Introducción

Endogeneidad y Variables Instrumentales

Múltiples instrumentos: 2SLS

Método Generalizado de Momentos (GMM)

Test de Sobre-identificación

Fallas en la identificación

Instrumentos Débiles

Clase 1: Selección en no obervables I

Tomás Rau Binder

8 de septiembre

Contenidos

1 Introducción

2 Endogeneidad y Variables Instrumentales

3 Múltiples instrumentos: 2SLS

4 Método Generalizado de Momentos (GMM)

5 Test de Sobre-identificación

6 Fallas en la identificación

7 Instrumentos Débiles

Introducción

- Endogeneidad es la contribución fundamental de la Econometría a la Estadística.
- El problema de endogeneidad, correlación entre una o más variables explicativas con los no observables, se le conoce como **selección en no observables** en la literatura de evaluación de programas.
- A diferencia del modelo de Heckman (1979), los individuos pueden (auto) seleccionarse en un programa por razones que no observamos y que están correlacionadas con la variable de resultado.

Introducción

- La presencia de endogeneidad hará que el estimador OLS sea sesgado e inconsistente
- Luego, es fundamental atacar el problema y el método más conocido es el de variables instrumentales (VI)
- La modelación estructural es la alternativa a VI pero siempre se necesitará de una variable excluida o instrumento
- Nosotros veremos el enfoque de VI

Introducción

Considere el siguiente modelo lineal:

$$y = \mathbf{x}\beta + u$$

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

El supuesto de identicación fundamental de Mínimos Cuadrados Ordinarios es

$$E(\mathbf{x}'u) = 0$$

Introducción

Note que el parámetro poblacional β puede ser expresado en momentos de las variables observables explotando el supuesto recién presentado:

$$\mathbf{x}'y = \mathbf{x}'\mathbf{x}\beta + \mathbf{x}'u$$

tomando valor esperado tenemos que:

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

lo que se conoce como *Regresión Poblacional*. Dado que $(y, \mathbf{x})$ es observable, β es *identificado*.

Introducción

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

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

que es básicamente lo propuesto por Gauss (1821).

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:

$$E(\mathbf{x}'u) \neq 0$$

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 $\mathbf{z}$ (en este caso de igual dimensión que $\mathbf{x}$) que cumpla las siguientes condiciones:

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

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

Endogeneidad y Variables Instrumentales

Podemos pre-multiplicar la ecuación estructural por $\mathbf{z}'$ y obtener un sistema de ecuaciones:

$$\mathbf{z}'y = \mathbf{z}'\mathbf{x}\beta + \mathbf{z}'u$$

y por lo tanto obtemos el siguiente sistema de ecuaciones:

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

donde $E(\mathbf{z}'\mathbf{x})$ es de orden $K \times K$ y $E(\mathbf{z}'y)$ es de orden $K \times 1$. Este sistema tiene solución única si la matriz $E(\mathbf{z}'\mathbf{x})$ es invertible, lo cual sucede si el rango de ésta es completo e igual a K.

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

Endogeneidad y Variables Instrumentales

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:

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

Este es el el estimador de VI para el caso **exactamente identificado**

Clase 1: Selección en no obervables I

Tomás Rau Binder

Introducción

Endogeneidad y Variables Instrumentales

Múltiples instrumentos: 2SLS

Método Generalizado de Momentos (GMM)

Test de Sobre-identificación

Fallas en la identificación

Instrumentos Débiles

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

$$\hat{\beta}_{IV} = \beta + \left[\frac{1}{n}\sum_{i=1}^{n} \mathbf{z}_i' \mathbf{x}_i\right]^{-1} \left[\frac{1}{n}\sum_{i=1}^{n} \mathbf{z}_i' u_i\right]$$

Luego, se deduce que $\mathrm{plim}\beta_{IV} = \beta$ por álgebra de plims y la ley débil de los grandes números o *Weak law of large numbers* (WLLN).

Clase 1: Selección en no obervables I

Tomás Rau Binder

Introducción

Endogeneidad y Variables Instrumentales

Múltiples instrumentos: 2SLS

Método Generalizado de Momentos (GMM)

Test de Sobre-identificación

Fallas en la identificación

Instrumentos Débiles

Ahora, podemos generar la expresión típica ajustada por $\sqrt{n}$:

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

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

El segundo término converge en distribución a una Normal por Teorema central del límite (CLT), en particular Lindeberg-Levy.

