# Variables Instrumentales

Walter Sosa-Escudero

March 28, 2019

# Recordatorio

$$P = X(X'X)^{-1}X', \quad M = I - P$$

- $\hat{Y} = X\hat{\beta} = X(X'X)^{-1}X'Y = PY$
- $P$ simetrica e idempotente

# Metodo de Momentos

$$X, \quad E(X) = \mu_0$$

$$\underbrace{E(X-\mu)=0,}_{A} \quad \text{sii } \mu = \mu_0$$

$$X_i, i = 1, \ldots, n, \quad \text{iid }, E(X_i) = \mu_0$$

$$\underbrace{\frac{1}{n}\sum_{i=1}^{n}(X_i - \hat{\mu}) = 0}_{B}$$

Idea: si $B \xrightarrow{p} A$ entonces $\hat{\mu} \xrightarrow{p} \mu_0$

Es facil chequear que $\hat{\mu} = \bar{X}$

En terminos generales, si para una funcion $m(.)$

$$E\left[m(X;\theta)\right] = 0, \ \text{sii } \theta = \theta_0 \tag{1}$$

El estimador de momentos de $\theta_0$ es $\hat{\theta}$ tal que

$$\frac{1}{n}\sum_{i=1}^{n} m(X_i;\hat{\theta}) = 0, \tag{2}$$

para una muestra $X_i$ iid de $X$

Terminologia: (1) es la *condicion de momentos* y (2) es el *analogo muestral*.

# Ejemplo: MCO via MM

$$y_i = x_i'\beta_0 + u_i, \qquad\qquad i = 1, \ldots, n$$

Condicion de momentos: $E(x_i u_i) = 0$, o:

$$E\big(x_i(y_i - x_i'\beta)\big) = 0, \text{ sii } \beta = \beta_0$$

El estimador MM es $\hat{\beta}$ tal que

$$\frac{1}{n}\sum_{i=1}^{n} x_i(y_i - x_i'\hat{\beta}) = 0$$

Despejando

$$\hat{\beta} = \Big(\sum x_i x_i\Big)^{-1} \sum x_i y_i = (X'X)^{-1}X'Y$$

# Comentarios

- MM define al estimador en forma implicita.
- Propiedades: asintoticas ($m(.)$ demasiado general).
- Consistencia y normalidad asintotica: condiciones sobre $m(.)$
- Detalles: van der Vaart (1998, Cap. 4)

# Endogeneidades

- Consistencia de MCO depends de $E(u_i x_i) = 0$.
- En general, cuando esta condicion no vale diremos que $x_i$ es endogena.

# Ejemplos

**Ecuaciones simultaneas**

Oferta y demanda:

$$\begin{cases} q_i^s & = & x_i^s \beta_1^s + \beta_2^s p_i + \epsilon_i^s \\ q_i^d & = & x_i^d \beta_1^d + \beta_2^d p_i + \epsilon_i^d \\ q_i^s & = & q_i^d \end{cases}$$

En equilibrio:

$$p_i = (\beta_2^s - \beta_2^d)^{-1}(x_i^d \beta_1^d - x_i^s \beta_1^s + \epsilon_i^d - \epsilon_i^s)$$

*En la oferta y en la demanda, $p_i$ depende del termino de error. MCO no es consistente*

## Variables omitidas

$$Y = \beta_1 X_1 + \beta_2 X_2 + u$$

si omitimos' $X_2$

$$Y = \beta_1 X_1 + \nu$$

$\nu \equiv \beta_2 X_2 + u$. $X_2$ endogena a menos que $X_1$ y $X_2$ sean ortogonales o $\beta_2 = 0$.

## Variable explicativa medida con error

$$C_i = \beta_1 + \beta_2 X_i^* + u_i$$

valen todas las condiciones para consistencia. Supongamos que solo se observa una version 'ruidosa' de $X_i$:

$$X_i = X_i^* + \omega_i$$

$\omega_i$ es un error de medicion. Supondremos $\omega_i$ iid, $E(\omega_i) = 0$, $V(\omega_i) = \sigma_\omega^2$ y no correlacionado con $X_i^*$ y $u_i$.
Reemplazando $X_i^* = X_i - \omega_i$

$$C_i = \beta_1 + \beta_2 X_i + \nu_i$$

con $\nu = -\beta_2 \omega_i + u_i$

Pero....

$$\begin{aligned} Cov(X_i, \nu_i) = E(X_i, \nu_i) &= E\left[\left(X_i^* + \omega_i\right)\left(-\beta_2 \omega_i + u_i\right)\right] \\ &= -\beta_2 \sigma_\omega^2 \quad \neq 0 \end{aligned}$$

MCO que regresa $Y_i$ en $X_i$ es inconsistente.

