

2 – Variable Instrumental (VI)

Evaluación de Impacto

Rachid Laajaj

MECA 4402



Universidad de
los Andes

Facultad
de Economía

Educación
Continua
Vicerrectoría Académica

Universidad de los Andes | Vigilada Mineducación

Reconocimiento como Universidad: Decreto 1297 del 30 de mayo de 1964.

Reconocimiento personería jurídica: Resolución 28 del 23 de febrero de 1949 Minjusticia



El problema de endogeneidad

Partimos de un ejemplo simple: nos interesa saber el efecto causal del riego sobre el valor de la producción agrícola por hectárea (rendimiento).

Uno puede estimarlo con una regresión:

$$\text{Rendimiento}_i = \alpha + \beta_1 \text{riego}_i + X_i \gamma + \varepsilon_i$$

Donde riego_i es una dummy = 1 si el individuo i usa riego

➤ ¿En este caso se puede interpretar el β_1 como el efecto causal del riego sobre el rendimiento? ¿Por qué?

Entonces lo podríamos intentar con efectos fijos de parcela i :

$$\text{Rendimiento}_{it} = \alpha_i + \beta_1 \text{riego}_{it} + X_{it} \gamma + \varepsilon_{it}$$

➤ ¿y ahora, se puede interpretar el β_1 como el efecto causal del riego sobre el rendimiento? ¿Por qué?

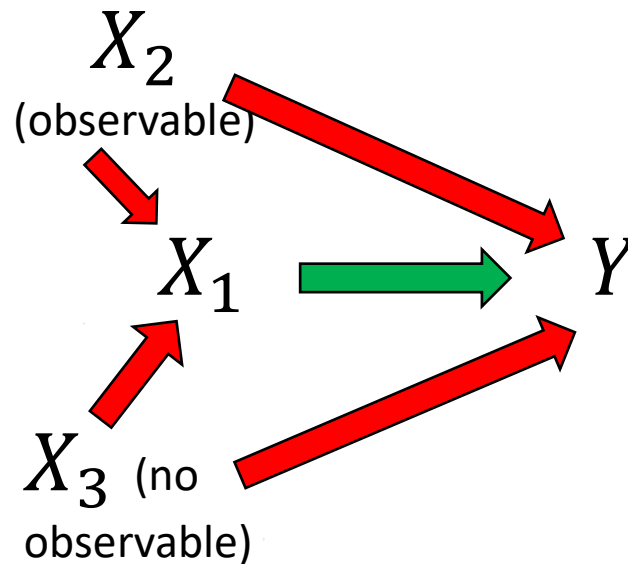
El instrumento busca resolver este problema aprovechando una parte exógena en los cambios de X_{1i}

El Método de Variable instrumental para resolver el problema de endogeneidad

- Existen tres (3) causas típicas de la endogeneidad que generan un sesgo en β_1 : variable omitida, causalidad inversa y errores de medición en la variable de interés.
- Como X_{1i} es una variable endógena, esta correlacionada con ε_i
- El instrumento busca resolver este problema aprovechando una parte exógena en los cambios de X_{1i}
- Pero solo ofrece una solución viable bajo ciertas condiciones
- En esta sección revisamos estas condiciones y como aplicar el método de variable instrumental.

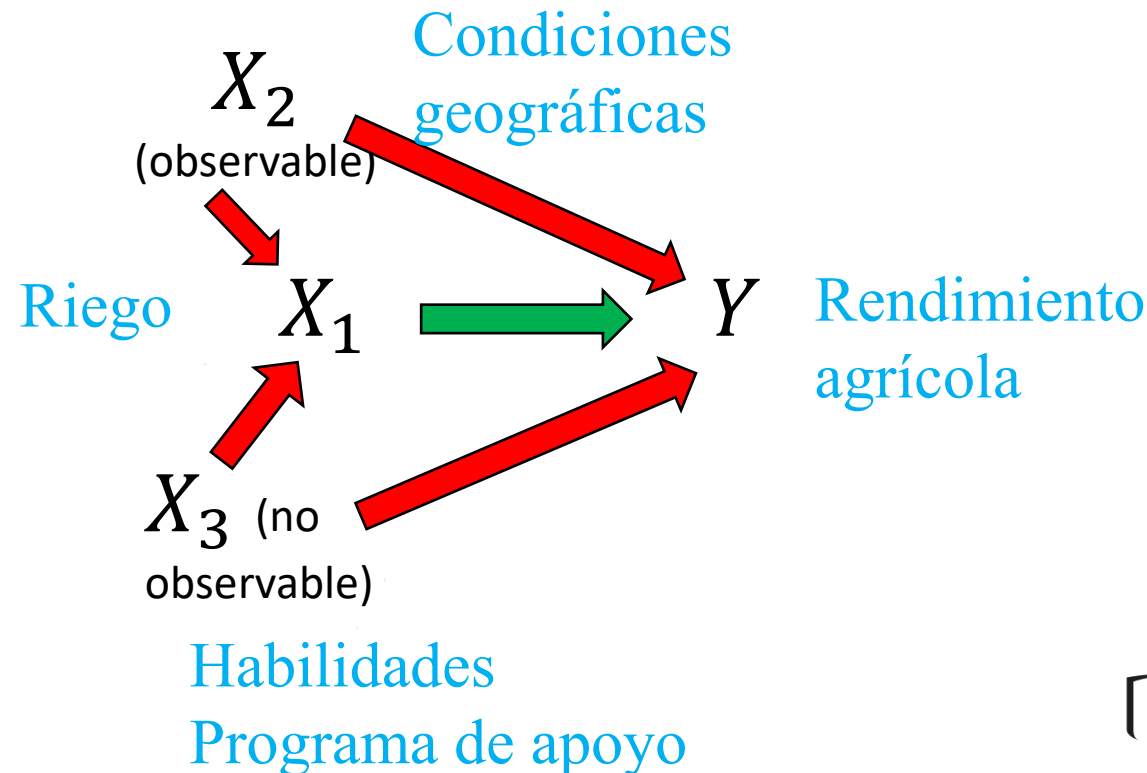
El problema inicial: la Endogeneidad (1)

- Suponemos que nos interesa el efecto de X_1 sobre Y , pero existen otras variables que afectan a X_1 y Y simultáneamente.
- X_2 es observable, entonces se puede controlar por X_2 , pero si X_3 no es observable entonces es una variable omitida y se mantiene un sesgo.



