey como la corrección de Cochrane-Orcutt pueden generalizarse a correlación serial de orden mayor a uno.