IV, 2SLS, GMM

Advanced Microeconometrics

Instructor: Cristián Sánchez

Spring 2022

Introducción

Declaración de Principios

- Endogeneidad es la contrubición fundamental de la Econometría a la Estadística.
- Endogeneidad es lo que diferencia la optimización de la economía.
- Un curso de econometría sin revisar este tópico en detalle no es econometría.

Introducción

Considere el siguiente modelo lineal:

$$y = x\beta + u$$

donde y es escalar y x representa una observación de dimensión $1 \times K$, β es un vector de parámetros de $K \times 1$ y u es un término de error no observable.

El supuesto de identicación fundamental de Mínimos Cuadrados Ordinarios es

$$E(x'u) = 0$$

Note que el parámetro poblacional β puede ser expresado en momentos de las variables

observables explotando el supuesto recién presentado:

$$x'y = x'x\beta + x'u$$

Introducción

Tomando valor esperado tenemos que:

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

lo que se conoce como *Regresión Poblacional*. Dado que (y, x) es observable, β es *identificado*.

El principio de la analogía o *analogy principle* (Goldberger (1968), Manski (1968)) implica.

$$\hat{\beta}_{MCO} = \left[\frac{1}{n}\sum_{i=1}^{n} x_i' x_i\right]^{-1} \left[\frac{1}{n}\sum_{i=1}^{n} x_i' y_i\right] \equiv (X'X)^{-1}X'Y$$

Endogeneidad y Variables Instrumentales

Cuando se viola el supuesto de exogeneidad de las variables explicativas con respecto al término de error, las variables x incluyen un subconjunto de variables que son endógenas lo que significa que:

$$E(x'u) \neq 0$$

esto genera un problema de identificación. No es posible encontrar una expresión del parámetro poblacional β en función de momentos poblacionales de variables observables a no ser que contemos con otro set de variables z (en este caso de igual dimensión que x) que cumpla las siguientes condiciones:

$$E(z'u) = 0 \tag{1}$$

$$E(z'x) \neq 0 \tag{2}$$

Podemos pre-multiplicar la ecuación estructural por z' y obtener un sistema de ecuaciones:

$$z'y = z'x\beta + z'u$$

y por lo tanto obtenemos el siguiente sistema de ecuaciones:

$$E(z'x)\beta = E(z'y) \tag{3}$$

donde $E(z'x)$ es de orden $K \times K$ y $E(z'y)$ es de orden $K \times 1$.

Este sistema tiene solución única si la matriz $E(z'x)$ es invertible, lo cual sucede si el rango de ésta es completo e igual a K.

$$\beta = E(z'x)^{-1}E(z'y) \tag{4}$$

Luego si tenemos una muestra aleatoria (y_i, x_i, z_i) y siguiendo el *analogy principle* tenemos que el estimador de variables instrumentales está dado por:

$$\hat{\beta}_{IV} = \left[\frac{1}{n}\sum_{i=1}^{n} z_i' x_i\right]^{-1} \left[\frac{1}{n}\sum_{i=1}^{n} z_i' y_i\right] \equiv (Z'X)^{-1} Z'Y$$

La consistencia de este estimador sigue de la ecuación (4) y de alguna ley de grandes números. Note que podemos escribir el estimador de variables instrumentales como sigue:

$$\hat{\beta}_{IV} = \beta + \left[\frac{1}{n}\sum_{i=1}^{n} z_i' x_i\right]^{-1}\left[\frac{1}{n}\sum_{i=1}^{n} z_i' u_i\right]$$

Luego, se deduce claramente que $\text{plim}\beta_{IV} = \beta$ por álgebra de plims y la ley débil de los grandes números o *Weak law of large numbers* (WLLN). Ahora, podemos generar la expresión típica ajustada por $\sqrt{n}$:

$$\sqrt{n}(\hat{\beta}_{IV} - \beta) = \left[\frac{1}{n}\sum_{i=1}^{n} z_i' x_i\right]^{-1}\left[\frac{1}{n}\sum_{i=1}^{n} z_i' u_i\right]$$

donde el primer término del lado derecho de la ecuación convergerá a $E(z'x) = M_{zx}$ por WLLN.