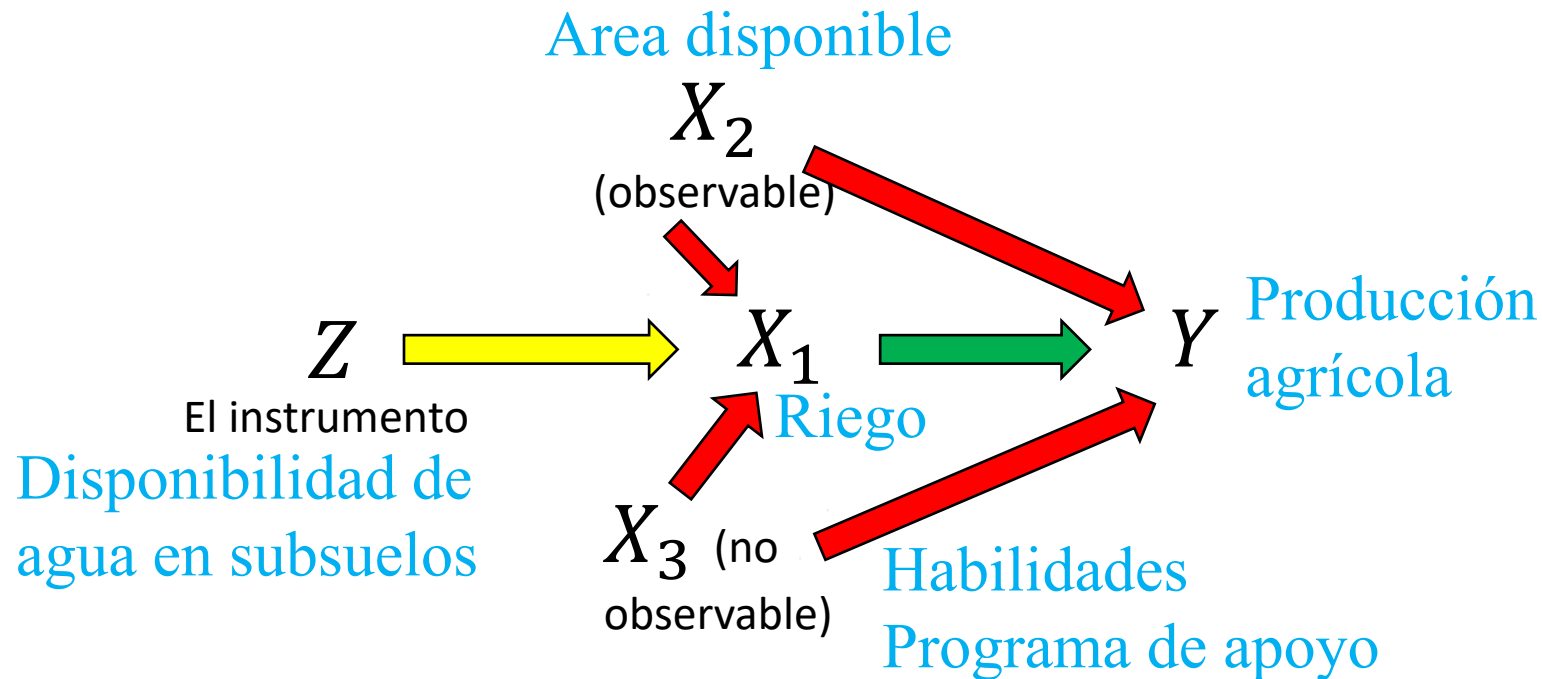
El problema inicial: la Endogeneidad (2)

- Suponemos que nos interesa el efecto de X_1 sobre Y , pero existen otras variables que afectan a X_1 y Y simultáneamente.
- X_2 es observable, entonces se puede controlar por X_2 , pero si X_3 no es observable entonces es una variable omitida y se mantiene un sesgo.



El problema inicial: la Endogeneidad (2)

- El método de VI requiere un instrumento (que llamamos Z), el cual afecta a X_1 y no afecta a Y directamente.
- Su validez tiene 2 condiciones: 1) Relevancia y 2) Exclusión



Los Mínimos Cuadrados en 2 Etapas (1)