# IV bajo identificacion exacta

Modelo. $z_i$ es un vector de $K$ variables instrumentales.

1. Linealidad: $y_i = x_i'\beta_0 + u_i \qquad i = 1, \ldots, n.$
2. Muestra aleatoria: $\{x_i, z_i, u_i\}$ es iid.
3. Validez de VI 1): rango. $E(z_i x_i') = \Sigma_{zx}$ invertible.
4. Validez de VI II): ortogonalidad. $E(z_{ik} u_i) = 0$ para todo $i$ y $k = 1, \ldots, K$.
5. $V(z_i u_i) = S$ finita y positiva definida.

Permite $x_i$ endogena.

# Estimador de VI via MM

Condicion de momentos (ortogonalidad)

$$E(z_i u_i) = E\left[z_i(y_i - x_i'\beta)\right] = 0 \text{ sii } \beta = \beta_0$$

Son $K$ condiciones para $K$ parametros (identificacion exacta).

El estimador MM resuelve:

$$\frac{1}{n}\sum_{i=1}^{n} z_i(y_i - x_i'\hat{\beta}_{IV}) = 0$$

Despejando:

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

- Existencia: asintotica, por la condicion de rango.
- Identificacion exacta: mismo numero de parametros que de instrumentos.
- Si $x_i$ exogena, instrumento valida para si misma ($Z = X$). Entonces $\hat{\beta}_{IV} = \hat{\beta}_{MCO}$.

# Propiedades asintoticas

**Consistencia:** $\hat{\beta}_{IV} \xrightarrow{p} \beta_0$

*Prueba (bosquejo):* $\hat{\beta}_{IV} = (Z'X)^{-1}Z'Y$. Replacing:

$$\begin{array}{rcll} \hat{\beta}_{IV} & = & \beta_0 + & (Z'X)^{-1} \quad Z'u \\ & = & \beta_0 + & \left(\frac{Z'X}{n}\right)^{-1} \quad \left(\frac{Z'u}{n}\right) \\ & & & \xrightarrow{p} \Sigma_{zx} < \infty \quad \xrightarrow{p} 0 \end{array}$$

usando las condiciones de rango y ortogonalidad.

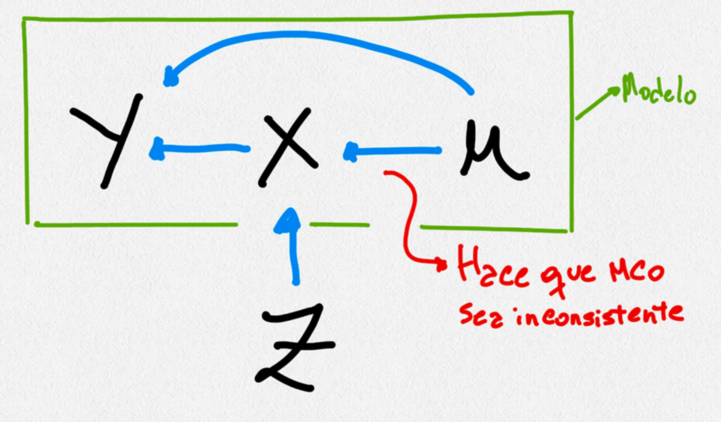
**Normalidad asintotica:** $\sqrt{n}(\hat{\beta}_{IV} - \beta_0) \xrightarrow{d} N(0, \Sigma_{zx}'^{-1} S \Sigma_{zx}'^{-1})$.

*Es un buen ejercicio hacer las pruebas. Ver las que puse para MCO*

# Fuente de instrumentos validos

**Validez de instrumentos**

- Validez de VI 1): rango. $E(z_i x_i') = \Sigma_{zx}$, finita e invertible. Instrumentos no correlacionados con las variables a instrumentar. Se puede testear empiricamente.
- Validez de VI 2): ortogonalidad. $E(z_{ik} u_i) = 0$. Instrumentos no correlacionados con no observables. Depende de inobservables y de como esta armado el modelo.




Instrumento valido: 'afecta a $Y$ solo a traves de $X$'.