Clase 1: Selección en no obervables I

Tomás Rau Binder

Introducción

Endogeneidad y Variables Instrumentales

Múltiples instrumentos: 2SLS

Método Generalizado de Momentos (GMM)

Test de Sobre-identificación

Fallas en la identificación

Instrumentos Débiles

Veamos un poco en detalle el segundo término:

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

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

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

Clase 1: Selección en no obervables I
Tomás Rau Binder
Introducción
Endogeneidad y Variables Instrumentales
Múltiples instrumentos: 2SLS
Método Generalizado de Momentos (GMM)
Test de Sobre-identificación
Fallas en la identificación
Instrumentos Débiles

Luego, para obtener un estimador de la varianza debemos usar el análogo muestral de M_{zx}

$$= \hat{M}_{zx} = \frac{1}{n}\sum_{i=1}^{n} \mathbf{z}_i'\mathbf{x}_i = \frac{1}{n}Z'X$$

y un estimador para V_0. Si asumimos $\mathbf{z}$ independiente de u tenemos que $V_0 = \sigma^2 E(\mathbf{z}'\mathbf{z})$ y

$$\hat{V}_0 = \hat{\sigma}^2 \frac{1}{n}\sum_{i=1}^{n} \mathbf{z}_i'\mathbf{z}_i = \frac{1}{n}\hat{\sigma}^2 Z'Z$$

El comando en STATA **ivreg** usa por *default* esta matriz de varianzas y covarianzas.

Endogeneidad y Variables Instrumentales

Instrumentos Binarios: Estimador de Wald

Suponga que tenemos el modelo lineal:

$$y_i = \alpha + \beta x_i + \epsilon_i$$

donde x_i es binaria (que puede ser 1 si es tratado T y cero sino). Si $E(\epsilon_i|x_i) = 0$, se puede demostrar que (hacerlo en la casa)

$$\begin{aligned}\hat{\beta}_{OLS} &= \frac{\sum(y_i - \bar{y})(x_i - \bar{x})}{\sum(x_i - \bar{x})^2} = \frac{\sum x_i y_i}{\sum x_i} - \frac{\sum(1 - x_i)y_i}{\sum(1 - x_i)} \\ &= \hat{E}[y|x = 1] - \hat{E}[y|x = 0]\end{aligned}$$

donde $\hat{E}[y|x = 1] = \frac{\sum_{i=1}^n y_i x_i}{\sum_{i=1}^n x_i}$ y $\hat{E}[y|x = 0] = \frac{\sum_{i=1}^n y_i(1-x_i)}{\sum_{i=1}^n(1-x_i)}$

Endogeneidad y Variables Instrumentales

Instrumento Binarios: Estimador de Wald

Ahora si $E(\epsilon_i|x_i) \neq 0$ pero tenemos un instrumento binario z_i tenemos que

$$\hat{\beta}_{IV} = \frac{\sum(y-\bar{y})(z-\bar{z})}{\sum(x-\bar{x})(z-\bar{z})}$$

un poco de álgebra implica

$$\hat{\beta}_{IV} = \frac{\hat{E}[y|z=1] - \hat{E}[y|z=0]}{\hat{E}[x|z=1] - \hat{E}[x|z=0]}$$

donde $\hat{E}[y|z=1] = \frac{\sum_{i=1}^n y_i z_i}{\sum_{i=1}^n z_i}$ y $\hat{E}[y|z=0] = \frac{\sum_{i=1}^n y_i(1-z_i)}{\sum_{i=1}^n (1-z_i)}$ y $\hat{E}[x|z=1] = \frac{\sum_{i=1}^n x_i z_i}{\sum_{i=1}^n z_i}$ y $\hat{E}[x|z=0] = \frac{\sum_{i=1}^n x_i(1-z_i)}{\sum_{i=1}^n (1-z_i)}$

Ejemplo: La lotería de Vietnam, Angrist (1990)

- Durante la Guerra de Vietnam se realizaron 5 loterías en USA para enviar jóvenes a la guerra.
- Por ejemplo, la lotería de 1970 cubrió a jóvenes entre 19 y 26 años.
- Se sortearon números (sin reemplazo) del 1 al 365. A cada número se le asignó una fecha de nacimiento (día, mes).
- La gente iba siendo llamada de acuerdo a una secuencia de números (de 1 a 365) con un techo (ejemplo 125) hasta cumplir la cuota requerida por el departamento de defensa.
- Después se le hacían exámenes médicos y se quedaba con quienes irían a la guerra

Ejemplo: La lotería de Vietnam, Angrist (1990)

- Angrist (1990) estudia el impacto de servir en Vietnam en el salario.
- Esto es interesante porque existe la hipótesis que servir en vietnam compensa la pérdida de experiencia en el mercado laboral
- El problema de endogeneidad es claro, gente que "sirve" en una guerra voluntariamente tiene otras caractéristicas no-observables que pueden estar correlacionadas con habilidad, capital social, etc.