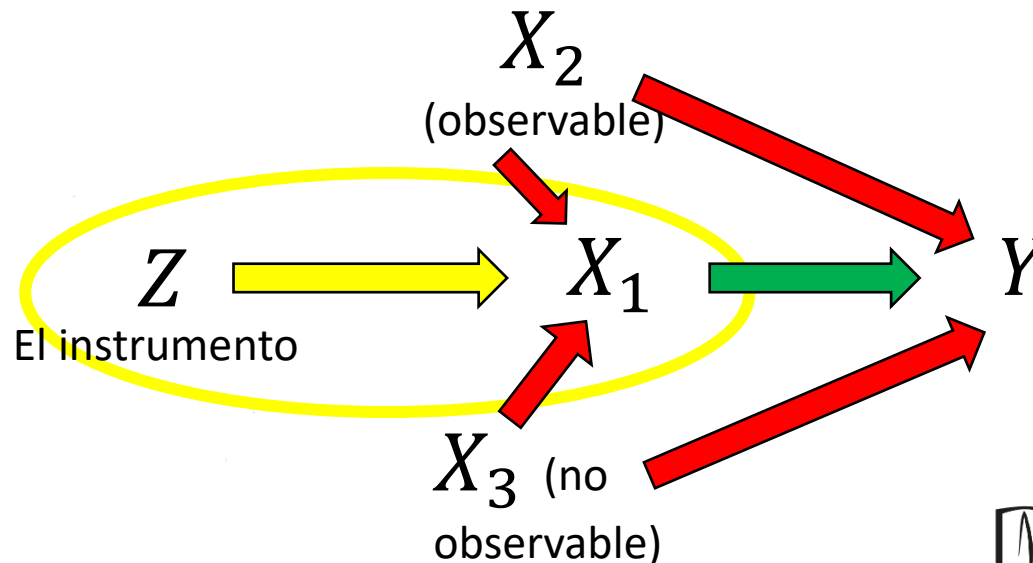
- La estimación se realiza en 2 etapas:
- La primera etapa está representada por la flecha amarilla:

$$X_{1i} = \alpha_0 + \alpha_1 Z_i + \alpha_2 X_{2i} + \varepsilon_i$$

- Una vez estimado, permite obtener la predicción de X_{1i} :

$$\widehat{X}_{1i} = \widehat{\alpha}_0 + \widehat{\alpha}_1 Z_i + \widehat{\alpha}_2 X_{2i}$$

- \widehat{X}_{1i} solo tiene la variación de X_{1i} que proviene de Z_i y de las variables de control, dejando de lado la varianza no explicada que proviene de X_{2i} y otras potenciales variables omitidas.

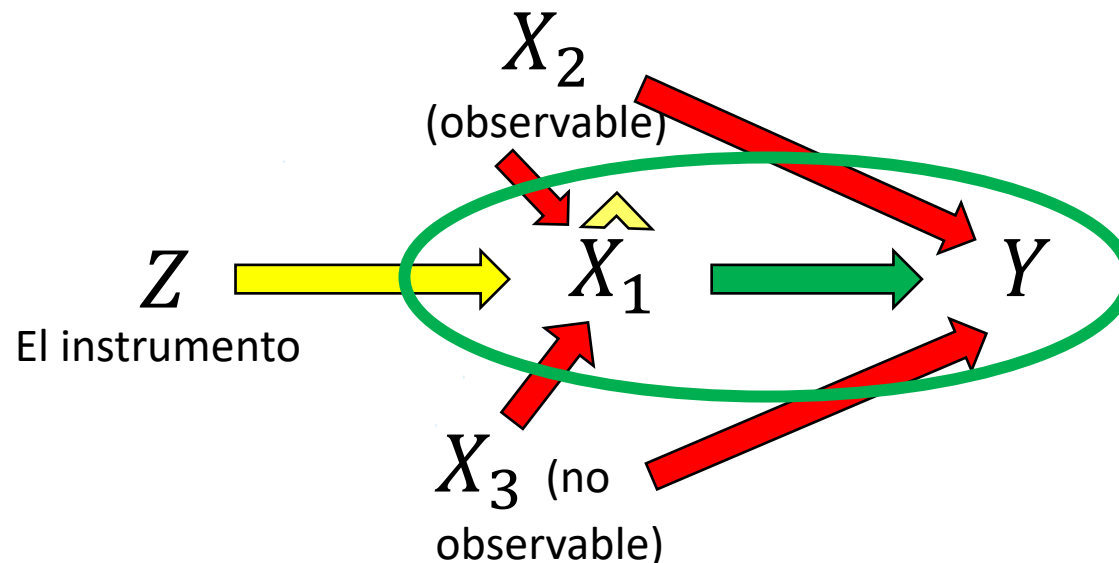


Los Mínimos Cuadrados en 2 etapas (2)

- La segunda etapa está representada por la flecha verde.

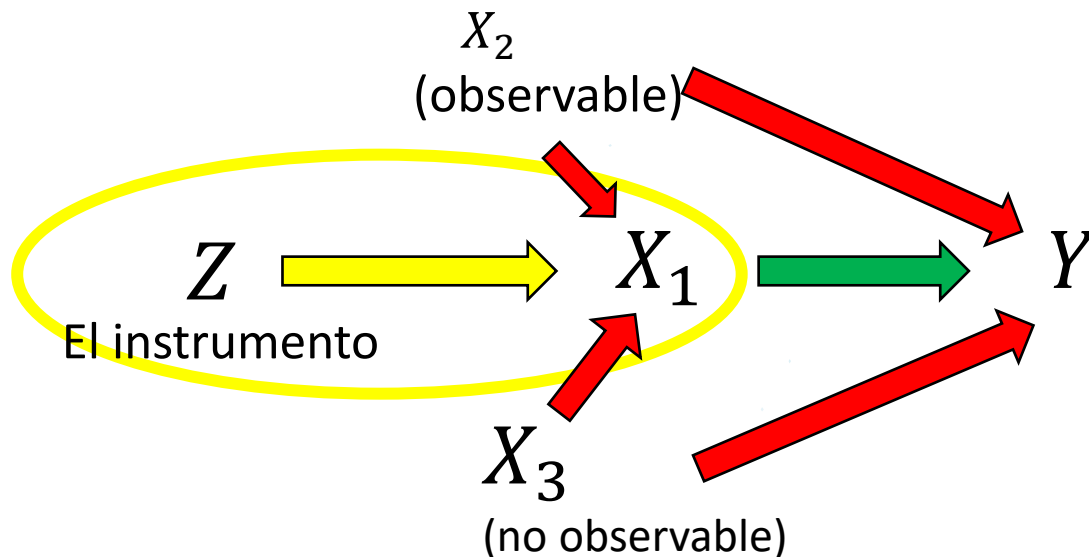
$$Y_i = \beta_0 + \beta_1 \widehat{X}_{1i} + \beta_2 X_{2i} + \varepsilon_i$$

