

Clase 10: Mapa de ruta

Econometría

Dr. Christian M.
García-Witulski

MC2E con una X
y múltiples Z's

MC2E con
Múltiples
Regresores
Exógenos

MC2E: Caso
General

Supuestos de VI

Propiedades de
MC2E

Tests sobre la
Validez de los
Instrumentos

Test sobre la
Relevancia de los
Instrumentos

Test sobre la
Exogeneidad de
los Instrumentos

Test sobre la
Endogeneidad de
un Regresor

- 1 MC2E con una X y múltiples Z's
- 2 MC2E con Múltiples Regresores Exógenos
- 3 MC2E: Caso General
- 4 Supuestos de VI
- 5 Propiedades de MC2E
- 6 Tests sobre la Validez de los Instrumentos
- 7 Test sobre la Relevancia de los Instrumentos
- 8 Test sobre la Exogeneidad de los Instrumentos
- 9 Test sobre la Endogeneidad de un Regresor

MC2E con una X y una Z : Repaso

Econometría

Dr. Christian M.
García-Witulski

MC2E con una X
y múltiples Z 's

MC2E con
Múltiples
Regresores
Exógenos

MC2E: Caso
General

Supuestos de VI

Propiedades de
MC2E

Tests sobre la
Validez de los
Instrumentos

Test sobre la
Relevancia de los
Instrumentos

Test sobre la
Exogeneidad de
los Instrumentos

Test sobre la
Endogeneidad de
un Regresor

- Consideremos el modelo

$$Y_i = \beta_0 + \beta_1 X_i + u_i$$

con $\mathbb{E}(u_i | X_i) \neq 0$ y un único instrumento, Z .

- **Mínimos Cuadrados en Dos Etapas (MC2E)**

- **Primera etapa:** usar MCO para estimar

$$X_i = \pi_0 + \pi_1 Z_i + v_i \quad \Rightarrow \quad \hat{X}_i = \hat{\pi}_0 + \hat{\pi}_1 Z_i$$

- **Segunda Etapa:** usar MCO para estimar

$$Y_i = \beta_0 + \beta_1 \hat{X}_i + e_i$$

- Obtenemos:

$$\hat{\beta}_1^{\text{MC2E}} = \frac{\sum_{i=1}^n (\hat{X}_i - \bar{\hat{X}})(Y_i - \bar{Y})}{\sum_{i=1}^n (\hat{X}_i - \bar{\hat{X}})^2} = \frac{s_{\hat{X}Y}}{s_{\hat{X}}^2}$$

que es consistente y asintóticamente normal (bajo las hipótesis de IV introducidas previamente).

- Como ya explicamos, el error estándar debe ser computado con la fórmula apropiada.

MC2E con una X y múltiples Z 's

Econometría

Dr. Christian M.
García-Witulski

MC2E con una X y múltiples Z 's

MC2E con
Múltiples
Regresores
Exógenos

MC2E: Caso
General

Supuestos de VI

Propiedades de
MC2E

Tests sobre la
Validez de los
Instrumentos

Test sobre la
Relevancia de los
Instrumentos

Test sobre la
Exogeneidad de
los Instrumentos

Test sobre la
Endogeneidad de
un Regresor

- Hasta ahora hemos asumido que disponemos de una única variable instrumental. ¿Qué ocurre si tenemos más?
- Supongamos que tenemos dos, Z_1 y Z_2 . ¿Qué hacemos?
 - Si aplicamos nuestro procedimiento previo obtendremos dos estimaciones distintas de β_1 !
 - ¿Descartamos uno de los instrumentos?
- La intuición nos dice que sería bueno aprovechar la información contenida en los dos instrumentos (siempre que sean válidos).
 - ¿Cómo definimos la validez cuando hay más de un instrumento?
- Hay más de una manera (razonable) de combinar múltiples instrumentos. Luego, hay múltiples estimadores VI.
- Nosotros nos vamos a concentrar en MC2E, apropiadamente extendido para el caso de múltiples instrumentos y regresores.
 - Con una X y una Z teníamos $\hat{\beta}_1^{\text{MC2E}} = \hat{\beta}_1^{\text{vi}}$. Ahora escribiremos simplemente $\hat{\beta}_1^{\text{MC2E}}$ ya que hay otros estimadores que utilizan instrumentos y que darán estimaciones distintas (como LIML o GMM).

MC2E con una X y múltiples Z 's

Econometría

Dr. Christian M.
García-Witulski

MC2E con una X y múltiples Z 's

MC2E con
Múltiples
Regresores
Exógenos

MC2E: Caso
General

Supuestos de VI

Propiedades de
MC2E

Tests sobre la
Validez de los
Instrumentos

Test sobre la
Relevancia de los
Instrumentos

Test sobre la
Exogeneidad de
los Instrumentos

Test sobre la
Endogeneidad de
un Regresor

- La extensión de MC2E al caso de una X y múltiples instrumentos es sencilla.
- Consideremos el model

$$Y_i = \beta_0 + \beta_1 X_i + u_i \quad \text{con} \quad \mathbb{E}(u_i | X_i) \neq 0$$

y supongamos que tenemos m instrumentos: Z_1, Z_2, \dots, Z_m .

- **Primera Etapa:** usar MCO para estimar

$$X_i = \pi_0 + \pi_1 Z_{1i} + \pi_2 Z_{2i} + \dots + \pi_m Z_{mi} + v_i$$

donde $Cov(Z_{li}, v_i) = 0 \forall l$, y computar

$$\hat{X}_i = \hat{\pi}_0 + \hat{\pi}_1 Z_{1i} + \hat{\pi}_2 Z_{2i} + \dots + \hat{\pi}_m Z_{mi}$$

- **Segunda Etapa:** usar MCO para estimar

$$Y_i = \beta_0 + \beta_1 \hat{X}_i + e_i$$

Luego:

$$\hat{\beta}_1^{\text{MC2E}} = \frac{s_{\hat{X}Y}}{s_{\hat{X}}^2}$$

MC2E con una X y múltiples Z 's

Econometría

Dr. Christian M.
García-Witulski

MC2E con una X
y múltiples Z 's

MC2E con
Múltiples
Regresores
Exógenos

MC2E: Caso
General

Supuestos de VI

Propiedades de
MC2E

Tests sobre la
Validez de los
Instrumentos

Test sobre la
Relevancia de los
Instrumentos

Test sobre la
Exogeneidad de
los Instrumentos

Test sobre la
Endogeneidad de
un Regresor

- En este caso:
 - La **relevancia** de los instrumentos requiere que al menos uno de los coeficientes de las Z 's en la regresión poblacional de la primera etapa ($\pi_1, \pi_2 \dots \pi_m$) sea distinto de cero.
 - La **exogeneidad** de los instrumentos requiere que $Cov(Z_{1i}, u_i) = 0, Cov(Z_{2i}, u_i) = 0, \dots, Cov(Z_{mi}, u_i) = 0$.
- La lógica es similar a la anterior:
 - La primera etapa descompone a X_i en dos partes: $X_i = \hat{X}_i + \hat{v}_i$.
 - Se cumple $Cov(\hat{X}_i, \hat{v}_i) = 0$ (propiedad de MCO) y $Cov(\hat{X}_i, u_i) = 0$ (ya que \hat{X}_i es una combinación lineal de las Z 's, que son exógenas).
 - Sustituyendo $X_i = \hat{X}_i + \hat{v}_i$ en $Y_i = \beta_0 + \beta_1 X_i + u_i$ obtenemos $Y_i = \beta_0 + \beta_1 \hat{X}_i + e_i$, donde $e_i = \beta_1 \hat{v}_i + u_i$, y $Cov(\hat{X}_i, e_i) = 0$.
 - $Cov(\hat{X}_i, e_i) = 0$ se sigue de $e_i = \beta_1 \hat{v}_i + u_i$ y $Cov(\hat{X}_i, \hat{v}_i) = Cov(\hat{X}_i, u_i) = 0$.
 - Luego, la estimación de $Y_i = \beta_0 + \beta_1 \hat{X}_i + e_i$ por MCO producirá un estimador consistente de β_1 (ver el Apéndice I para un argumento formal).

MC2E con Múltiples Regresores Exógenos

Econometría

Dr. Christian M.
García-Witulski

MC2E con una X
y múltiples Z 's

MC2E con
Múltiples
Regresores
Exógenos

MC2E: Caso
General

Supuestos de VI

Propiedades de
MC2E

Tests sobre la
Validez de los
Instrumentos

Test sobre la
Relevancia de los
Instrumentos

Test sobre la
Exogeneidad de
los Instrumentos

Test sobre la
Endogeneidad de
un Regresor

- En muchas aplicaciones hay un único regresor endógeno (X) pero varios regresores exógenos (W 's):

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 W_{1i} + \dots + \beta_{r+1} W_{ri} + u_i$$

donde **endógeno** significa “correlacionado con u ” y **exógeno** significa “no correlacionado con u ”.

- Ejemplo: función de demanda con precio, ingreso y edad como regresores.