Origen de instrumentos validos?

- Dificil. Pregunta estadistica y modelistica.
- Angrist y Krueger (1991): Salarios como funcion de la educacion. Educacion endogena. Instrumento: mes de nacimiento! Validez?
- Growth regression: que es verdaderamente endogeno para el crecimiento?(Durlauf, 2001).

# VI: caso sobreidentificado

$p > K$ instrumentos validos. La logica de MM implica que

$$\frac{1}{n}\sum_{i=1}^{n} z_i(y_i - x_i'\hat{\beta}_{IV}) = 0$$

es un sistema de $p$ ecuaciones lineales con $K$ incognitas.

- Si solo nos preocupase la consistencia, hay algo obvio por hacer.
- Menos obvio?

Idea: estrategia consistente y tal vez mas eficiente.

# Modelo

$z_i^0$ es un vector de $p > K$ instrumentos.

1. Linealidad: $y_i = x_i'\beta_0 + u_i \qquad i = 1, \ldots, n.$
2. Muestra aleatoria: $\{x_i, z_i^0, u_i\}$ conjuntamete iid.
3. Validez de VI 1): rango. $E(z_i^0 x_i') = \Sigma_{z^0x}$ es una matrix $p \times K$ finita y con rango columna completo.
4. Validez de VI 2): ortogonalidad. $E(z_i^0 u_i) = 0$ para todo $i$.
5. $V(z_i^0 u_i) = \sigma^2 E(z_i^0 z_i^{0'})$ finita y positiva definida.

Cuidado, hemos supuesto homocedasticidad.

## Variables instrumentales: $p > K$ instrumentos

$z_i^0$ vector de $p > K$ instrumentos.

- Podriamos descartar $p - K$ instrumentos. Consistente.
- Cualquier combinacion lineal de VI's es un instrumento valido: combinar los $p$ instrumentos para obtener $K$:

$$Z = Z^0 A$$

con $Z_{n\times K}, Z^0_{n\times p}, A_{p\times K}$. El estimador seria:

$$\hat{\beta}_{IV} = (Z'X)^{-1} Z'Y$$

tal como en el caso anterior.

*Problema*: ¿como elegir $A$?

*Resultado:* La combinacion lineal de VI's que tiene varianza asintotica minima corresponde a:

$$A = (Z^{0\prime} Z^0)^{-1} Z^{0\prime} X$$

VI optimo:

$$\hat{\beta}_{IV} = (Z'X)^{-1} Z'X$$

con $Z = Z^0 A = Z^0 (Z^{0\prime} Z^0)^{-1} Z^{0\prime} X = P_{Z^0} X$.

Reemplazando, y dado que $P_{Z^0}$ es simetrica:

$$\begin{aligned} \hat{\beta}_{IV} &= (Z'X)^{-1} Z'X \\ &= (X' P_{Z^0} X)^{-1} X' P_{Z^0} Y \end{aligned}$$

Dado que $P_{Z^0}$ es idempotente y simetrica:

$$\begin{aligned}\hat{\beta}_{IV} &= (X'P_{Z^0}X)^{-1}X'P_{Z^0}Y \\ &= (X'P'_{Z^0}P_{Z^0}X)^{-1}X'P'_{Z^0}P_{Z^0}Y \\ &= (X^{*\prime}X^*)^{-1}X^{*\prime}Y^*\end{aligned}$$

con $X^* \equiv P_{Z^0}X$ y $Y^* \equiv P_{Z^0}Y$.

$$\hat{\beta}_{IV} = (X^{*\prime}X^*)^{-1}X^{*\prime}Y^*, \qquad X^* \equiv P_{Z^0}X, \;\; Y^* \equiv P_{Z^0}Y$$

VI optimo:

1. Obtener $X^*$ y $Y^*$.
2. MCO con $X^*$ y $Y^*$.

VI optimo: *minimos cuadrados en dos etapas*.

Intuicion: MC2E reemplaza las variables por las mejores predicciones en base a $Z$.

# Variables instrumentales: cuestiones operativas

- Si $E(x_i u_i) = 0$, trivialmente $z_i = x_i$ y $\hat{\beta}_{VI} = \hat{\beta}_{MCO}$.
- $K$ variables explicativas, solo 1 de ellas endogena: $x_i^1$ exogena, $x_i^2$ endogena. $z_i$ puede incluir $x_i^1$. Falta al menos 1 instrumento. Trivialmente, en la primera etapa, $x_i^{1*} = x_i^1$ (porque?).