- Es nuestra regresión inicial después de remplazar X_{1i} por \widehat{X}_{1i}
- Ya que \widehat{X}_{1i} fue “purgado” de la parte endógena, β_1 es un estimador insesgado del efecto causal de X_1 sobre Y .
- Pero (repito) su validez tiene 2 condiciones:
 - 1) Relevancia
 - 2) Exclusión



1) La Condición de Relevancia

- Matemáticamente la relevancia requiere $cov(Z_i, X_{1i}) \neq 0$
- O sea, Z debe predecir las variaciones de X de manera suficientemente fuerte.
- Para eso 2 criterios típicos son tener:
 1. un α_1 muy significativo en la primera etapa (y de dirección esperada) y
 2. un estadístico F mayor a 10 en la regresión de la primera etapa



2) La Condición de Exclusión (e Independencia)

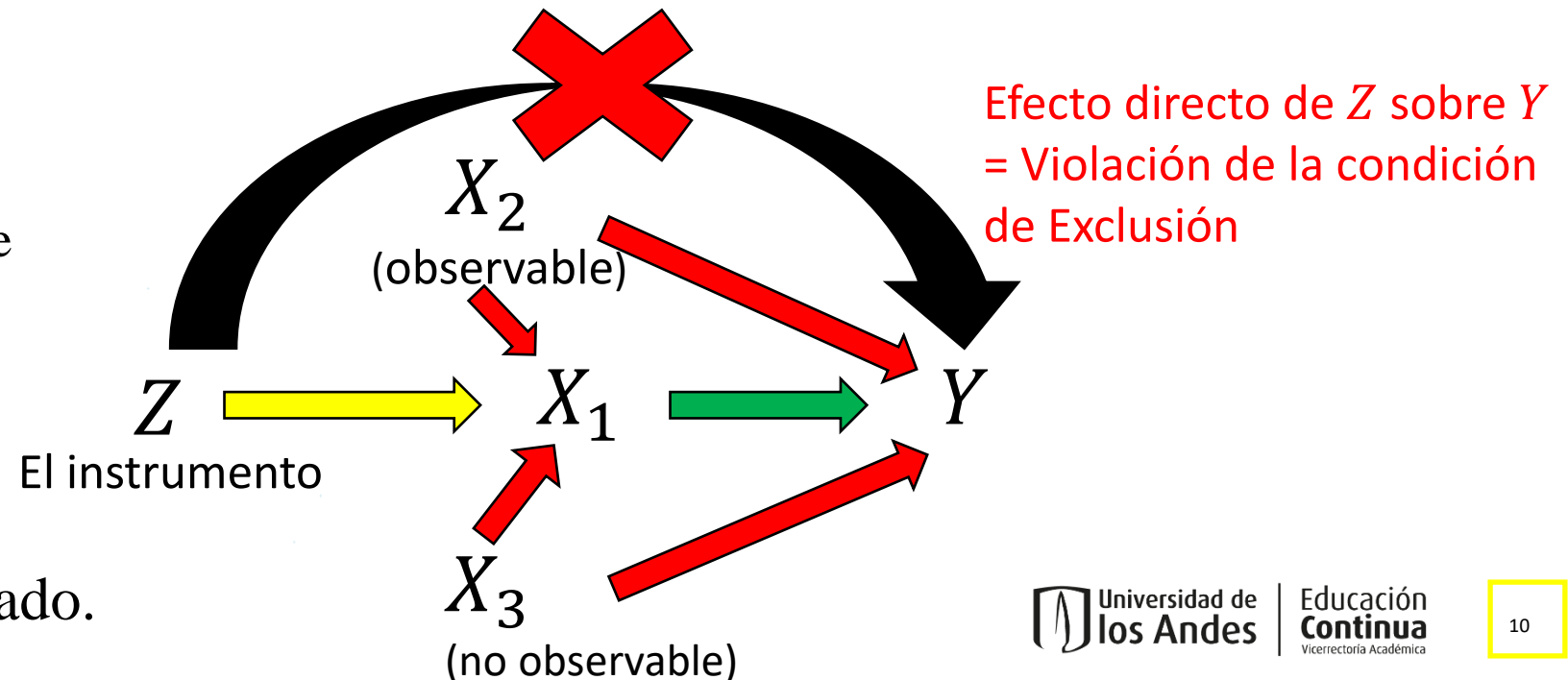
Matemáticamente la Exclusión requiere $cov(Z_i, \varepsilon_i) = 0$

O sea, **Z debe afectar a Y únicamente a través de X_1** (o eventualmente a través de las variables de control).