MC2E con Múltiples Regresores Exógenos

Econometría

Dr. Christian M.
García-Witulski

MC2E con una X
y múltiples Z 's

MC2E con
Múltiples
Regresores
Exógenos

MC2E: Caso
General

Supuestos de VI

Propiedades de
MC2E

Tests sobre la
Validez de los
Instrumentos

Test sobre la
Relevancia de los
Instrumentos

Test sobre la
Exogeneidad de
los Instrumentos

Test sobre la
Endogeneidad de
un Regresor

- Para utilizar MC2E necesitamos al menos un instrumento (Z) para el regresor endógeno (X_1).
 - Sin instrumentos, el modelo está sub-identificado.
 - Con un único instrumento, el modelo está exactamente identificado.
 - Con más de un instrumento, el modelo está sobre-identificado.
- Con $m \geq 1$ instrumentos, el estimador MC2E se computa de la siguiente manera:
 - **Primera etapa:** correr una regresión de X_{1i} en los instrumentos (Z_{1i}, \dots, Z_{mi}) y los regresores exógenos (W_{1i}, \dots, W_{ri}), usando MCO; computar luego los valores de predicción \hat{X}_{1i} .
 - **Segunda etapa:** correr una regresión de Y_i en \hat{X}_{1i} y los regresores exógenos (W_{1i}, \dots, W_{ri}), usando MCO.
- ¿Por qué funciona este procedimiento?
 - Responderemos esta pregunta después de plantear el modelo general, con múltiples regresores endógenos..
- En la práctica, R computa las dos etapas conjuntamente.

MC2E: Caso General

Econometría

Dr. Christian M.
García-Witulski

MC2E con una X
y múltiples Z's

MC2E con
Múltiples
Regresores
Exógenos

MC2E: Caso
General

Supuestos de VI

Propiedades de
MC2E

Tests sobre la
Validez de los
Instrumentos

Test sobre la
Relevancia de los
Instrumentos

Test sobre la
Exogeneidad de
los Instrumentos

Test sobre la
Endogeneidad de
un Regresor

- Consideremos ahora un modelo con múltiples regresores endógenos (X 's) y múltiples regresores exógenos (W 's):

$$Y_i = \beta_0 + \beta_1 X_{1i} + \dots + \beta_k X_{ki} + \beta_{k+1} W_{1i} + \dots + \beta_{k+r} W_{ri} + u_i$$

- En este caso, el estimador MC2E se obtiene de la siguiente manera:

- Primera etapa:** para cada regresor endógeno, correr una regresión de X_{li} en Z_{1i}, \dots, Z_{mi} y W_{1i}, \dots, W_{ri} , usando MCO; computar luego los valores de predicción, $\hat{X}_{1i}, \dots, \hat{X}_{ki}$.
- Segunda etapa:** correr una regresión de Y_i en $\hat{X}_{1i}, \dots, \hat{X}_{ki}$ y W_{1i}, \dots, W_{ri} , usando MCO.

- R estima las dos etapas conjuntamente y computa los errores estándar apropiados.

MC2E: Caso General

Econometría

Dr. Christian M.
García-Witulski

MC2E con una X
y múltiples Z 's

MC2E con
Múltiples
Regresores
Exógenos

MC2E: Caso
General

Supuestos de VI

Propiedades de
MC2E

Tests sobre la
Validez de los
Instrumentos

Test sobre la
Relevancia de los
Instrumentos

Test sobre la
Exogeneidad de
los Instrumentos

Test sobre la
Endogeneidad de
un Regresor

- ¿Por qué funciona este procedimiento?

- En la primera etapa:

- la regresión de X_{1i} en Z_{1i}, \dots, Z_{mi} y W_{1i}, \dots, W_{ri} produce \hat{X}_{1i} y $\hat{v}_{1i} \equiv X_{1i} - \hat{X}_{1i}$;
- la regresión de X_{2i} en Z_{1i}, \dots, Z_{mi} y W_{1i}, \dots, W_{ri} produce \hat{X}_{2i} y $\hat{v}_{2i} \equiv X_{2i} - \hat{X}_{2i}$;
- ...
- la regresión de X_{ki} en Z_{1i}, \dots, Z_{mi} y W_{1i}, \dots, W_{ri} produce \hat{X}_{ki} y $\hat{v}_{ki} \equiv X_{ki} - \hat{X}_{ki}$.

- Por las propiedades algebraicas de MCO, sabemos que \hat{v}_{1i} no correlaciona con ninguno de los regresores $Z_{1i}, \dots, Z_{mi}, W_{1i}, \dots, W_{ri}$. Luego, \hat{v}_{1i} no correlaciona con ninguna combinación lineal de $Z_{1i}, \dots, Z_{mi}, W_{1i}, \dots, W_{ri}$.

- En particular, \hat{v}_{1i} no correlaciona con $\hat{X}_{1i}, \hat{X}_{2i}, \dots, \hat{X}_{ki}$ (ya que todos los valores de predicción son combinaciones lineales de $Z_{1i}, \dots, Z_{mi}, W_{1i}, \dots, W_{ri}$).

MC2E: Caso General

Econometría

Dr. Christian M.
García-Witulski

MC2E con una X
y múltiples Z's

MC2E con
Múltiples
Regresores
Exógenos

MC2E: Caso
General

Supuestos de VI

Propiedades de
MC2E

Tests sobre la
Validez de los
Instrumentos

Test sobre la
Relevancia de los
Instrumentos

Test sobre la
Exogeneidad de
los Instrumentos

Test sobre la
Endogeneidad de
un Regresor

- Análogamente:

- \hat{v}_{2i} no correlaciona con $\hat{X}_{1i}, \hat{X}_{2i}, \dots, \hat{X}_{ki}$.
- ...
- \hat{v}_{ki} no correlaciona con $\hat{X}_{1i}, \hat{X}_{2i}, \dots, \hat{X}_{ki}$.

- Sustituyendo $X_{1i} = \hat{X}_{1i} + \hat{v}_{1i}, \dots, X_{ki} = \hat{X}_{ki} + \hat{v}_{ki}$ en el modelo original obtenemos:

$$Y_i = \beta_0 + \beta_1(\hat{X}_{1i} + \hat{v}_{1i}) + \dots + \beta_k(\hat{X}_{ki} + \hat{v}_{ki}) + \beta_{k+1}W_{1i} + \dots + \beta_{k+r}W_{ri} +$$

$$Y_i = \beta_0 + \beta_1\hat{X}_{1i} + \dots + \beta_k\hat{X}_{ki} + \beta_{k+1}W_{1i} + \dots + \beta_{k+r}W_{ri} + \\ (u_i + \beta_1\hat{v}_{1i} + \dots + \beta_k\hat{v}_{ki})$$

- Luego:

$$Y_i = \beta_0 + \beta_1\hat{X}_{1i} + \dots + \beta_k\hat{X}_{ki} + \beta_{k+1}W_{1i} + \dots + \beta_{k+r}W_{ri} + e_i$$

$$\text{donde } e_i \equiv u_i + \beta_1\hat{v}_{1i} + \dots + \beta_k\hat{v}_{ki}.$$

MC2E: Caso General

Econometría

Dr. Christian M.
García-Witulski

MC2E con una X
y múltiples Z 's

MC2E con
Múltiples
Regresores
Exógenos

MC2E: Caso
General

Supuestos de VI

Propiedades de
MC2E

Tests sobre la
Validez de los
Instrumentos

Test sobre la
Relevancia de los
Instrumentos

Test sobre la
Exogeneidad de
los Instrumentos

Test sobre la
Endogeneidad de
un Regresor

- Ya sabemos que \hat{X}_{1i} no correlaciona con $\hat{v}_{1i}, \dots, \hat{v}_{ki}$.
- Sabemos también que \hat{X}_{1i} no correlaciona con u_i . La razón es que \hat{X}_{1i} es una combinación lineal de $Z_{1i}, \dots, Z_{mi}, W_{1i}, \dots, W_{ri}$, y ninguna de estas variables correlaciona con u_i (por hipótesis).
- Como e_i es una combinación lineal de $u_i, \hat{v}_{1i}, \dots, \hat{v}_{ki}$, concluimos que \hat{X}_{1i} no correlaciona con e_i .
- El mismo razonamiento se aplica a $\hat{X}_{2i}, \dots, \hat{X}_{ki}$: ninguna de ellas correlaciona con e_i .
- Recordemos que ninguna de las W 's correlaciona con u_i (por hipótesis); además, ninguna de las W 's correlaciona con $\hat{v}_{1i}, \dots, \hat{v}_{ki}$ (por las propiedades algebraicas de MCO). Luego, ninguna de las W 's correlaciona con e_i .
- Concluimos que ninguno de los regresores de la segunda etapa ($\hat{X}_{1i}, \dots, \hat{X}_{ki}, W_{1i}, \dots, W_{ri}$) correlaciona con el término de error (e_i).
- Luego, MC2E producirá estimadores consistentes para los β 's.