# Variables instrumentales: resultados de muestra finita

La propiedad clave del metodo de VI es *consistencia*. ¿Que sucede en muestras finitas?

- VI es *sesgado* (Sawa, 1962).
- El sesgo se aproxima al de MCO cuando el $R^2$ entre los instrumentos y las variables endogenas tiende a cero (Bound, Jaeger y Baker, 1995).
- Si la correlacion entre los instrumentos y las variables endogenas es baja, una pequeña correlacion entre los instrumentos y el error puede causar una gran inconsistencia en VI (BJB, 1995).
- En muestras grandes, $p$ grande es mejor (porque?), pero en muestras chicas, $p$ grande sesga a VI en la direccion de MCO.

# Test de Hausman

Es un test de 'endogenidad' (cuidado, no es tan asi)

- Bajo $H_0$, $\hat{\beta}_{MCO}$ es consistente y eficiente.
- Bajo $H_A$, $\hat{\beta}_{MCO}$ es inconsistente.
- $\hat{\beta}_{IV}$ es siempre consistente.

$$H = (\hat{\beta}_{MCO} - \hat{\beta}_{IV})' \left( V(\hat{\beta}_{IV}) - V(\hat{\beta}_{MCO}) \right)^{-1} (\hat{\beta}_{MCO} - \hat{\beta}_{IV})$$

tiene distribucion asintotica $\chi^2(K)$ bajo $H_0$.
Caso particular: un solo coeficiente $\beta_1$

$$H = \frac{(\hat{\beta}_{1,MCO} - \hat{\beta}_{1,IV})^2}{(V(\hat{\beta}_{1,IV}) - V(\hat{\beta}_{1,MCO}))^{1/2}} \sim N(0,1)$$

# Test de Restricciones de Sobreidentificacion

Si $p > K$ hay mas instrumentos que los necesarios.
'Validez':

- $Z$ no esta correlacionado con $u$
- Especificacion correcta: los $Z$ que no pertenecen a la regresion efectivamente no pertenecen.

Test: $nR^2 \sim \chi^2(p-k)$
$R^2$, de regresar: $\hat{r} = Z\pi +$ residuo
$\hat{r} \equiv y - x'\hat{\beta}_{VI}$
$R^2 \neq 0$ : instrumentos correlacionados con el error o modelo mal especificado.

# Teoria asintotica para estimadores VI

## Consistencia

$$
\begin{aligned}
\hat{\beta}_{VI} &= (X'P_{Z^0}X)^{-1}X'P_{Z^0}Y \\
&= (X'P_{Z^0}X)^{-1}X'P_{Z^0}(X'\beta_0 + u) \\
&= \beta_0 + (X'P_{Z^0}X)^{-1}X'P_{Z^0}u \\
&= \beta_0 + (X'Z(Z'Z)^{-1}Z'X)^{-1}X'Z(Z'Z)^{-1}Z'u \\
&= \beta_0 + \left(\frac{X'Z}{n}\left(\frac{Z'Z}{n}\right)^{-1}\frac{Z'X}{n}\right)^{-1}\frac{X'Z}{n}\left(\frac{Z'Z}{n}\right)^{-1}\frac{Z'u}{n}
\end{aligned}
$$

- A partir de aqui los dejo solos. Usar los supuestos y la estrategia que usamos en la clase anterior.

## Normalidad Asintotica

Del resultado anterior:

$$\sqrt{n}\left(\hat{\beta}_{VI}-\beta_0\right) = \left(\frac{X'Z}{n}\left(\frac{Z'Z}{n}\right)^{-1}\frac{Z'X}{n}\right)^{-1}\frac{X'Z}{n}\left(\frac{Z'Z}{n}\right)^{-1}\left[\sqrt{n}\frac{Z'u}{n}\right]$$

- A partir de aqui, tal como lo hicimos para MCO, hay que establecer normalidad asintotica de $\sqrt{n}\frac{Z'u}{n}$ y utilizar el teorema de Slutzky, se los dejo como ejercicio.

El resultado es:

$$\sqrt{n}\left(\hat{\beta}_{VI}-\beta_0\right) \xrightarrow{d} N(0,V)$$

con

$$V = \sigma^2\left[E(xz')E(zz')^{-1}E(zx')\right]^{-1}$$