La **Independencia** = El instrumento no puede estar relacionado con variables relacionadas con Y. (debe ser “tan bueno como aleatorio”)
(puede mencionarse como parte de la exclusión o como un 3er supuesto)

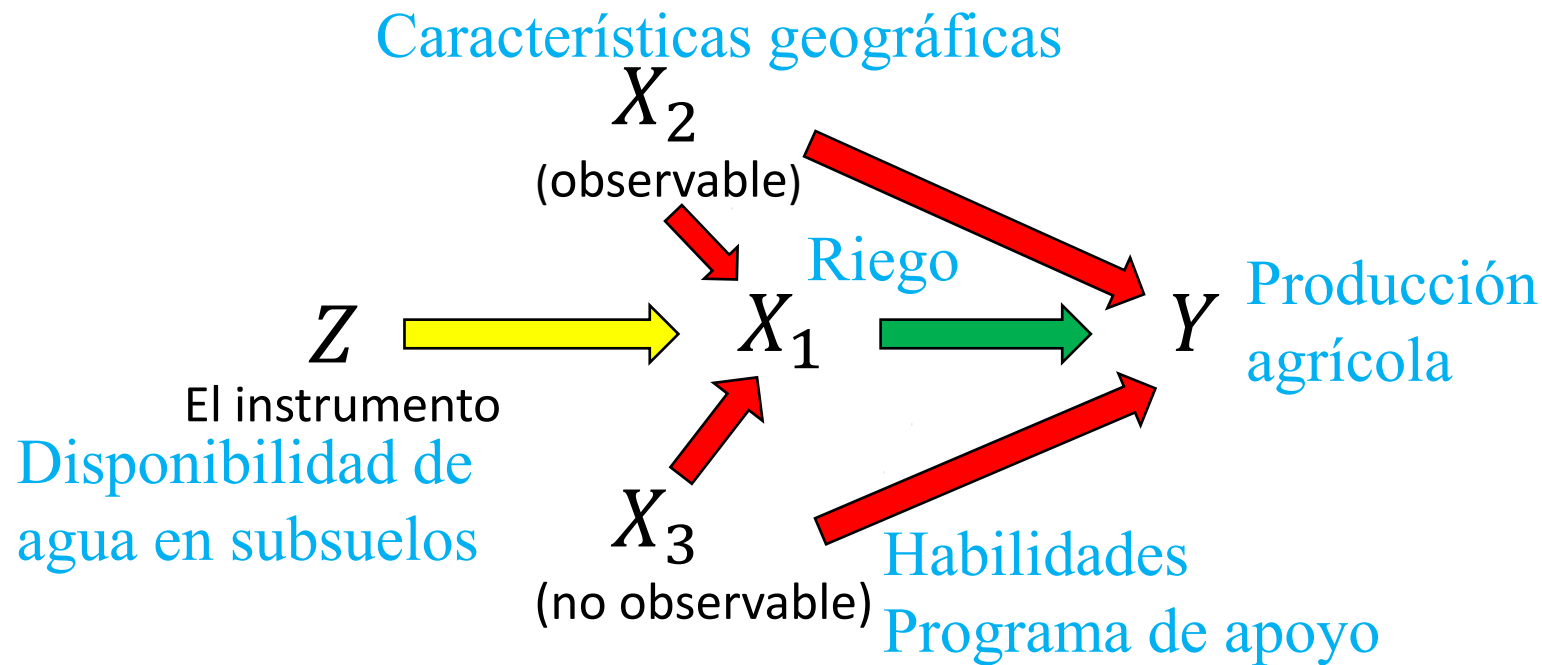
No hay buena prueba de la condición de exclusión,
(no se considera que las pruebas de Sargan y Hansen comprueben realmente la exogeneidad)

Es un supuesto que el investigador debe poder defender de manera convincente en el contexto dado.