Ejemplo: La lotería de Vietnam, Angrist (1990)

- Angrist (1990) decide usar el número de la lotería como un instrumento de servir en Vietnam mediante un estimador de Wald.
- Se puede crear un instrumento Binario (1 si tienes un número de sorteo bajo, 0 si es alto)
- Este instrumento está correlacionado con "servir" en la guerra pero no con otras características

así,

$$\hat{\beta}_{IV} = \frac{\hat{E}[y|z=1] - \hat{E}[y|z=0]}{\hat{E}[x|z=1] - \hat{E}[x|z=0]}$$

Ejemplo: La lotería de Vietnam, Angrist (1990)

Table 4.1.3: Wald estimates of the effects of military service on the earnings of white men born in 1950

	Earnings		Veteran Status		Wald Estimate of Veteran Effect
Earnings year	Mean (1)	Eligibility Effect (2)	Mean (3)	Eligibility Effect (4)	(5)
1981	16,461	-435.8	0.267	0.159	-2,741
		(210.5)		(0.040)	(1,324)
1971	3,338	-325.9			-2050
		(46.6)			(293)
1969	2,299	-2.0			
		(34.5)			

Notes: Adapted from Angrist (1990), Tables 2 and 3. Standard errors are shown in parentheses. Earnings data are from Social Security administrative records. Figures are in nominal dollars. Veteran status data are from the Survey of Program Participation. There are about 13,500 individuals in the sample.

Ejemplo: La lotería de Vietnam, Angrist (1990)

- El resultado más importante que encuentran es que para veteranos blancos, el efecto de servir en Vietnam reduce el salario en alrededor de US$2.000-US$2.700
- Luego la sustición servicio-experiencia laboral es parcial
- Interpretación: LATE (Imbens y Rubin, 1996). El efecto encontrado es para la subpoblación inducida por la loteria a servir en Vietnam, no para toda la población (en especial no para los voluntarios ni para aquellos que se casaron o entratona estudiar para no ir a Vietnam)

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 $\mathbf{z}\Pi$ es de $1 \times K$ al igual que $\mathbf{x}$. Así pre-multiplicamos (por la derecha) a ambos lados de la ecuación por $\Pi'\mathbf{z}'$

$$\mathbf{y} = \mathbf{x}\beta + u \quad | \times \Pi'\mathbf{z}'$$

Múltiples instrumentos: 2SLS

Obteniendo

$$\Pi'\mathbf{z}'y = \Pi'\mathbf{z}'\mathbf{x}\beta + \Pi'\mathbf{z}'u$$

Note que podemos identificar β tomando valor esperado a ambos lados explotando la condición de identificación $E(\mathbf{z}'u) = 0$

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

siempre que $[E(\Pi'\mathbf{z}'\mathbf{x})]^{-1}$ exista.

Múltiples instrumentos: 2SLS

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

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

pero Π es desconocida luego este estimador no es factible.

Supongamos que tenemos un estimador de Π dado por $\hat{\Pi}$, luego el estimador *generalizado de variables instrumentales* (GIV) está dado por

Múltiples instrumentos: 2SLS

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

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,

$$\hat{\beta}_{2SLS} = (X'Z(Z'Z)^{-1}Z'X)^{-1}X'Z(Z'Z)^{-1}Z'y \tag{6}$$

$$\tag{7}$$

Múltiples instrumentos: 2SLS

Note que el nombre de mínimos cuadrados en dos etapas (2SLS) viene de la interpretación de Theil (1953) 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.

Clase 1: Selección en no obervables I

Tomás Rau Binder

Introducción

Endogeneidad y Variables Instrumentales

Múltiples instrumentos: 2SLS

Método Generalizado de Momentos (GMM)

Test de Sobre-identificación

Fallas en la identificación

Instrumentos Débiles

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 $V_0 = E(u^2 \mathbf{z}' \mathbf{z})$ 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}.