MC2E: Caso General

Econometría

Dr. Christian M.
García-Witulski

MC2E con una X
y múltiples Z 's

MC2E con
Múltiples
Regresores
Exógenos

MC2E: Caso
General

Supuestos de VI

Propiedades de
MC2E

Tests sobre la
Validez de los
Instrumentos

Test sobre la
Relevancia de los
Instrumentos

Test sobre la
Exogeneidad de
los Instrumentos

Test sobre la
Endogeneidad de
un Regresor

- Para utilizar MC2E necesitamos al menos tantos instrumentos (Z 's) como regresores endógenos (X_{1i}, \dots, X_{ki}):
 - Si $m < k$, el modelo está sub-identificado.
 - Si $m = k$, el modelo está exactamente identificado.
 - Si $m > k$, el modelo está sobre-identificado.
- Para usar MC2E necesitamos $m \geq k$.
- ¿Qué ocurre si $m < k$?
 - Tendremos multicolinealidad en la segunda etapa (ver Apéndice II).

Validez de los Instrumentos en el Modelo General

Econometría

Dr. Christian M.
García-Witulski

MC2E con una X
y múltiples Z 's

MC2E con
Múltiples
Regresores
Exógenos

MC2E: Caso
General

Supuestos de VI

Propiedades de
MC2E

Tests sobre la
Validez de los
Instrumentos

Test sobre la
Relevancia de los
Instrumentos

Test sobre la
Exogeneidad de
los Instrumentos

Test sobre la
Endogeneidad de
un Regresor

- Consideremos el modelo

$$Y_i = \beta_0 + \beta_1 X_{1i} + \dots + \beta_k X_{ki} + \beta_{k+1} W_{1i} + \dots + \beta_{k+r} W_{ri} + u_i$$

con m instrumentos, Z_1, \dots, Z_m .

- Las dos **condiciones para la validez de los instrumentos** son:
 - **Exogeneidad:** $Cov(Z_{1i}, u_i) = 0, \dots, Cov(Z_{mi}, u_i) = 0$.
 - **Relevancia:**
 - Caso general con múltiples X 's.
Supongamos que la regresión de la segunda etapa se pudiera estimar usando los valores de predicción de la primera etapa computados con los coeficientes poblacionales (i.e., con los verdaderos π 's en lugar de los $\hat{\pi}$'s). Entonces, no debería haber multicolinealidad perfecta en la regresión de la segunda etapa.
 - Caso particular con una sola X .
Al menos un instrumento debe tener un coeficiente (π) distinto de cero en la versión poblacional de la regresión de la primera etapa.

Supuestos de VI para Regresión Múltiple

Econometría

Dr. Christian M.
García-Witulski

MC2E con una X
y múltiples Z's

MC2E con
Múltiples
Regresores
Exógenos

MC2E: Caso
General

Supuestos de VI

Propiedades de
MC2E

Tests sobre la
Validez de los
Instrumentos

Test sobre la
Relevancia de los
Instrumentos

Test sobre la
Exogeneidad de
los Instrumentos

Test sobre la
Endogeneidad de
un Regresor

- Consideremos el modelo

$$Y_i = \beta_0 + \beta_1 X_{1i} + \dots + \beta_k X_{ki} + \beta_{k+1} W_{1i} + \dots + \beta_{k+r} W_{ri} + u_i$$

con m instrumentos (Z_{1i}, \dots, Z_{mi}) .

- IV Supuesto 1: **Exogeneidad**

$$\mathbb{E}(u_i \mid W_{1i}, \dots, W_{ri}) = 0$$

- IV Supuesto 2: **Muestreo Aleatorio Simple**

$(Y_i, X_{1i}, \dots, X_{ki}, W_{1i}, \dots, W_{ri}, Z_{1i}, \dots, Z_{mi})$ son *iid*

- IV Supuesto 3: **Improbabilidad de Outliers**

$$(Y_i, X_{1i}, \dots, X_{ki}, W_{1i}, \dots, W_{ri}, Z_{1i}, \dots, Z_{mi})$$

tienen momentos de cuarto orden positivos y finitos

- IV Supuesto 4: **Validez de los Instrumentos**

(Z_{1i}, \dots, Z_{mi}) son instrumentos válidos (exógenos y relevantes)

Propiedades de MC2E

Econometría

Dr. Christian M.
García-Witulski

MC2E con una X
y múltiples Z 's

MC2E con
Múltiples
Regresores
Exógenos

MC2E: Caso
General

Supuestos de VI

**Propiedades de
MC2E**

Tests sobre la
Validez de los
Instrumentos

Test sobre la
Relevancia de los
Instrumentos

Test sobre la
Exogeneidad de
los Instrumentos

Test sobre la
Endogeneidad de
un Regresor

- Bajo los supuestos de VI mencionados previamente, el estimador MC2E es consistente y asintóticamente normal.
- Las demostraciones son similares a las del caso particular de una X y una Z , pero las expresiones son más complicadas y no las derivaremos en este curso.
- Como la distribución muestral de los estimadores MC2E es aproximadamente normal para muestras grandes, el procedimiento para hacer inferencia (tests de hipótesis, intervalos de confianza) es análogo al que hemos usado con MCO.
 - Por ejemplo, un intervalo de confianza del 95 % para β_1 se construye de la siguiente manera: $\hat{\beta}_1^{\text{MC2E}} \pm 1.96 \times SE(\hat{\beta}_1^{\text{MC2E}})$.
 - El contraste de hipótesis conjuntas sobre los coeficientes poblacionales se puede hacer con el estadístico F .

Tests sobre la Validez de los Instrumentos

Econometría

Dr. Christian M.
García-Witulski

MC2E con una X
y múltiples Z's

MC2E con
Múltiples
Regresores
Exógenos

MC2E: Caso
General

Supuestos de VI

Propiedades de
MC2E

**Tests sobre la
Validez de los
Instrumentos**

Test sobre la
Relevancia de los
Instrumentos

Test sobre la
Exogeneidad de
los Instrumentos

Test sobre la
Endogeneidad de
un Regresor

- La utilidad de VI depende crucialmente de la validez de los instrumentos.
- ¿Cómo sabemos si nuestros instrumentos son válidos?
- Recordemos que los instrumentos serán válidos si son relevantes y exógenos.
- Consideremos las consecuencias de violar cada una de estas hipótesis.

Test sobre la Relevancia de los Instrumentos

Econometría

Dr. Christian M.
García-Witulski

MC2E con una X
y múltiples Z 's

MC2E con
Múltiples
Regresores
Exógenos

MC2E: Caso
General

Supuestos de VI

Propiedades de
MC2E

Tests sobre la
Validez de los
Instrumentos

**Test sobre la
Relevancia de los
Instrumentos**

Test sobre la
Exogeneidad de
los Instrumentos

Test sobre la
Endogeneidad de
un Regresor

- Con relación a la relevancia, el problema no es simplemente si los instrumentos son relevantes o no, sino *cuán relevantes* son.
- Cuanto mayor sea la variabilidad de X explicada por el instrumento, mayor será la información disponible para aprovechar en la segunda etapa.
- Cuando los instrumentos explican una parte pequeña de la variación de X se dice que hay un problema de **instrumentos débiles**.
- Tener un instrumento más relevante es como tener una muestra más grande, en el sentido que permite una estimación más precisa (menores errores estándar) y justifica la aproximación normal para hacer inferencia estadística.
- Por lo tanto, de la misma manera que una muestra pequeña, tener instrumentos débiles presenta un problema para la inferencia.
- Este problema es bastante serio en la práctica porque muchos trabajos aplicados usan instrumentos exógenos pero débiles.
- Más adelante veremos un ejemplo específico.

Test sobre la Relevancia de los Instrumentos

Econometría

Dr. Christian M.
García-Witulski

MC2E con una X
y múltiples Z 's

MC2E con
Múltiples
Regresores
Exógenos

MC2E: Caso
General

Supuestos de VI

Propiedades de
MC2E

Tests sobre la
Validez de los
Instrumentos

Test sobre la
Relevancia de los
Instrumentos

Test sobre la
Exogeneidad de
los Instrumentos

Test sobre la
Endogeneidad de
un Regresor

- ¿Por qué es problemática la debilidad de los instrumentos?
- Consideremos, por simplicidad, el caso de una X y una Z .

Luego:

$$\hat{\beta}_1^{\text{MC2E}} = \frac{s_{ZY}}{s_{ZX}}$$

- Si la covarianza poblacional entre Z y X es muy cercana a cero ($\sigma_{ZX} \cong 0$), pequeños cambios en la covarianza muestral s_{ZX} (de una muestra a otra) pueden generar cambios drásticos en $\hat{\beta}_1^{\text{MC2E}}$ (¿qué pasaría si en una muestra $s_{ZX} = 0,001$ y en otra $s_{ZX} = -0,001$?).
- En este caso, la aproximación normal para n grande resulta ser una mala aproximación a la distribución muestral de $\hat{\beta}_1^{2\text{sls}}$.
 - De hecho, para el caso particular con $\sigma_{ZX} = 0$ se puede demostrar que $\hat{\beta}_1^{\text{MC2E}}$ converge en distribución a un cociente de variables normales correlacionadas entre sí).
- Luego, si los instrumentos son débiles, el método usual de inferencia no es confiable.