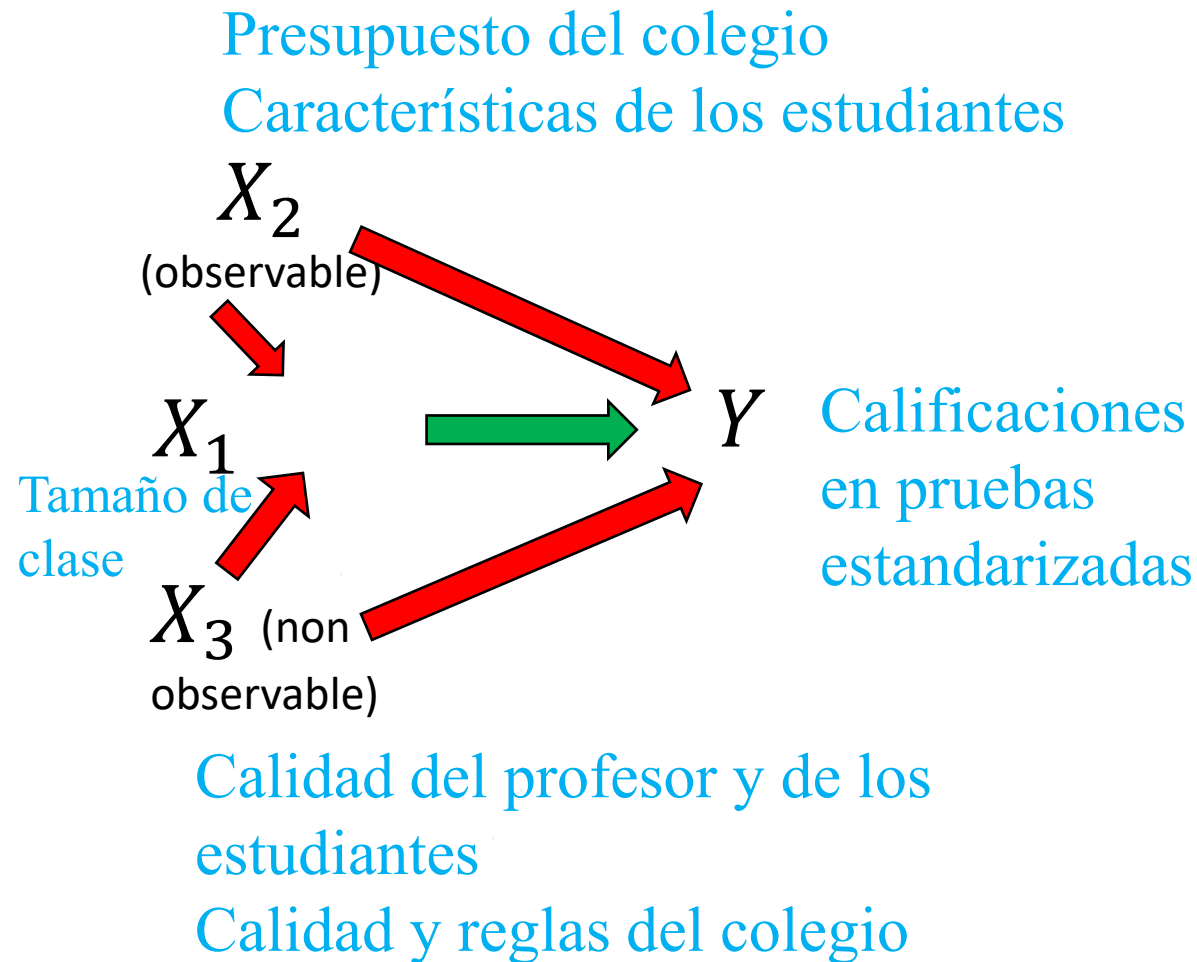


Aplicación de los supuestos a nuestro ejemplo de instrumentación de riego por agua en subsuelos

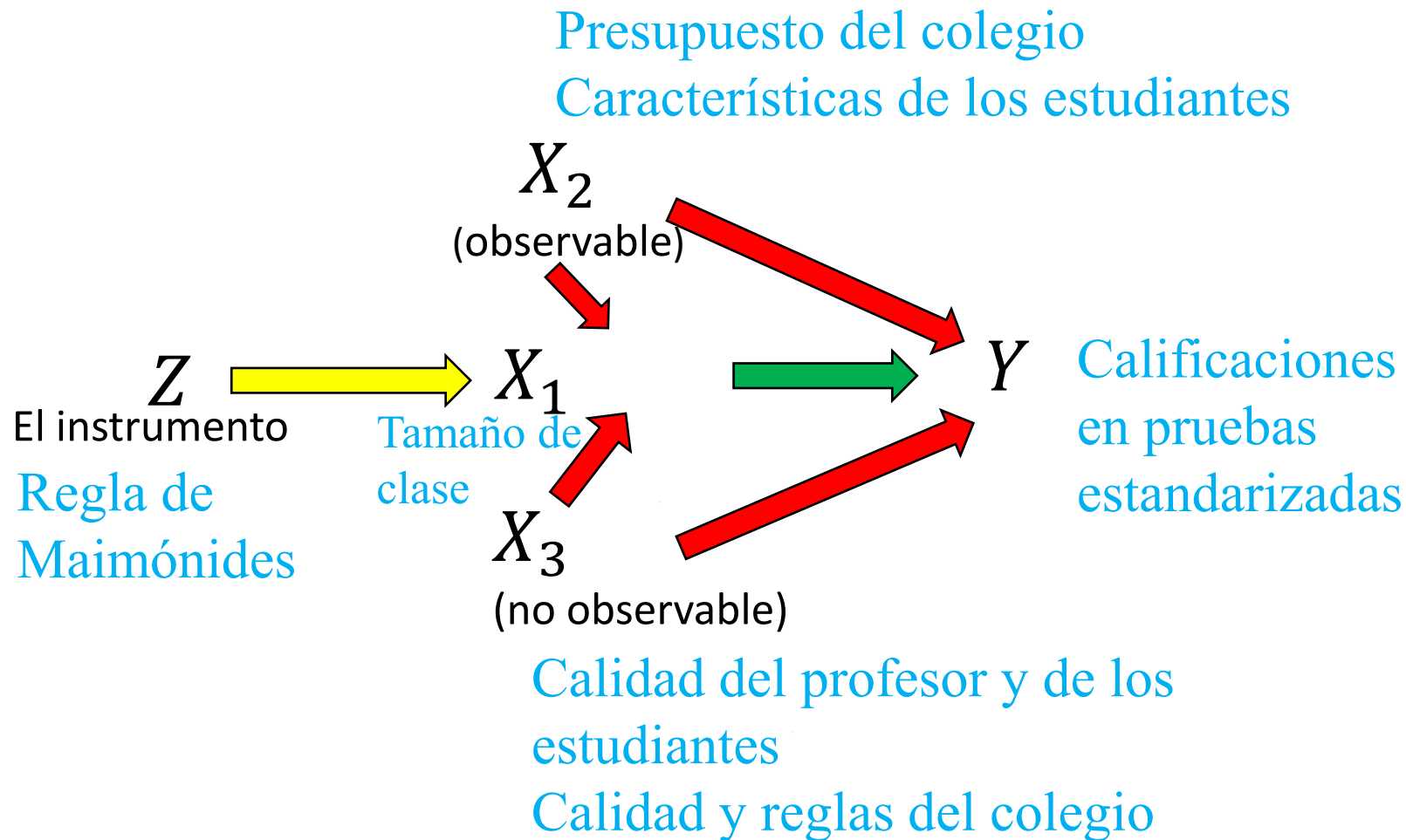
- ¿En nuestro ejemplo anterior de instrumentación de riego por agua en subsuelos, que implican los supuestos de 1) relevancia, 2a) exclusión y 2b) independencia?
- ¿Que tan plausibles les parece?
- ¿Se pueden comprobar? ¿De qué manera?
- ¿Como los controles pueden contribuir a satisfacer el supuesto de independencia?