Múltiples instrumentos: 2SLS

Normalidad y Consistencia Asintótica

Si bien no conocemos Π, podemos estimarla como la proyección ortogonal de X sobre Z

Luego obtendremos la distribución asintótica de el estimador de 2SLS (el cual es factible). Por ley de grandes números sabemos que

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

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:

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

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

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

Este es el *default* de STATA con el comando **ivreg**

Múltiples instrumentos: 2SLS

Normalidad y Consistencia Asintótica

Para estimar σ^2 definamos el residuo de 2SLS como $\hat{u} = y - \mathbf{x}\hat{\beta}_{2SLS}$. 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 $\mathbf{z}$ no son independientes (a nivel de distribución, pero $E(\mathbf{z}'u) = 0$) se puede estimar consistentemente la matriz de varianzas y covarianzas usando **Eicker-Huber-White** o **Newey-West**.

Múltiples instrumentos: 2SLS
Normalidad y Consistencia Asintótica

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

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

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, Hansen (1982).

Definamos las **condiciónes de momento** como

$$\begin{aligned} m(\mathbf{z}, \mathbf{x}, \beta) &= \mathbf{z}' u \\ &= \mathbf{z}'(y - \mathbf{x}\beta) \end{aligned}$$

donde $\mathbf{z}$ es una realización del vector de $L \times 1$ de instrumentos y $\mathbf{x}$ es una realización del vector de variables endógenas de $K \times 1$.

Método Generalizado de Momentos: GMM

Suponiendo que $L > K$ tenemos un sistema sobre identificado. Dado el supuesto de identificación $E(\mathbf{z}'u) = 0$ tenemos que,

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

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

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

Método Generalizado de Momentos: GMM

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

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

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

Método Generalizado de Momentos: GMM

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

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

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

$$\min_{\beta} \frac{1}{n} (y - X\beta)' Z V_0^{-1} Z' (y - X\beta)$$

luego se puede demostrar que

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 $V_0 = \sigma^2 M_{zz}$ la varianza asíntotica es la misma que 2SLS, así

$$\sqrt{n}(\hat{\beta}_{EGMM} - \beta) \xrightarrow{d} \mathcal{N}(0, \sigma^2(M'_{zx}M_{zz}^{-1}M_{zx})^{-1}) \tag{16}$$

Si no estamos seguros del supuesto de homocedasticidad podemos usar el estimador de **Eicker-Huber-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 $\hat{u} = y - \mathbf{x}\hat{\beta}_{2SLS}$.
2. Construya la matriz $\hat{V}_0 = \frac{1}{n}\sum \hat{u}_i^2 \mathbf{z}_i' \mathbf{z}_i$.
3. Estime mediante EGMM usando $\hat{V}_0$ como weighting matrix

En STATA esto se puede implementar con el comando **ivreg2** con la opción **gmm2s**. Si no está seguro de la presencia de heterocedasticidad se recomienda usar EGMM.

Restricciones de Sobre-identificación

- En el caso que tenemos más instrumentos que variables endógenas podemos testear la *validez* de los instrumentos.
- La gracia de tener más instrumentos es que nos ayudan a reducir la varianza de nuestro estimador
- En realidad se puede testear que el modelo esté bien especificado con dichos instrumentos, luego es un test que en el evento que se rechace la nula de correcta especificación no sabremos con seguridad si es por los instrumentos o por el modelo.
- En el caso de 2SLS este test se llama test de Sargan y en el caso de GMM es el test J.

Restricciones de Sobre-identificación

En el contexto de GMM este test se realiza usando el estadístico J de Hansen (1982). Este estadístico no es más que la función objetivo evaluada en $\hat{\beta}_{EGMM}$ y se distribuye como χ^2_{L-k}, así

$$J(\hat{\beta}_{EGMM}) = n\,\overline{m}(\mathbf{z},\mathbf{x},\hat{\beta}_{EGMM})'\hat{V}_0^{-1}\overline{m}(\mathbf{z},\mathbf{x},\hat{\beta}_{EGMM}) \xrightarrow{d} \chi^2_{L-K}$$

claramente un valor "grande" de $J(\hat{\beta}_{EGMM})$, es decir, mayor al valor crítico obtenido de la tabla, nos da indicios que el modelo está mal especificado o que los instrumentos no son limpios.