Test sobre la Relevancia de los Instrumentos

Econometría

Dr. Christian M.
García-Witulski

MC2E con una X
y múltiples Z 's

MC2E con
Múltiples
Regresores
Exógenos

MC2E: Caso
General

Supuestos de VI

Propiedades de
MC2E

Tests sobre la
Validez de los
Instrumentos

**Test sobre la
Relevancia de los
Instrumentos**

Test sobre la
Exogeneidad de
los Instrumentos

Test sobre la
Endogeneidad de
un Regresor

- Test sobre debilidad de los instrumentos para el caso de **un único regresor endógeno** (una sola X).
 - Regla aproximada: (i) computar el estadístico F para la hipótesis nula de que los coeficientes de los instrumentos en la primera etapa son conjuntamente iguales a cero; (ii) si $F < 10$, hay un problema de instrumentos débiles.
 - Una justificación para esta regla se ofrece en el Apéndice 12.5 de S&W.
 - Nota: para el caso de un solo instrumento, se cumple $F = t^2$.
- ¿Qué hacer cuando los instrumentos son débiles?
 - Encontrar nuevos instrumentos.
 - En el caso de sobre-identificación, se puede descartar algún instrumento (nótese, sin embargo, que el test anterior es un test conjunto que no nos dice cuál o cuáles son los instrumentos problemáticos).
 - Usar un método de estimación alternativo.

Test sobre la Exogeneidad de los Instrumentos

Econometría

Dr. Christian M.
García-Witulski

MC2E con una X
y múltiples Z 's

MC2E con
Múltiples
Regresores
Exógenos

MC2E: Caso
General

Supuestos de VI

Propiedades de
MC2E

Tests sobre la
Validez de los
Instrumentos

Test sobre la
Relevancia de los
Instrumentos

Test sobre la
Exogeneidad de
los Instrumentos

Test sobre la
Endogeneidad de
un Regresor

- Si los instrumentos no son exógenos ($Cov(Z_i, u_i) \neq 0$), las \hat{X} 's estarán correlacionadas con u y los estimadores MC2E serán inconsistentes.
 - Esto conspira contra el objetivo principal de VI que consiste en aislar la “parte buena” de X .
- ¿Podemos hacer un test sobre la exogeneidad de los regresores?
 - Si el número de instrumentos es igual al número de variables endógenas (identificación exacta), no es posible hacer el test (hay que usar la teoría económica, la intuición, o algún otro criterio).
 - Si el número de instrumentos es mayor que el número de variables endógenas (sobre-identificación), entonces es posible hacer un test.
 - Veamos cómo.

Test sobre la Exogeneidad de los Instrumentos

Econometría

Dr. Christian M.
García-Witulski

MC2E con una X
y múltiples Z's

MC2E con
Múltiples
Regresores
Exógenos

MC2E: Caso
General

Supuestos de VI

Propiedades de
MC2E

Tests sobre la
Validez de los
Instrumentos

Test sobre la
Relevancia de los
Instrumentos

Test sobre la
Exogeneidad de
los Instrumentos

Test sobre la
Endogeneidad de
un Regresor

- La intuición es la siguiente:
 - Si algunos instrumentos son exógenos y otros no, entonces diferentes grupos de instrumentos generarán estimaciones significativamente distintas, algo que podemos testear.
 - Si $m = k$ hay un único grupo disponible de instrumentos (todos), de manera que el test no puede realizarse.
- El **test de restricciones de sobre-identificación** (OIR = *Over-Identifying Restrictions*) hace esta comparación (en forma implícita).

Test sobre la Exogeneidad de los Instrumentos

Econometría

Dr. Christian M.
García-Witulski

MC2E con una X
y múltiples Z 's

MC2E con
Múltiples
Regresores
Exógenos

MC2E: Caso
General

Supuestos de VI

Propiedades de
MC2E

Tests sobre la
Validez de los
Instrumentos

Test sobre la
Relevancia de los
Instrumentos

Test sobre la
Exogeneidad de
los Instrumentos

Test sobre la
Endogeneidad de
un Regresor

- Sabemos que la exogeneidad requiere ausencia de correlación entre las Z 's y u .
- En la práctica no observamos el error u_i , pero podemos estimarlo usando el residuo

$$\begin{aligned}\hat{u}_i^{\text{MC2E}} &= Y_i - \hat{Y}_i \\ &= Y_i - \hat{\beta}_0^{\text{MC2E}} - \hat{\beta}_1^{\text{MC2E}} X_{1i} - \dots - \hat{\beta}_{k+r}^{\text{MC2E}} W_{ri}\end{aligned}$$

- Nótese que \hat{Y}_i se calcula usando las verdaderas X 's, no las \hat{X} 's.
- Ahora podemos usar MCO para estimar los coeficientes de la siguiente regresión:

$$\hat{u}_i^{\text{MC2E}} = \delta_0 + \delta_1 Z_{1i} + \dots + \delta_m Z_{mi} + \delta_{m+1} W_{1i} + \dots + \delta_{m+r} W_{ri} + e_i$$

- Consideremos el estadístico F para la siguiente hipótesis nula:

$$H_0 : \delta_1 = \dots = \delta_m = 0$$

- El estadístico F se calcula asumiendo homocedasticidad de los errores.

Test sobre la Exogeneidad de los Instrumentos

Econometría

Dr. Christian M.
García-Witulski

MC2E con una X
y múltiples Z's

MC2E con
Múltiples
Regresores
Exógenos

MC2E: Caso
General

Supuestos de VI

Propiedades de
MC2E

Tests sobre la
Validez de los
Instrumentos

Test sobre la
Relevancia de los
Instrumentos

Test sobre la
Exogeneidad de
los Instrumentos

Test sobre la
Endogeneidad de
un Regresor

- Luego, bajo la hipótesis nula de que los instrumentos son exógenos, el estadístico para el test OIR es

$$J = mF \xrightarrow{d} \chi_{m-k}^2$$

con m = nro. de instrumentos y k = nro. de regresores endógenos.

- Nótese que $m = k$ implica la pérdida total de grados de libertad (no se puede hacer el test).
- Para el caso homocedástico existe un estadístico alternativo, conocido como estadístico de Sargan.
- También existe una variante para el caso no homocedástico.
- Si algunos instrumentos son exógenos y otros endógenos, el valor del estadístico J será grande y H_0 se rechazará.
- Si se rechaza H_0 , al menos un instrumento es endógeno (pero el test no nos dice cuál/es). ¿Qué podemos hacer?
 - Buscar instrumentos mejores.
 - Si hay sobre-identificación es posible descartar uno o más instrumentos (pero el test no nos dice cuál/es descartar).

Test sobre la Endogeneidad de un Regresor

Econometría

Dr. Christian M.
García-Witulski

MC2E con una X
y múltiples Z 's

MC2E con
Múltiples
Regresores
Exógenos

MC2E: Caso
General

Supuestos de VI

Propiedades de
MC2E

Tests sobre la
Validez de los
Instrumentos

Test sobre la
Relevancia de los
Instrumentos

Test sobre la
Exogeneidad de
los Instrumentos

Test sobre la
Endogeneidad de
un Regresor

- Consideremos el modelo de regresión simple

$$Y_i = \beta_0 + \beta_1 X_i + u_i$$

y supongamos que contamos con un instrumento válido, Z .

- Queremos testear si X es endógena. Si no lo es (*i.e.*, si es exógena) podríamos estimar el modelo directamente por MCO.
- Consideremos la regresión poblacional correspondiente a la primera etapa de MC2E,

$$X_i = \pi_0 + \pi_1 Z_i + v_i$$

- Si tuviésemos más instrumentos simplemente los agregaríamos como variables explicativas en la regresión anterior.

Test sobre la Endogeneidad de un Regresor

Econometría

Dr. Christian M.
García-Witulski

MC2E con una X
y múltiples Z's

MC2E con
Múltiples
Regresores
Exógenos

MC2E: Caso
General

Supuestos de VI

Propiedades de
MC2E

Tests sobre la
Validez de los
Instrumentos

Test sobre la
Relevancia de los
Instrumentos

Test sobre la
Exogeneidad de
los Instrumentos

Test sobre la
Endogeneidad de
un Regresor

- Ahora notemos que

$$\begin{aligned}Cov(X_i, u_i) &= Cov(\pi_0 + \pi_1 Z_i + v_i, u_i) \\&= \pi_1 Cov(Z_i, u_i) + Cov(v_i, u_i) \\&= Cov(v_i, u_i)\end{aligned}$$

donde hemos usado $Cov(Z_i, u_i) = 0$ (exogeneidad del instrumento).

- Luego:

$$Cov(X_i, u_i) \neq 0 \Leftrightarrow Cov(v_i, u_i) \neq 0$$

- Notemos que $Cov(v_i, u_i) \neq 0$ implica $\alpha \neq 0$ en la siguiente regresión:

$$u_i = \alpha v_i + \epsilon_i$$

donde ϵ_i es un término de error, con media cero, no correlacionado con v_i .

Test sobre la Endogeneidad de un Regresor

Econometría

Dr. Christian M.
García-Witulski

MC2E con una X
y múltiples Z 's

MC2E con
Múltiples
Regresores
Exógenos

MC2E: Caso
General

Supuestos de VI

Propiedades de
MC2E

Tests sobre la
Validez de los
Instrumentos

Test sobre la
Relevancia de los
Instrumentos

Test sobre la
Exogeneidad de
los Instrumentos

Test sobre la
Endogeneidad de
un Regresor

- Pero, según nuestro modelo: $u_i = Y_i - \beta_0 - \beta_1 X_i$.