El segundo término converge en distribución a una Normal por Teorema Central del Límite (CLT), en particular Lindeberg-Levy. Veamos un poco en detalle el segundo término:

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n} z_i' u_i \xrightarrow{d} \mathcal{N}(0, V_0)$$

donde $V_0 = E(u^2 z' z)$. Por lo tanto, por WLLN y CLT tenemos que

$$\sqrt{n}(\hat{\beta}_{IV} - \beta) \xrightarrow{d} \mathcal{N}(0, M_{zx}^{-1} V_0 (M_{zx}^{-1})')$$

Luego, para obtener un estimador de la varianza debemos usar el *sample analogue* de M_{zx} que es trivial y un estimador para V_0.

Múltiples instrumentos: 2SLS

Sea Z la matriz de instrumentos de orden $n \times L$ y X es la matriz de variables independientes o explicativas de orden $n \times K$ donde $L > K$. Sistema con infinitas soluciones.

Una manera de resolver este problema es post-multiplicando la matriz de instrumentos Z por otra matriz Π de orden $L \times K$. Luego, la matriz $Z\Pi$ es de dimensión $n \times K$ y el vector $z\Pi$ es de $1 \times K$ al igual que x.

Explotando la condición de identificación $E(z'u) = 0$ tenemos que se cumple que

$$\Pi' z' y = \Pi' z' x\beta + \Pi' z' u$$

Note que podemos identificar β tomando valor esperado,

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

Múltiples instrumentos: 2SLS

Por lo tanto, siguiendo el *analogy principle* el estimador estaría dado por

$$\hat{\beta} = \left[\frac{1}{n} \sum_{i=1}^{n} \Pi' z_i' x_i \right]^{-1} \left[\frac{1}{n} \sum_{i=1}^{n} \Pi' z_i' y_i \right] \equiv (\Pi' Z' X)^{-1} (\Pi' Z' y)$$

pero no hemos dicho nada de la matriz Π. Esta puede ser desconocida para lo cual necesitaremos un estimador de Π.

Antes de ir con eso, supongamos que tenemos un estimador de Π dado por $\hat{\Pi}$, luego el estimador *generalizado de variables instrumentales* (GIV) o de método de momentos generalizado (GMM) está dado por (5)

Múltiples instrumentos: 2SLS

$$\hat{\beta}_{GIV} = \left[\frac{1}{n}\sum_{i=1}^{n}\hat{\Pi}' z_i' x_i\right]^{-1}\left[\frac{1}{n}\sum_{i=1}^{n}\hat{\Pi}' z_i' y_i\right] \equiv (\hat{\Pi}' Z' X)^{-1}(\hat{\Pi}' Z' y) \tag{5}$$

La elección clásica de matriz es $\hat{\Pi} = (Z'Z)^{-1}Z'X$ que corresponde al estimador MCO de la regresión de X sobre Z, así el estimador de mínimos cuadrados en dos etapas (2SLS) es

$$\begin{aligned}\hat{\beta}_{2SLS} &= (X'Z(Z'Z)^{-1})Z'X)^{-1}(X'Z(Z'Z)^{-1}Z'y) && (6)\\ &= (\hat{X}'X)^{-1}\hat{X}'y && (7)\end{aligned}$$

Múltiples instrumentos: 2SLS

Note que el nombre de mínimos cuadrados en dos etapas (2SLS) viene de la interpretación de Theil que muestra que el estimador se puede obtener de la siguiente manera

$$\hat{\beta}_{2SLS} = (\hat{X}'\hat{X})^{-1}\hat{X}'y \tag{8}$$

En consecuencia, se puede escribir como un sistema de ecuaciones donde tenemos la primera etapa y luego la ecuación estructural:

$$X = Z\Pi + e \tag{9}$$

$$y = X\beta + u \tag{10}$$

Así, se estima la primera etapa y se ocupan los valores estimados $\hat{X} = Z\hat{\Pi} = Z(Z'Z)^{-1}Z'X$ en la segunda etapa.