En Stata el comando **ivreg2** con la opción **robust** estima por EGMM y además entrega el estadístico J.

Restricciones de Sobre-identificación

En el caso de 2SLS existe el test de Sargan (1958) para restricciones de sobre-identificación. Es muy simple,

$$Sargan = \frac{\hat{u}' P_Z \hat{u}}{\hat{u}'\hat{u}/n} \xrightarrow{d} \chi^2_{L-K}$$

donde $P_Z = Z(Z'Z)^{-1}Z'$

El númerador mide cuánto explican **x** y **z** la variación de los errores estimados $\hat{u}$ (y nos gustaría que fuera bajo)

El denominador es sólo un estimador (consistente) de σ^2 o varianza de los errores.

Restricciones de Sobre-identificación

- Una manera sencilla de obtener el estadístico de Sargan es correr el modelo por 2SLS
- obtener $\hat{u}$.
- Correr la regressión auxiliar de $\hat{u}$ sobre todas las variables exógenas ($\mathbf{x}$ y $\mathbf{z}$'s)
- obtener el R^2. Se puede demostrar que $Sargan = n \times R^2$.

- En Stata, el comando **ivreg** con la opción **overid** nos entrega el estadístico de Sargan.

Testeando Exogeneidad Débil

- En el caso de IV o 2SLS existe el test de Hausman (1978).
- Si bien, $H_0 : E(\mathbf{x}'u) = 0$ no puede ser testeado, se puede comparar IV con OLS.
- IV es consistente bajo la hipótesis nula (de endogeneidad) y la hipótesis alternativa (de exogeneidad), no así OLS que no es consistente bajo endogeneidad.
- Luego, la hipótesis nula factible es $H_0 : \hat{\beta}_{IV} - \hat{\beta}_{OLS} = 0$.
- Una manera de hacer un test es crear una forma cuadrádica con $\hat{\beta}_{IV} - \hat{\beta}_{OLS}$ pero necesitamos la varianza de eso para crear el *sandwhich*

Testeando Exogeneidad Débil

Luego,

$$H = (\hat{\beta}_{IV} - \hat{\beta}_{OLS})'[V(\hat{\beta}_{IV} - \hat{\beta}_{OLS})]^{-1}(\hat{\beta}_{IV} - \hat{\beta}_{OLS})$$

Hausman (1978) demuestra que

$$V(\hat{\beta}_{IV} - \hat{\beta}_{OLS}) = V(\hat{\beta}_{IV}) - V(\hat{\beta}_{OLS})$$

y que

$$H \xrightarrow{d} \chi(rank(V(\hat{\beta}_{IV}) - V(\hat{\beta}_{OLS}))$$

y se rechaza la hip. nula de exogeneidad para valores grandes de H de acuerdo a la tabla χ^2.

Fallas en la identificación e identificación débil

Fallas en la identificación

Considere el caso en que $K = L = 1$. Luego, el modelo puede ser escrito de la siguiente manera

$$y_i = x_i\beta + e_i$$
$$x_i = z_i\pi + u_i$$

donde e y u son i.i.d normalmente distribuidos. Si $E(z_i e_i) = 0$,

β está *identificado* si y sólo si $E(z_i x_i) \neq 0$ dado que

$$z_i y_i = z_i x_i \beta + z_i e_i$$
$$E(z_i y_i) = E(z_i x_i)\beta$$

luego

$$\beta = E(z_i y_i)/E(z_i x_i)$$

existe sólo si $E(z_i x_i) \neq 0$.

Fallas en la identificación

- Luego, $E(z_i x_i) \neq 0$ es una condición necesaria para poder "despejar" β.
- Si $E(z_i x_i) = 0$ no podemos escribir β en función de momentos, lo que no significa que β no exista, este existe pero el método no es capaz de identificarlo.
- Note que $E(z_i x_i) = 0$ es equivalente que $\pi = 0$
- Luego, necesitamos que la variable instrumental esté correlacionada con la variable endógena. Este es el supuesto (2) de variables instrumentales que si se puede testear.

Fallas en la identificación

Suponga que esta condición falla, luego $E(z_i x_i) = 0$.
El estimador IV está dado por

$$\begin{aligned} \hat{\beta}_{iv} &= \frac{\sum_i^n z_i y_i}{\sum_i^n z_i x_i} \\ &= \beta + \frac{\sum_i^n z_i e_i}{\sum_i^n z_i x_i} \end{aligned}$$

Por Ley de Grandes Números: $1/n \sum_i^n z_i e_i = 0$ y $1/n \sum_i^n z_i x_i = 0$. Luego Plim $\hat{\beta}_{iv}$ NO EXISTE.

