Instrumental Variables

Walter Sosa-Escudero

Econ 507. Econometric Analysis. Spring 2009

July 2, 2009

Motivation

- Consistency of OLS crucially depends on $E(u_i x_i) = 0$.
- In general, when this assumption does not hold x_i will be endogenous

Endogeneities: Examples

Simultaneous Equations

The simplest supply and demand system:

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

In equilibrium:

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

In both, supply and demand, p_i as an explanatory variable depends on the error term. Simple OLS is not consistent

Omitted Variables

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

if we ‘omit’ X_2

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

$\nu \equiv \beta_1 X_2$. Then the predeterminedness assumption will be violated unless X_1 and X_2 are orthogonal or $\beta_2 = 0$.

Explanatory variable measured with error

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

and assume all assumptions for consistency hold. Now suppose we observe a 'noisy' version of X_i:

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

ω_i is a measurement error. We will assume ω_i is iid, with $E(\omega_i) = 0$, $V(\omega_i) = \sigma^2_\omega$ and uncorrelated with X_i^* and u_i. Replacing $X_i^* = X_i - \omega_i$, the regression model can be written as

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

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

Now

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

Then the OLS estimator that regresses Y_i on X_i is inconsistent.

We will derive more details in the next homework.

IV under Exact Identification

Model as before. z_i is a vector of K instrumental variables.

1. Linearity: $y_i = x_i'\beta_0 + u_i \qquad i = 1, \ldots, n.$
2. Random sample: $\{x_i, z_i, u_i\}$ is a jointly i.i.d. process.
3. IV validity 1): rank. $E(z_i x_i') = \Sigma_{zx}$ is a finite, invertible matrix
4. IV validity 2): orthogonality. $E(z_{ik} u_i) = 0$ for all i and $k = 1, \ldots, K$.
5. $V(z_i u_i) = S$ finite positive definite.

The Method-of-Moments and the IV estimator.

The orthogonality condition

$$E(z_i u_i) = E\left[z_i(y_i - x_i'\beta_0)\right] = 0$$

is a set of K moment conditions for the K unknown parameters.

Method-of-moments: choose as estimator the values of the parameters that force the sample moments to hold. $\hat{\beta}_{IV}$ satisfies:

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

This is a system of K linear equations with K unknowns with solution

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

- Existence: asympotically guaranteed by the rank condition.
- Exact identification: same number of variables and instruments (K).
- If x_i is exogenous, it is a valid vector of instruments for itself. Then $\hat{\beta}_{IV} = \hat{\beta}_{OLS}$.

Large sample properties

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

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

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

using the rank and ortogonality conditions.

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

Finish both proofs as a useful homework

Sources of valid instruments

Instrument Validity

- IV validity 1): rank. $E(z_i x_i') = \Sigma_{zx}$, finite and invertible. Intuitively this requires the instruments to be correlated with the variables to be instrumented. This might be checked empirically. More later.
- IV validity 2): orthogonality. $E(z_{ik} u_i) = 0$. Instruments must be uncorrelated with whatever is not observed that is a determinant of y_i. This depends on things we do not observe and on how we setup the model.

Where do valid instruments came from?

- Complex question. It is an econometric and a modeling issue.
- Instrument for prices in the demand function: price determinants related to the supply side (costs).
- Angrist and Krueger (1991): wages as a function of education. Education endogenous. Instrument: month of birth!: individuals born in the first months of the year may abandon school earlier, then they should have *less* education than the rest. Validity?
- Growth regression: What is truly exogenous for growth?(Durlauf, 2001).

IV: the Overidentified Case

Suppose now that there are $p > K$ instruments. Then the MM logic implies that

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

is a system of p linear equations with K unknowns.

- If we only care about consistency there is something obvious we can do.
- Less obvious?
- We will explore a consistent and hopefully more efficient strategy.

Model as before. z_i^0 is a vector of $p > K$ instruments.

1. Linearity: $y_i = x_i'\beta_0 + u_i \qquad i = 1, \ldots, n.$
2. Random sample: $\{x_i, z_i^0, u_i\}$ is a jointly i.i.d. process.
3. IV validity 1): rank. $E(z_i^0 x_i') = \Sigma_{z^0x}$ is a $p \times K$ matrix, that exists, is finite, an has full column rank.
4. IV validity 2): orthogonality. $E(z_i^0 u_i) = 0$ for all i.
5. $V(z_i^0 u_i) = \sigma^2 E(z_i^0 z_i^{0'})$ finite positive definite.

Variables instrumentales: $p > K$ instrumentos

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

- Podria descartar $p - K$ instrumentos. Voy a obtener un estimador consistente.
- Es facil probar que *cualquier combinacion lineal de VI's* es un instrumento valido: podria combinar los p instrumentos de alguna manera, 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:

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

Entonces, el estimador de VI optimo es:

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

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 $Y^* \equiv P_{Z^0}Y$.
Entonces: el estimador de VI optimo para el caso de $p > K$ instrumentos se puede computar en dos etapas.

1. Obtener X^* y Y^*. Pregunta: ¿que es intuitivamente esto?
2. Correr MCO con X^* y Y^*.

A este estimador de VI optimo se lo llama estimador de *minimos cuadrados en dos etapas*.

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 β_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 MV

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

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