Múltiples instrumentos: 2SLS

Normalidad y Consistencia Asintótica

Podemos escribir

$$\hat{\beta}_{GIV} - \beta = (\hat{\Pi}' Z' X)^{-1} \hat{\Pi}' Z' u \tag{11}$$

Si $\frac{1}{\sqrt{N}} Z' u \xrightarrow{d} (0, V_0)$ donde $V_0 = E(u^2 z' z)$ y $\hat{\Pi} \xrightarrow{p} \Pi$ es fácil demostrar que

$$\sqrt{n}(\hat{\beta}_{GIV} - \beta) \xrightarrow{d} \mathcal{N}(0, [\Pi' M_{zx}]^{-1} \Pi' V_0 \Pi [\Pi' M_{zx}^{-1}]'), \tag{12}$$

expresión que depende de Π, V_0 y M_{zx}. Si bien no conocemos Π, podemos estimarla como la proyección ortogonal de X sobre Z y en consecuencia obtendremos la distribución asintótica de el estimador de 2SLS (el cual es factible). Por ley de grandes números sabemos que

$$\hat{\Pi} \xrightarrow{p} \Pi \equiv [E(z' z)]^{-1} E(z' x) = M_{zz}^{-1} M_{zx}$$

Múltiples instrumentos: 2SLS

Normalidad y Consistencia Asintótica

Reemplazando la expresión de Π en (12) la varianza asintótica del estimador de 2SLS está dada por la imponente fórmula:

$$\sqrt{n}(\hat{\beta}_{2SLS} - \beta) \xrightarrow{d} \mathcal{N}(0, [M'_{xz} M_{zz}^{-1} M_{zx}]^{-1} M'_{xz} M_{zz}^{-1} V_0 M_{zz}^{-1} M_{zx} [M'_{xz} M_{zz}^{-1} M_{zx}]') \tag{13}$$

La buena noticia es que si u es independiente de z (como tradicionalmente se asume) y bajo homcedasticidad, la fórmula se simplifica debido a que $V_0 = \sigma^2 M_{zz}$,

$$\sqrt{n}(\hat{\beta}_{2SLS} - \beta) \xrightarrow{d} \mathcal{N}(0, [M'_{xz} M_{zz}^{-1} M_{zx}]^{-1}) \tag{14}$$

Para estimar σ^2 definamos el residuo de 2SLS como $\hat{u} = y - x\hat{\beta}_{2SLS}$. Note que este NO es el residuo de la segunda etapa. Así, el estimador de σ^2 se define de la forma tradicional

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

Si u y z no son independientes (a nivel de distribución, pero $E(z'u) = 0$)) se puede estimar consistentemente la matriz de varianzas y covarianzas usando Eicker-Huber-White o Newey-West dependiendo si los errores son autocorrelacionados.

El estimador de Eicker-Huber-White de la varianza asintótica de $\hat{\beta}_{2SLS}$ está dado por

$$A\hat{var}(\beta_{2SLS}) = (\hat{X}'\hat{X})^{-1}\left(\sum_{i=1}^{n}\hat{u}_i^2 z_i' z_i\right)(\hat{X}'\hat{X})^{-1}$$

donde $\hat{X} = Z\hat{\Pi}$.

En Stata, usando el comando `ivreg` o `ivreg2` con la opción `robust` nos entregará los errores estándar de la matriz recién descrita.

Método Generalizado de Momentos: GMM

Un estimador alternativo a 2SLS en presencia de endogeneidad y variables instrumentales es el estimador de GMM.

Definamos las **condiciones de momento** como

$$m(z, x, \beta) = z'(y - x\beta)$$

donde z es una realización del vector de $L \times 1$ de instrumentos y x es una realización del vector de variables endógenas de $K \times 1$. Suponiendo que $L > K$ tenemos un sistema sobre identificado. Dado el supuesto de identificación $E(z'u) = 0$ tenemos que,

$$E(m(z, x, \beta)) = 0$$

luego el valor esperado de cada condición de momento es cero. Cada condición de momento poblacional tiene su contraparte muestral dada por,

$$\overline{m}(z, x, \beta) = \frac{1}{n} \sum z'(y - x\beta) = \frac{1}{n} Z'u$$

Método Generalizado de Momentos: GMM

Para el caso sobre-identificado el estimador de GMM es aquel que minimiza la siguiente forma cuadrádica,

$$\min_{\beta} n\overline{m}(z,x,\beta)'W^{-1}\overline{m}(z,x,\beta)$$

donde W^{-1} es una matriz de $L \times L$ con lo cual el sistema es de $K \times K$.