Fallas en la identificación

Más encima, tampoco existe su esperanza puesto que por CLT *Lindeberg-Levy*,

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n} z_i e_i \xrightarrow{d} N_1 \sim \mathcal{N}(0, E(z_i^2 e_i^2))$$

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n} z_i x_i = \frac{1}{\sqrt{n}}\sum_{i=1}^{n} z_i u_i \xrightarrow{d} N_2 \sim \mathcal{N}(0, E(z_i^2 u_i^2))$$

luego

$$\hat{\beta}_{iv} - \beta = \frac{\frac{1}{\sqrt{n}}\sum_{i=1}^{n} z_i e_i}{\frac{1}{\sqrt{n}}\sum_{i=1}^{n} z_i u_i} = \frac{N_1}{N_2} \sim Cauchy$$

Fallas en la identificación

dado que el cuociente de dos variables aleatorias normalmente distribuidas sigue una distribución de Cauchy.

Esto es grave puesto que $(\beta - \hat{\beta})$ no converge a cero y el estimador es inconsistente.

Esto sucede puesto que la distribución de Cauchy no tiene primer momento finito.

Este resultado se puede extender para otros casos y fue examinado por Phillips (1989) y Choi (1992).

Identificación débil

Instrumentos Débiles

- Cuando los instrumentos están débilmente correlacionados con las variables endógenas, el uso de variables instrumentales puede ser "peor que la enfermedad" .
- En presencia de *weak instruments* el estimador por variables instrumentales está sesgado en la misma dirección que el estimador por MCO e incluso puede no ser consistente.
- Por otra parte, en presencia de *weak instruments* los tests tienen una medida o significancia incorrecta y los intervalos de confianza tienen "coverage probabilities" incorrectas.

Instrumentos Débiles

Considere el siguiente *setup* para un regresor endógeno

$$y = X\beta + e \tag{17}$$

$$X = Z\Pi + u \tag{18}$$

donde, y es un vector de $n \times 1$, X is una matriz de $n \times 1$, Z is a una matriz de $n \times l$ y e, u son vectores de errores de $n \times 1$ con varianzas σ_e^2 y σ_u^2 respectivamente y con coeficiente de correlación iguales a ρ.

Una medida de "fortaleza" de los instrumentos está dada por el parámetro de concentración (*signal to noise ratio*)

$$\mu^2 = \Pi' Z' Z \Pi / \sigma_u^2 \tag{19}$$

Instrumentos Débiles

- Este parámetro está muy relacionado con el estadístico F de la primera etapa para testear la hipótesis de irrelevancia $\Pi = 0$.
- Se puede demostrar que valores grandes de μ^2/k mueven la distribución del estadístico F y $F - 1$ es un estimador de μ^2/k.
- Rothenberg(1984) muestra que a medida que μ^2 crece, el estimador 2SLS converge en probabilidad y su distribución es estándar.

Instrumentos Débiles

- El siguiente ejercicio de **Monte Carlo** considerado por Staigner y Stock (1997), y Nelson y Starz (1990) , y Stock, Wright y Yogo (2002) nos puede ayudar a ver cuán sensible es la distribución al parámetro de concentración.
- Un ejercicio de **Monte Carlo** se parece a un bootstrap, con la diferencia que no se remuestrea de una muestra existente sino que se inventan nuevas muestras a partir de un *diseño* estadístico.

Instrumentos Débiles

El modelo es

$$y = \beta x + e$$
$$x = \pi z + v$$

- El verdadero valor de $\beta = 0$, los errores i.i.d normalmente distribuidos con medias iguales a cero, varianzas iguales a uno y covarianza $\rho = 0{,}99$.
- Para cada muestra generada fijamos $n = 20$