Reemplazando en $u_i = \alpha v_i + \epsilon_i$ obtenemos:

$$Y_i - \beta_0 - \beta_1 X_i = \alpha v_i + \epsilon_i \Rightarrow$$

$$Y_i = \beta_0 + \beta_1 X_i + \alpha v_i + \epsilon_i$$

- Si pudiésemos correr la regresión anterior podríamos testear la endogeneidad de X contrastando las hipótesis

$$H_0 : \alpha = 0 \quad (X \text{ es exógena})$$

$$H_A : \alpha \neq 0 \quad (X \text{ es endógena})$$

- El problema con esta regresión es que v_i no es observable.
- Afortunadamente, α se puede estimar de manera consistente reemplazando v_i por \hat{v}_i (el residuo obtenido al estimar la primera etapa de MC2E, $X_i = \pi_0 + \pi_1 Z_i + v_i$).

Test sobre la Endogeneidad de un Regresor

Econometría

Dr. Christian M.
García-Witulski

MC2E con una X
y múltiples Z 's

MC2E con
Múltiples
Regresores
Exógenos

MC2E: Caso
General

Supuestos de VI

Propiedades de
MC2E

Tests sobre la
Validez de los
Instrumentos

Test sobre la
Relevancia de los
Instrumentos

Test sobre la
Exogeneidad de
los Instrumentos

Test sobre la
Endogeneidad de
un Regresor

- Por lo tanto, para testear la endogeneidad de X corremos la regresión

$$Y_i = \beta_0 + \beta_1 X_i + \alpha \hat{v}_i + \varepsilon_i$$

y contrastamos $H_0 : \alpha = 0$ (X es exógena) vs. $H_A : \alpha \neq 0$ (X es endógena) usando el estadístico t (o el estadístico F , que en este caso será igual al cuadrado de t).

- A este procedimiento se lo conoce como test de Hausman-Wu (o Durbin-Hausman-Wu).
- El test se generaliza fácilmente para el caso de k regresores endógenos: agregamos $\hat{v}_{1i}, \hat{v}_{2i}, \dots, \hat{v}_{ki}$ y usamos el estadístico F para testear la nula de que los coeficientes de los residuos son conjuntamente iguales a cero.
- Es importante recordar que para obtener este test hemos asumido que los instrumentos son válidos. Si no lo son, el test no funciona.
- Además, la muestra debe ser suficientemente grande para que el test funcione bien (ya que hemos reemplazado v_i por \hat{v}_i).

- Recordemos nuestra estimación de la demanda de cigarrillos obtenida con datos de 1995 para los 48 estados continentales de EEUU. Usando *SalesTax* como instrumento para $\ln P$ obtuvimos:

```
cig_ivreg <- ivreg(log(packs) ~ log(rprice) | salestax, data=c1995)
coeftest(cig_ivreg, vcov = vcovHC, type = "HC1")
```

```
##
## t test of coefficients:
##
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  9.71988      1.52832  6.3598 8.346e-08 ***
## log(rprice) -1.08359      0.31892 -3.3977 0.001411 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

- Luego:

$$\widehat{\ln Q_i} = \underset{(1,53)}{9,72} - \underset{(0,32)}{1,08} \ln P_i$$

Ejemplo: Demanda de Cigarrillos

Econometría

Dr. Christian M.
García-Witulski

MC2E con una X
y múltiples Z's

MC2E con
Múltiples
Regresores
Exógenos

MC2E: Caso
General

Supuestos de VI

Propiedades de
MC2E

Tests sobre la
Validez de los
Instrumentos

Test sobre la
Relevancia de los
Instrumentos

Test sobre la
Exogeneidad de
los Instrumentos

Test sobre la
Endogeneidad de
un Regresor

- Como el cigarrillo provoca adicción, una elasticidad de $-1,08$ parece realmente alta: una suba del precio del 1 % reduciría el consumo de cigarrillos en 1.08 %.
- Pero la exogeneidad podría ser un problema...
- Además, deberíamos prestar atención a un posible sesgo por variables omitidas.
- Probablemente deberíamos incluir más variables explicativas (W 's).
- Un candidato obvio para variable omitida es el ingreso.
 - Es posible que los estados con mayor ingreso recurran más al impuesto al ingreso que al impuesto a las ventas ($Cov(Ingreso, SalesTax) < 0$).
 - Además, es muy probable que la demanda de cigarrillos dependa del ingreso ($Cov(Ingreso, u) \neq 0$).
- De ahora en más agregaremos como regresor al logaritmo del ingreso real per cápita de cada estado, $\ln Inc_i$.

- Los resultados de MC2E son los siguientes:

```
cig_ivreg2 <- ivreg(log(packs) ~ log(rprice) + log(rincome)
                    | salestax + log(rincome), data = c1995)
coeftest(cig_ivreg2, vcov = vcovHC, type = "HC1")
```

```
##
## t test of coefficients:
##
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)   9.43066    1.25939   7.4883 1.935e-09 ***
## log(rprice)  -1.14338    0.37230  -3.0711 0.003611 **
## log(rincome)  0.21452    0.31175   0.6881 0.494917
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

- Luego:

$$\widehat{\ln Q_i} = 9,43 - 1,14 \ln P_i + 0,21 \ln Inc_i$$

(1,26) (0,37) (0,31)

Ejemplo: Demanda de Cigarrillos

Econometría

Dr. Christian M.
García-Witulski

MC2E con una X
y múltiples Z's

MC2E con
Múltiples
Regresores
Exógenos

MC2E: Caso
General

Supuestos de VI

Propiedades de
MC2E

Tests sobre la
Validez de los
Instrumentos

Test sobre la
Relevancia de los
Instrumentos

Test sobre la
Exogeneidad de
los Instrumentos

Test sobre la
Endogeneidad de
un Regresor

- ¿Relevancia? Tenemos que analizar el resultado de la primera etapa:

$$\widehat{\ln P_i} = \underset{(0,173)}{3,59} + \underset{(0,0041)}{0,0274}SalesTax_i + \underset{(0,065)}{0,389} \ln Inc_i$$

- El estadístico F para el test de instrumentos débiles es $F = t^2 = \left(\frac{0,0274}{0,0041}\right)^2 = 44,7 > 10$. No hay señales de debilidad del instrumento.
- ¿Exogeneidad?
 - No podemos hacer el test: el modelo está exactamente identificado.
- Sin embargo, en la base de datos tenemos otra variable que podría servir como instrumento: el impuesto específico a los cigarrillos ($CigTax$).
- Veamos qué ocurre cuando usamos los dos impuestos ($SalesTax$ y $CigTax$) como instrumentos.

- Los resultados de MC2E son los siguientes:

```
cig_ivreg3 <- ivreg(log(packs) ~ log(rprice) + log(rincome)
                    | salestax + cigtax + log(rincome),
                    data = c1995)
coeftest(cig_ivreg3, vcov = vcovHC, type = "HC1")
```

```
##
## t test of coefficients:
##
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)   9.89496    0.95922  10.3157 1.947e-13 ***
## log(rprice)  -1.27742    0.24961  -5.1177 6.211e-06 ***
## log(rincome)  0.28040    0.25389   1.1044  0.2753
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

- Luego:

$$\widehat{\ln Q_i} = \underset{(0,96)}{9,89} - \underset{(0,25)}{1,28} \ln P_i + \underset{(0,25)}{0,28} \ln Inc_i$$

- Nótese que seguimos estimando una elasticidad de la demanda muy alta.

- ¿Relevancia? Nuevamente, tenemos que analizar los resultados de la primera etapa:

$$\widehat{\ln P_i} = 4,10 + 0,011 \text{ SalesTax}_i + 0,009 \text{ CigTax}_i + 0,108 \ln \text{Inc}_i$$

(0,088) (0,002) (0,001) (0,040)

- Ahora usamos R para contrastar la hipótesis de que los coeficientes de *SalesTax* & *CigTax* son conjuntamente iguales a cero.

```
mod_relevance <- lm(log(rprice) ~ salestax + cigtax + log(rincome),
                     data = c1995)
coeftest(mod_relevance, vcov = vcovHC, type = "HC1")
```

```
##
## t test of coefficients:
##
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  4.1030339  0.0883803  46.4248 < 2.2e-16 ***
## salestax      0.0108898  0.0021366   5.0968 6.999e-06 ***
## cigtax        0.0093517  0.0008698  10.7515 6.820e-14 ***
## log(rincome)  0.1083449  0.0396526   2.7324 0.009019 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
linearHypothesis(mod_relevance,
c("salestax = 0", "cigtax = 0"),
white.adjust = "hcl")
```

```
## Linear hypothesis test
##
## Hypothesis:
## salestax = 0
## cigtax = 0
##
## Model 1: restricted model
## Model 2: log(rprice) ~ salestax + cigtax + log(rincome)
##
## Note: Coefficient covariance matrix supplied.
##
##   Res.Df Df      F    Pr(>F)
## 1      46
## 2      44  2 209.68 < 2.2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Vemos que $F = 209,7 > 10$: los instrumentos no son débiles.