Se define el estimador de GMM eficiente o EGMM aquel que utiliza como weighting matrix

$$W = Var(m(z,x,\beta)) = E(u^2 z'z) = V_0$$

Por lo tanto, en el caso que $W = V_0$ tenemos que $\hat{\beta}_{GMM}$ minimiza la siguiente expresión

$$\min_{\beta} \frac{1}{n}(y - X\beta)'ZV_0^{-1}Z'(y - X\beta)$$

Luego,

Método Generalizado de Momentos: GMM

$$\hat{\beta}_{EGMM} = [X'ZV_0^{-1}Z'X]^{-1}X'ZV_0^{-1}Z'y \tag{15}$$

Sólo nos falta un estimador consistente de V_0. Bajo el supuesto de homocedasticidad la varianza asintótica estará dada por

$$Avar(\hat{\beta}_{EGMM}) = (M'_{zx}V_0^{-1}M_{zx})^{-1}$$

Si no estamos seguros del supuesto de homocedasticidad podemos usar el estimador de Eicker-White.

Método Generalizado de Momentos: GMM

Luego, se puede implementar el estimador EGMM en tres etapas:

1. Estime el modelo por 2SLS y obtenga los residuos de la manera antes descrita $\hat{u} = y - x\hat{\beta}_{2SLS}$.
2. Construya la matriz $\hat{V}_0 = \frac{1}{n} \sum \hat{u}_i^2 z_i' z_i$.
3. Estime mediante EGMM usando $\hat{V}_0$ como weighting matrix.

En Stata esto se puede implementar con el comando `ivreg2` con la opción `gmm`.

En el caso general, cuando los errores son heterocedásticos y/o autocorrelacionados y $V_0 \neq \sigma^2 M_{zz}$ el estimador 2SLS (caso particular de GIV) no tendrá la menor varianza asintótica. Para obtener un estimador eficiente necesitamos escoger una matriz Π que minimize la varianza asintótica.

Luego, queremos minimizar con respecto a Π la siguiente expresión,

$$Avar(\hat{\beta}_{GIV}) = [\Pi' M_{zx}]^{-1} \Pi' V_0 \Pi ([\Pi M_{zx}]^{-1})'$$

Se puede demostrar (bastante engorroso) que

$$\Pi^* = V_0^{-1} M_{zx} = \text{argmin}_{\Pi} Avar(\hat{\beta}_{GIV}(\Pi))$$

Pero en la práctica, no ponemos disponer de Π^* incluso si suponemos V_0 conocida. Necesitamos un estimador consistente de M_{zx}. La ley débil de los grandes números nos garantiza que si $\{x_i z_i\}$ son $i.i.d$ con primer y segundo momento acotados,

$$\hat{\Pi}^* = V_0^{-1}\left(\frac{1}{n}\sum z_i' x_i\right) = V_0^{-1}\frac{1}{n}Z'X \xrightarrow{p} V_0^{-1}M_{zx} = \Pi^*$$

por lo tanto, el estimador generalizado de variables instrumentales eficiente corresponde al estimador eficiente de método de momentos y es igual a

$$\hat{\beta}_{EGIV} = \hat{\beta}_{EGMM} = [X'ZV_0^{-1}Z'X]^{-1}X'ZV_0^{-1}Z'y \tag{16}$$

con distribución asintótica

$$\sqrt{n}(\hat{\beta}_{EGMM} - \beta) \xrightarrow{d} \mathcal{N}(0, (M_{zx}'V_0^{-1}M_{zx})^{-1}) \tag{17}$$