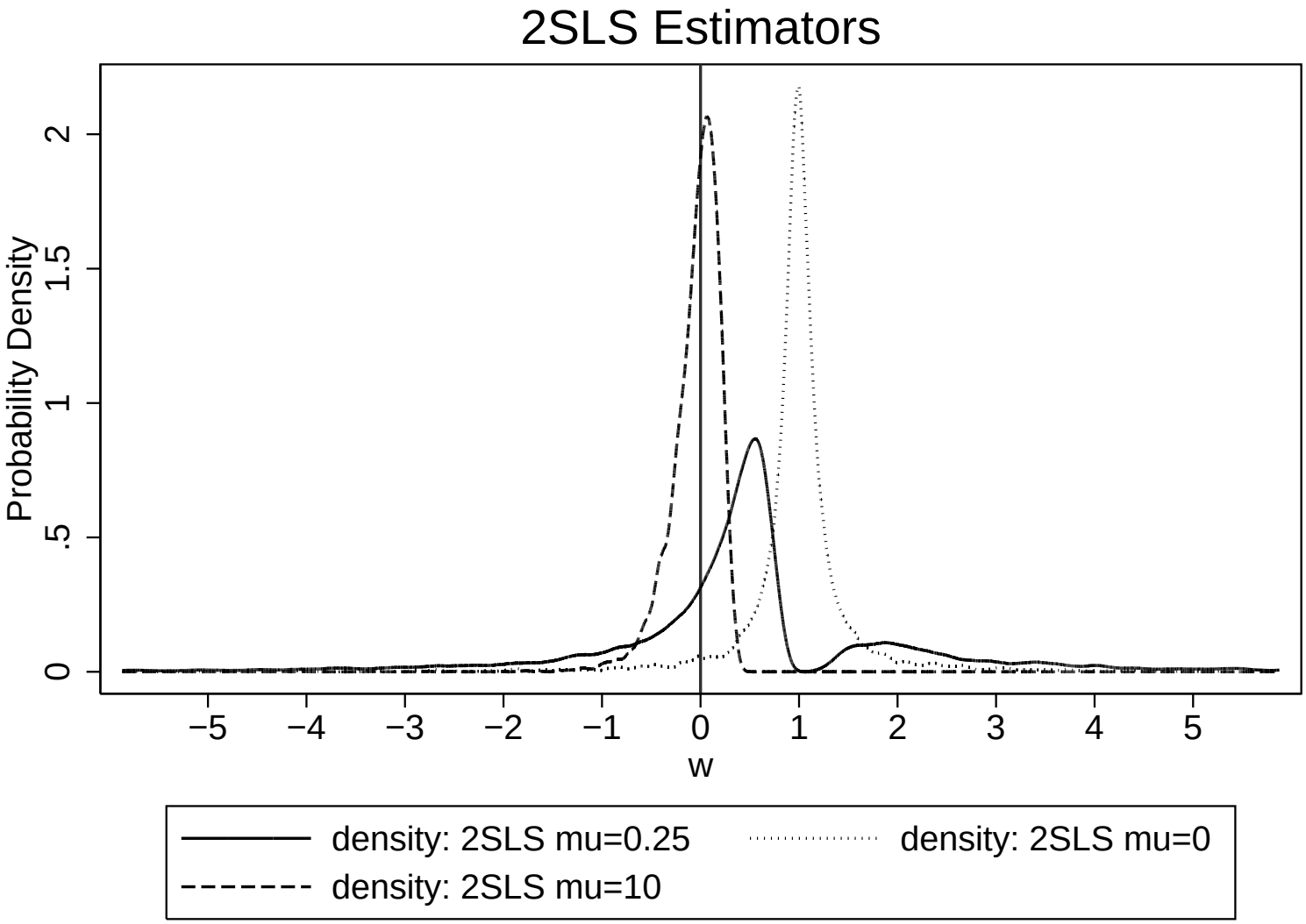
Instrumentos Débiles

- Se fija un valor para μ^2 y se generan números aleatorios para **z**
- Con ellos se obtiene un valor para π

$$\pi^2 = \mu^2 / \sum_{i=1}^{n} z_i^2 \tag{20}$$

  y se construye **x** y luego **y** generando (e,v) de una normal bivariada
- Se estima β mediante 2SLS, esto se hace 10,000 veces
- Para distintos valores del parámetro de concentración graficamos la distribución de $\hat{\beta}_{2SLS}$.

Instrumentos Débiles



Instrumentos Débiles

- Podemos ver que cuando $\mu^2 = 10$ la distribución se centra en el verdadero valor del parámetro
- Sin embargo, cuando $\mu^{=}$0,25 existe un sesgo asintótico y una extraña *bimodalidad*
- Este se ve exacerbado cuando $\mu^2 = 0$
- Si bien no se aprecia mucho, las colas de la última distribución son *pesadas* por eso no tiene valor esperado