Otra Aplicación El efecto del tamaño de clase sobre el aprendizaje de los estudiantes



Otra Aplicación El efecto del tamaño de clase sobre el aprendizaje de los estudiantes



La regla de Maimónides para predecir el tamaño de clase

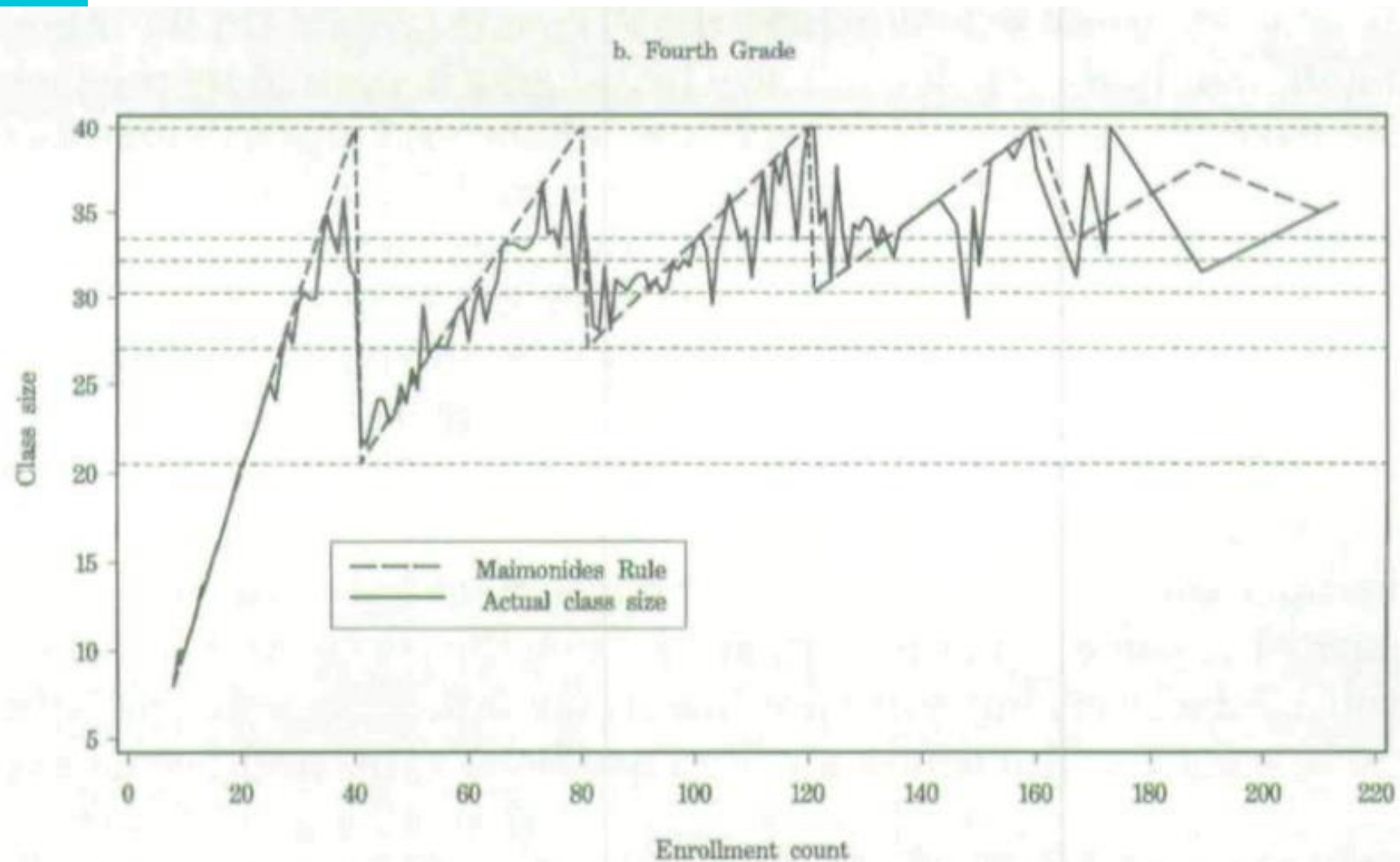


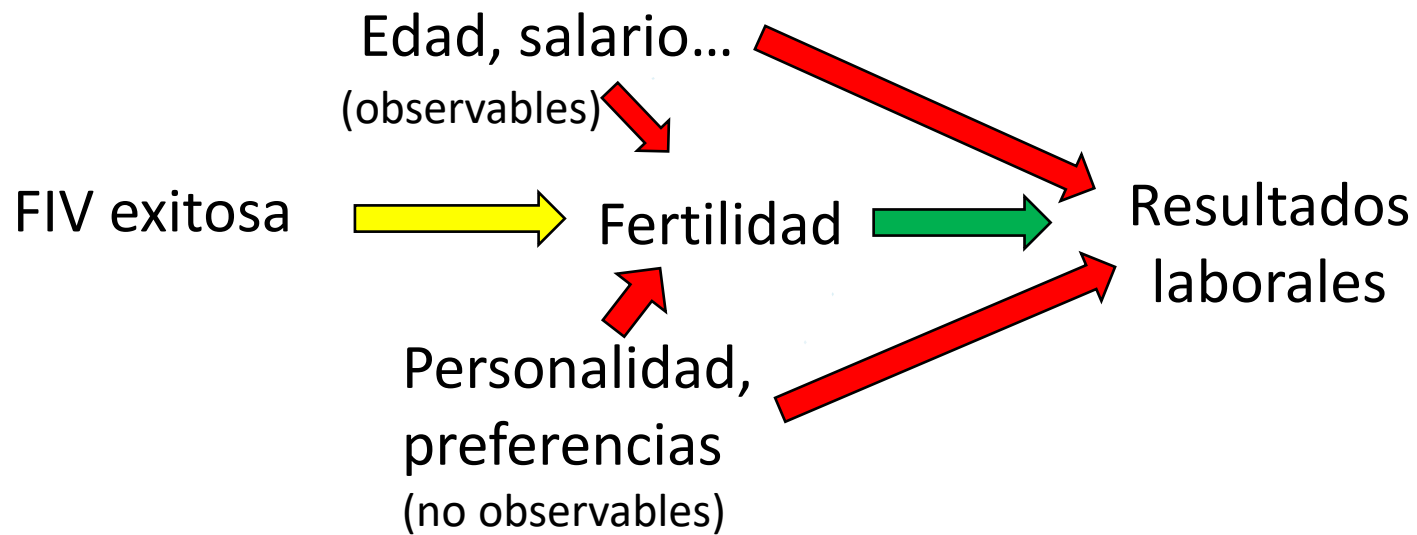
FIGURE I
Class Size in 1991 by Initial Enrollment Count, Actual Average Size and as Predicted by Maimonides' Rule