- Como ahora tenemos un modelo sobre-identificado, podemos testear la exogeneidad de los instrumentos con el estadístico J .

```
cig_iv_OR <- lm(residuals(cig_ivreg3) ~ log(rincome) + salestax
               + cigtax, data = c1995)
cig_OR_test <- linearHypothesis(cig_iv_OR,
c("salestax = 0", "cigtax = 0"),
test = "Chisq")
cig_OR_test
```

```
## Linear hypothesis test
##
## Hypothesis:
## salestax = 0
## cigtax = 0
##
##   Res.Df    RSS Df Sum of Sq  Chisq Pr(>Chisq)
## 1      46 1.588
## 2      44 1.577  2  0.011005 0.307    0.8577
```

- Obtenemos: $J_{value} = 0,307$

- Cuidado!: En este caso, el valor-p informado por **linearHypothesis()** es incorrecto porque los grados de libertad se establecen en 2. Esto difiere del grado de sobreidentificación ($m - k = 2 - 1 = 1$) por lo que el estadístico-J está distribuido como χ_1^2 en lugar de χ_2^2 como asume por defecto **linearHypothesis()**. Podemos calcular el valor-p correcto usando **pchisq()**.

```
pchisq(cig_OR_test[2, 5], df = 1, lower.tail = FALSE)
```

```
## [1] 0.5795077
```

- Obtenemos:

$$p = 0,58$$

- No rechazamos la hipótesis nula de exogeneidad de los instrumentos

Ejemplo: Demanda de Cigarrillos

Econometría

Dr. Christian M.
García-Witulski

MC2E con una X
y múltiples Z's

MC2E con
Múltiples
Regresores
Exógenos

MC2E: Caso
General

Supuestos de VI

Propiedades de
MC2E

Tests sobre la
Validez de los
Instrumentos

Test sobre la
Relevancia de los
Instrumentos

Test sobre la
Exogeneidad de
los Instrumentos

Test sobre la
Endogeneidad de
un Regresor

- Finalmente, podemos usar el test de Hausman-Wu para testear si $\ln P_i$ es endógena.
 - Recordemos que, según los tests previos, no rechazamos la validez de los instrumentos (requisito indispensable para que el test de Hausman-Wu funcione bien).
- Comenzamos corriendo la regresión

$$\ln Q_i = \beta_0 + \beta_1 \ln P_i + \beta_2 \ln Inc_i + \alpha \hat{v}_i + e_i$$

donde \hat{v}_i es el residuo de

$$\ln P_i = \pi_0 + \pi_1 SalesTax_i + \pi_2 CigTax_i + \pi_3 \ln Inc_i + v_i.$$

Econometría

Dr. Christian M.
García-Witulski

MC2E con una X
y múltiples Z's

MC2E con
Múltiples
Regresores
Exógenos

MC2E: Caso
General

Supuestos de VI

Propiedades de
MC2E

Tests sobre la
Validez de los
Instrumentos

Test sobre la
Relevancia de los
Instrumentos

Test sobre la
Exogeneidad de
los Instrumentos

Test sobre la
Endogeneidad de
un Regresor

```
# Test de Hausman-Wu
# Estimamos el modelo en su primer etapa
mod_relevance <- lm(log(rprice) ~ salestax + cigtax + log(rincome),
                    data = c1995)
# Estimamos el modelo con los residuos
mod_test_HW <- lm(log(packs) ~ log(rprice) + log(rincome)
                  + residuals(mod_relevance), data = c1995)
coeftest(mod_test_HW, vcov = vcovHC, type = "HC1")
```

```
##
## t test of coefficients:
##
##
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    9.89496   0.86554 11.4321 9.194e-15 ***
## log(rprice)   -1.27742   0.22809 -5.6006 1.299e-06 ***
## log(rincome)    0.28040   0.24725  1.1341  0.26290
## residuals(mod_relevance) -1.56495   0.83592 -1.8721  0.06785 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Ejemplo: Demanda de Cigarrillos

Econometría

Dr. Christian M.
García-Witulski

MC2E con una X
y múltiples Z's

MC2E con
Múltiples
Regresores
Exógenos

MC2E: Caso
General

Supuestos de VI

Propiedades de
MC2E

Tests sobre la
Validez de los
Instrumentos

Test sobre la
Relevancia de los
Instrumentos

Test sobre la
Exogeneidad de
los Instrumentos

Test sobre la
Endogeneidad de
un Regresor

- Obtenemos:

$$\widehat{\ln Q}_i = 9,89 - 1,28 \ln P_i + 0,28 \ln Inc_i - 1,57 \widehat{v}_i$$

(0,87) (0,23) (0,25) (0,84)

- Nótese que los coeficientes estimados para $\ln P_i$ y $\ln Inc_i$ coinciden con los obtenidos usando MC2E. Esto no es una casualidad, ocurrirá siempre.

- Test t sobre α :

$$t_{value} = \frac{-1,57}{0,84} = -1,87$$

$$p = 0,07$$

- La hipótesis nula de exogeneidad se rechaza para niveles de significatividad iguales a superiores a 0,07.
 - Con la significatividad típica de 0,05 no se rechaza H_0 .
 - Notemos, sin embargo, que el tamaño de nuestra muestra es relativamente chico: $n = 48$.

Instrumentos Débiles: Un Ejemplo

Econometría

Dr. Christian M.
García-Witulski

MC2E con una X
y múltiples Z's

MC2E con
Múltiples
Regresores
Exógenos

MC2E: Caso
General

Supuestos de VI

Propiedades de
MC2E

Tests sobre la
Validez de los
Instrumentos

Test sobre la
Relevancia de los
Instrumentos

Test sobre la
Exogeneidad de
los Instrumentos

Test sobre la
Endogeneidad de
un Regresor

- Mucha gente está interesada en estimar el efecto de la educación sobre el ingreso (el retorno de la educación).
- El problema es que la educación está potencialmente correlacionada con variables no osbservables, como la capacidad.
- ¿Podemos resolver este problema con VI?
- En principio sí, ¿pero cuál sería nuestro instrumento?
- Joshua Angrist y Alan Krueger (1991) tuvieron una idea interesante:
 - En EEUU, la ley obliga a los alumnos a permanecer en la escuela hasta su cumpleaños N^o 16.
 - Ahora bien, un alumno que cumple 16 en enero, mientras está en el 10^o grado, puede abandonar el colegio sin completar dicho grado.
 - En cambio, un alumno que cumple 16 en julio, estará obligado a completar el 10^o grado.
- Luego, la fecha de cumpleaños puede funcionar como instrumento para el nivel educativo: no debería tener efectos sobre el ingreso excepto a través de su impacto sobre el nivel educativo.

Instrumentos Débiles: Un Ejemplo

Econometría

Dr. Christian M.
García-Witulski

MC2E con una X
y múltiples Z's

MC2E con
Múltiples
Regresores
Exógenos

MC2E: Caso
General

Supuestos de VI

Propiedades de
MC2E

Tests sobre la
Validez de los
Instrumentos

Test sobre la
Relevancia de los
Instrumentos

Test sobre la
Exogeneidad de
los Instrumentos

Test sobre la
Endogeneidad de
un Regresor

- Angrist & Krueger usaron datos de EEUU para estimar

$$\ln Wage_i = \beta_0 + \beta_1 educ_i + OtherStuff + u_i$$

usando el trimestre de nacimiento como instrumento para la educación.

- Encontraron un retorno a la educación de aproximadamente 8 % (mayor que el que encontraron usando MCO, lo que resulta sorprendente ya que uno esperaría que MCO sobre-estime el efecto de la educación).
- El estadístico t correspondiente a $H_0 : \beta_1 = 0$ es muy grande (> 5).
- ¿Hay algún problema con estos resultados?

Instrumentos Débiles: Un Ejemplo

Econometría

Dr. Christian M.
García-Witulski

MC2E con una X
y múltiples Z's

MC2E con
Múltiples
Regresores
Exógenos

MC2E: Caso
General

Supuestos de VI

Propiedades de
MC2E

Tests sobre la
Validez de los
Instrumentos

Test sobre la
Relevancia de los
Instrumentos

Test sobre la
Exogeneidad de
los Instrumentos

Test sobre la
Endogeneidad de
un Regresor

- A pesar de tener una muestra enorme ($n = 329,000$!), el estadístico F de la primera etapa era (en algunas especificaciones) menor que 2.
- Luego, el instrumento sufre de debilidad.
- Es más, John Bound mostró que reemplazando el instrumento utilizado por Angrist & Krueger con una fecha de nacimiento generada aleatoriamente (irrelevante por diseño) se obtienen resultados muy similares.
- Este resultado muestra los peligros de usar instrumentos débiles.

Apéndice I: Consistencia de MC2E

Econometría

Dr. Christian M.
García-Witulski

MC2E con una X
y múltiples Z 's

MC2E con
Múltiples
Regresores
Exógenos

MC2E: Caso
General

Supuestos de VI

Propiedades de
MC2E

Tests sobre la
Validez de los
Instrumentos

Test sobre la
Relevancia de los
Instrumentos

Test sobre la
Exogeneidad de
los Instrumentos

Test sobre la
Endogeneidad de
un Regresor

- Mostraremos que $\hat{\beta}_1^{\text{MC2E}}$ es un estimador consistente de β_1 . Lo haremos para el caso de una X y múltiples instrumentos.

- A partir de $\hat{\beta}_1^{\text{MC2E}} = \frac{s_{\hat{X}Y}}{s_{\hat{X}}^2}$ obtenemos

$$\hat{\beta}_1^{\text{MC2E}} = \frac{\sum_{i=1}^n (\hat{X}_i - \bar{\hat{X}})(Y_i - \bar{Y})}{\sum_{i=1}^n (\hat{X}_i - \bar{\hat{X}})^2}$$

- Usando $Y_i = \beta_0 + \beta_1 \hat{X}_i + e_i$:

$$\begin{aligned}\hat{\beta}_1^{\text{MC2E}} &= \frac{\sum_{i=1}^n (\hat{X}_i - \bar{\hat{X}})(\beta_0 + \beta_1 \hat{X}_i + e_i - \beta_0 - \beta_1 \bar{\hat{X}} - \bar{e})}{\sum_{i=1}^n (\hat{X}_i - \bar{\hat{X}})^2} \\&= \frac{\sum_{i=1}^n (\hat{X}_i - \bar{\hat{X}})[\beta_1 (\hat{X}_i - \bar{\hat{X}}) + (e_i - \bar{e})]}{\sum_{i=1}^n (\hat{X}_i - \bar{\hat{X}})^2} \\&= \beta_1 \frac{\sum_{i=1}^n (\hat{X}_i - \bar{\hat{X}})^2}{\sum_{i=1}^n (\hat{X}_i - \bar{\hat{X}})^2} + \frac{\sum_{i=1}^n (\hat{X}_i - \bar{\hat{X}})(e_i - \bar{e})}{\sum_{i=1}^n (\hat{X}_i - \bar{\hat{X}})^2}\end{aligned}$$

Apéndice I: Consistencia de MC2E

Econometría

Dr. Christian M.
García-Witulski

MC2E con una X
y múltiples Z's

MC2E con
Múltiples
Regresores
Exógenos

MC2E: Caso
General

Supuestos de VI

Propiedades de
MC2E

Tests sobre la
Validez de los
Instrumentos

Test sobre la
Relevancia de los
Instrumentos

Test sobre la
Exogeneidad de
los Instrumentos

Test sobre la
Endogeneidad de
un Regresor

- Luego:

$$\hat{\beta}_1^{\text{MC2E}} = \beta_1 + \frac{\frac{1}{n} \sum_{i=1}^n (\hat{X}_i - \bar{\hat{X}})(e_i - \bar{e})}{\frac{1}{n} \sum_{i=1}^n (\hat{X}_i - \bar{\hat{X}})^2}$$

- Usando $e_i = \beta_1 \hat{v}_i + u_i$:

$$\begin{aligned}\hat{\beta}_1^{\text{MC2E}} &= \beta_1 + \frac{\frac{1}{n} \sum_{i=1}^n (\hat{X}_i - \bar{\hat{X}})[\beta_1(\hat{v}_i - \bar{\hat{v}}) + (u_i - \bar{u})]}{\frac{1}{n} \sum_{i=1}^n (\hat{X}_i - \bar{\hat{X}})^2} \\&= \beta_1 + \beta_1 \frac{\frac{1}{n} \sum_{i=1}^n (\hat{X}_i - \bar{\hat{X}})(\hat{v}_i - \bar{\hat{v}})}{\frac{1}{n} \sum_{i=1}^n (\hat{X}_i - \bar{\hat{X}})^2} + \frac{\frac{1}{n} \sum_{i=1}^n (\hat{X}_i - \bar{\hat{X}})(u_i - \bar{u})}{\frac{1}{n} \sum_{i=1}^n (\hat{X}_i - \bar{\hat{X}})^2} \\&= \beta_1 + \frac{\frac{1}{n} \sum_{i=1}^n (\hat{X}_i - \bar{\hat{X}})(u_i - \bar{u})}{\frac{1}{n} \sum_{i=1}^n (\hat{X}_i - \bar{\hat{X}})^2}\end{aligned}$$

donde hemos usado $\frac{1}{n} \sum_{i=1}^n (\hat{X}_i - \bar{\hat{X}})(\hat{v}_i - \bar{\hat{v}}) = \text{Cov}(\hat{X}_i, \hat{v}_i) = 0$
(propiedad de MCO).

Apéndice I: Consistencia de MC2E

Econometría

Dr. Christian M.
García-Witulski

MC2E con una X
y múltiples Z's

MC2E con
Múltiples
Regresores
Exógenos

MC2E: Caso
General

Supuestos de VI

Propiedades de
MC2E

Tests sobre la
Validez de los
Instrumentos

Test sobre la
Relevancia de los
Instrumentos

Test sobre la
Exogeneidad de
los Instrumentos

Test sobre la
Endogeneidad de
un Regresor

- Ahora supongamos, por simplicidad, que tenemos dos instrumentos (el caso general es similar). Luego:

$$\hat{X}_i = \hat{\pi}_0 + \hat{\pi}_1 Z_{1i} + \hat{\pi}_2 Z_{2i}$$

- Luego:

$$\overline{\hat{X}} = \hat{\pi}_0 + \hat{\pi}_1 \overline{Z}_1 + \hat{\pi}_2 \overline{Z}_2$$

- Luego:

$$\begin{aligned} \frac{1}{n} \sum_{i=1}^n (\hat{X}_i - \overline{\hat{X}})^2 &= \frac{1}{n} \sum_{i=1}^n [\hat{\pi}_1 (Z_{1i} - \overline{Z}_1) + \hat{\pi}_2 (Z_{2i} - \overline{Z}_2)]^2 \\ &= \frac{1}{n} \sum_{i=1}^n \left[\begin{aligned} &(\hat{\pi}_1)^2 (Z_{1i} - \overline{Z}_1)^2 + (\hat{\pi}_2)^2 (Z_{2i} - \overline{Z}_2)^2 \\ &+ 2\hat{\pi}_1 \hat{\pi}_2 (Z_{1i} - \overline{Z}_1)(Z_{2i} - \overline{Z}_2) \end{aligned} \right] \\ &= (\hat{\pi}_1)^2 \frac{1}{n} \sum_{i=1}^n (Z_{1i} - \overline{Z}_1)^2 + (\hat{\pi}_2)^2 \frac{1}{n} \sum_{i=1}^n (Z_{2i} - \overline{Z}_2)^2 \\ &\quad + 2\hat{\pi}_1 \hat{\pi}_2 \frac{1}{n} \sum_{i=1}^n (Z_{1i} - \overline{Z}_1)(Z_{2i} - \overline{Z}_2) \end{aligned}$$

Apéndice I: Consistencia de MC2E

Econometría

Dr. Christian M.
García-Witulski

MC2E con una X
y múltiples Z's

MC2E con
Múltiples
Regresores
Exógenos

MC2E: Caso
General

Supuestos de VI

Propiedades de
MC2E

Tests sobre la
Validez de los
Instrumentos

Test sobre la
Relevancia de los
Instrumentos

Test sobre la
Exogeneidad de
los Instrumentos

Test sobre la
Endogeneidad de
un Regresor

- Pero:

$$\frac{1}{n} \sum_{i=1}^n (Z_{1i} - \bar{Z}_1)^2 \xrightarrow{p} \text{Var}(Z_1) = \sigma_{Z_1}^2$$

$$\frac{1}{n} \sum_{i=1}^n (Z_{2i} - \bar{Z}_2)^2 \xrightarrow{p} \text{Var}(Z_2) = \sigma_{Z_2}^2$$

$$\frac{1}{n} \sum_{i=1}^n (Z_{1i} - \bar{Z}_1)(Z_{2i} - \bar{Z}_2) \xrightarrow{p} \text{Cov}(Z_1, Z_2) = \sigma_{Z_1 Z_2}$$

- Además:

$$\hat{\pi}_1 \xrightarrow{p} \pi_1$$

$$\hat{\pi}_2 \xrightarrow{p} \pi_2$$

ya que hemos asumido $\text{Cov}(Z_{1i}, v_i) = 0$ y $\text{Cov}(Z_{2i}, v_i) = 0$ en la primera etapa.

- Luego:

$$\frac{1}{n} \sum_{i=1}^n (\hat{X}_i - \bar{X})^2 \xrightarrow{p} (\pi_1)^2 \sigma_{Z_1}^2 + (\pi_2)^2 \sigma_{Z_2}^2 + 2\pi_1 \pi_2 \sigma_{Z_1 Z_2}$$

Apéndice I: Consistencia de MC2E

Econometría

Dr. Christian M.
García-Witulski

MC2E con una X
y múltiples Z's

MC2E con
Múltiples
Regresores
Exógenos

MC2E: Caso
General

Supuestos de VI

Propiedades de
MC2E

Tests sobre la
Validez de los
Instrumentos

Test sobre la
Relevancia de los
Instrumentos

Test sobre la
Exogeneidad de
los Instrumentos

Test sobre la
Endogeneidad de
un Regresor

- También se cumple

$$\begin{aligned}\frac{1}{n} \sum_{i=1}^n (\hat{X}_i - \bar{\hat{X}})(u_i - \bar{u}) &= \frac{1}{n} \sum_{i=1}^n [\hat{\pi}_1 (Z_{1i} - \bar{Z}_1) + \hat{\pi}_2 (Z_{2i} - \bar{Z}_2)](u_i - \bar{u}) \\ &= \hat{\pi}_1 \frac{1}{n} \sum_{i=1}^n (Z_{1i} - \bar{Z}_1)(u_i - \bar{u}) + \hat{\pi}_2 \frac{1}{n} \sum_{i=1}^n (Z_{2i} - \bar{Z}_2)(u_i - \bar{u})\end{aligned}$$

- Pero

$$\frac{1}{n} \sum_{i=1}^n (Z_{1i} - \bar{Z}_1)(u_i - \bar{u}) \xrightarrow{p} Cov(Z_1, u) = 0$$

$$\frac{1}{n} \sum_{i=1}^n (Z_{2i} - \bar{Z}_2)(u_i - \bar{u}) \xrightarrow{p} Cov(Z_2, u) = 0$$

donde $Cov(Z_1, u) = 0$ y $Cov(Z_2, u) = 0$ por la **exogeneidad de los instrumentos**.

- Luego:

$$\frac{1}{n} \sum_{i=1}^n (\hat{X}_i - \bar{\hat{X}})(u_i - \bar{u}) \xrightarrow{p} \pi_1 \times 0 + \pi_2 \times 0 = 0$$

Apéndice I: Consistencia de MC2E

Econometría

Dr. Christian M.
García-Witulski

MC2E con una X
y múltiples Z's

MC2E con
Múltiples
Regresores
Exógenos

MC2E: Caso
General

Supuestos de VI

Propiedades de
MC2E

Tests sobre la
Validez de los
Instrumentos

Test sobre la
Relevancia de los
Instrumentos

Test sobre la
Exogeneidad de
los Instrumentos

Test sobre la
Endogeneidad de
un Regresor

- Volvamos ahora a nuestra expresión para $\hat{\beta}_1^{\text{MC2E}}$:

$$\hat{\beta}_1^{\text{MC2E}} = \beta_1 + \frac{\frac{1}{n} \sum_{i=1}^n (\hat{X}_i - \bar{\hat{X}})(u_i - \bar{u})}{\frac{1}{n} \sum_{i=1}^n (\hat{X}_i - \bar{\hat{X}})^2}$$

- Usando los resultados previos:

$$\hat{\beta}_1^{\text{MC2E}} \xrightarrow{p} \beta_1 + \frac{0}{(\pi_1)^2 \sigma_{Z_1}^2 + (\pi_2)^2 \sigma_{Z_2}^2 + 2\pi_1 \pi_2 \sigma_{Z_1 Z_2}}$$

- La **relevancia de los instrumentos** requiere que $\pi_1 \neq 0$ o $\pi_2 \neq 0$ (o ambas). Por lo tanto, el denominador de la expresión anterior es distinto de cero. Luego:

$$\hat{\beta}_1^{\text{MC2E}} \xrightarrow{p} \beta_1$$

- Concluimos que, si los instrumentos son válidos (relevantes y exógenos), $\hat{\beta}_1^{\text{MC2E}}$ es un estimador consistente de β_1 .

Apéndice II: Sub-identificación

Econometría

Dr. Christian M.
García-Witulski

MC2E con una X
y múltiples Z 's

MC2E con
Múltiples
Regresores
Exógenos

MC2E: Caso
General

Supuestos de VI

Propiedades de
MC2E

Tests sobre la
Validez de los
Instrumentos

Test sobre la
Relevancia de los
Instrumentos

Test sobre la
Exogeneidad de
los Instrumentos

Test sobre la
Endogeneidad de
un Regresor

- Mostramos ahora que $m < k$ (menos instrumentos que regresores endógenos) implica multicolinealidad en la segunda etapa.
- Por simplicidad, consideremos un modelo con dos regresores endógenos ($k = 2$), un regresor exógeno ($r = 1$)

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \beta_3 W_i + u_i$$

y un solo instrumento Z_i ($m = 1$). Luego, $m = 1 < 2 = k$.

- A partir de la primera etapa obtenemos:

$$(1) \quad \hat{X}_{1i} = \hat{\pi}_0 + \hat{\pi}_1 Z_i + \hat{\pi}_2 W_i$$

$$(2) \quad \hat{X}_{2i} = \hat{\theta}_0 + \hat{\theta}_1 Z_i + \hat{\theta}_2 W_i$$

- La regresión de la segunda etapa es

$$(3) \quad Y_i = \beta_0 + \beta_1 \hat{X}_{1i} + \beta_2 \hat{X}_{2i} + \beta_3 W_i + e_i$$

Apéndice II: Sub-identificación

Econometría

Dr. Christian M.
García-Witulski

MC2E con una X
y múltiples Z's

MC2E con
Múltiples
Regresores
Exógenos

MC2E: Caso
General

Supuestos de VI

Propiedades de
MC2E

Tests sobre la
Validez de los
Instrumentos

Test sobre la
Relevancia de los
Instrumentos

Test sobre la
Exogeneidad de
los Instrumentos

Test sobre la
Endogeneidad de
un Regresor

- A partir de (2) obtenemos: $\hat{\theta}_1 Z_i = -\hat{\theta}_0 - \hat{\theta}_2 W_i + \hat{X}_{2i} \Rightarrow$

$$(4) \quad Z_i = -\frac{\hat{\theta}_0}{\hat{\theta}_1} - \frac{\hat{\theta}_2}{\hat{\theta}_1} W_i + \frac{1}{\hat{\theta}_1} \hat{X}_{2i}$$

- Sustituyendo (4) en (1):

$$\hat{X}_{1i} = \hat{\pi}_0 + \hat{\pi}_1 \left(-\frac{\hat{\theta}_0}{\hat{\theta}_1} - \frac{\hat{\theta}_2}{\hat{\theta}_1} W_i + \frac{1}{\hat{\theta}_1} \hat{X}_{2i} \right) + \hat{\pi}_2 W_i$$

$$\hat{X}_{1i} = \left(\hat{\pi}_0 - \hat{\pi}_1 \frac{\hat{\theta}_0}{\hat{\theta}_1} \right) + \frac{\hat{\pi}_1}{\hat{\theta}_1} \hat{X}_{2i} + \left(\hat{\pi}_2 - \hat{\pi}_1 \frac{\hat{\theta}_2}{\hat{\theta}_1} \right) W_i$$

$$(5) \quad \hat{X}_{1i} = a + b \hat{X}_{2i} + c W_i$$

donde $a \equiv \hat{\pi}_0 - \hat{\pi}_1 \frac{\hat{\theta}_0}{\hat{\theta}_1}$, $b \equiv \frac{\hat{\pi}_1}{\hat{\theta}_1}$ y $c \equiv \hat{\pi}_2 - \hat{\pi}_1 \frac{\hat{\theta}_2}{\hat{\theta}_1}$ son coeficientes conocidos.

- Como \hat{X}_{1i} es una combinación lineal exacta de \hat{X}_{2i} y W_i , la regresión de la segunda etapa tendrá un problema de multicolinealidad.

Apéndice II: Sub-identificación

Econometría

Dr. Christian M.
García-Witulski

MC2E con una X
y múltiples Z's

MC2E con
Múltiples
Regresores
Exógenos

MC2E: Caso
General

Supuestos de VI

Propiedades de
MC2E

Tests sobre la
Validez de los
Instrumentos

Test sobre la
Relevancia de los
Instrumentos

Test sobre la
Exogeneidad de
los Instrumentos

Test sobre la
Endogeneidad de
un Regresor

- Para entender por qué la multicolinealidad implica que β_0 , β_1 , β_2 y β_3 no están identificados, sustituyamos (5) en (3):

$$Y_i = \beta_0 + \beta_1(a + b\hat{X}_{2i} + cW_i) + \beta_2\hat{X}_{2i} + \beta_3W_i + e_i$$

$$Y_i = (\beta_0 + \beta_1a) + (\beta_1b + \beta_2)\hat{X}_{2i} + (\beta_1c + \beta_3)W_i + e_i$$

$$(6) \quad Y_i = \gamma_0 + \gamma_1\hat{X}_{2i} + \gamma_2W_i + e_i$$

donde $\gamma_0 \equiv \beta_0 + \beta_1a$, $\gamma_1 \equiv \beta_1b + \beta_2$ y $\gamma_2 \equiv \beta_1c + \beta_3$.

- Si estimamos (6) por MCO obtendremos

$$\hat{\gamma}_0 = \widehat{\beta_0 + \beta_1a}$$

$$\hat{\gamma}_1 = \widehat{\beta_1b + \beta_2}$$

$$\hat{\gamma}_2 = \widehat{\beta_1c + \beta_3}$$

Apéndice II: Sub-identificación

Econometría

Dr. Christian M.
García-Witulski

MC2E con una X
y múltiples Z's

MC2E con
Múltiples
Regresores
Exógenos

MC2E: Caso
General

Supuestos de VI

Propiedades de
MC2E

Tests sobre la
Validez de los
Instrumentos

Test sobre la
Relevancia de los
Instrumentos

Test sobre la
Exogeneidad de
los Instrumentos

Test sobre la
Endogeneidad de
un Regresor

- Nótese que tenemos un sistema de tres ecuaciones en cuatro incógnitas (β_0 , β_1 , β_2 y β_3).
- Luego, β_0 , β_1 , β_2 y β_3 no están identificados.