¿Como encontrar un buen instrumento?

- Puede provenir de:
 - ❖ Condiciones técnicas (como agua subterránea)
 - ❖ Una regla ad hoc (como la regla Maimónides sobre el tamaño de la clase),
 - ❖ Variaciones lejanas (precios o choque de demanda en un mercado global).
- En todo caso, debe provenir de “lejos” para que sea creíble que no tenga muchas relaciones con las otras variables de decisión de la unidad de observación.
- ¡Por esto la evaluación de impacto requiere conocer el contexto y creatividad!

Otro ejemplo concreto: Instrumentación de la Fertilidad con el Éxito de la Fecundación in Vitro (FIV)

- Lundborg et al. quieren entender el efecto de tener hijos sobre salario, empleo, horas de trabajo, depresión y divorcios.
- Pero los hogares que tienen y los que no tienen hijos difieren de varias maneras que pueden estar relacionadas con resultados laborales
- Proponen usar el Éxito de la Fecundación in Vitro (FIV) como instrumento (dentro de hogares que intentaron una FIV).
- En este otro ejemplo, ¿que implican los supuestos de relevancia, exclusión e independencia? ¿Que tan plausibles les parece? ¿Se pueden comprobar? ¿Como?



Un ejemplo concreto: Instrumentación de la Fertilidad con el Éxito de la Fecundación in Vitro (FIV)

- La relevancia es que el FIV aumente la probabilidad de tener al menos 1 hijo. Se puede estimar directamente a través de la primera etapa.
- La exclusión es que el éxito de FIV solo afecte a los resultados laborales a través de su efecto sobre fertilidad (y tampoco por variable omitida). No se puede comprobar econométricamente

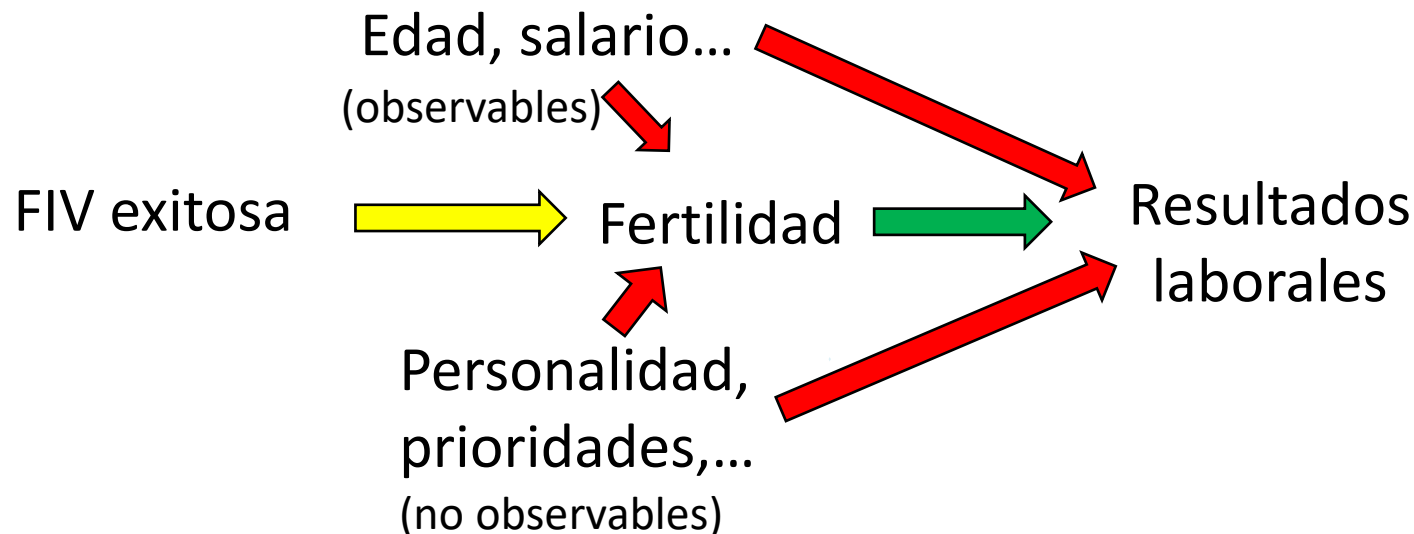


TABLE 3—FERTILITY EFFECTS ON FEMALE LABOR MARKET OUTCOMES:
RESULTS FROM INSTRUMENTAL VARIABLE REGRESSIONS

	Earnings (1)	Positive earnings (2)	Weekly hours (3)	Wages (4)
<i>Panel A. Years 0–1</i>				
Fertility	–70,088 (2,054)	–0.072 (0.006)	–5.911 (0.190)	4.244 (3.235)
Percent impact	–31	–8	–21	2
Observations	18,538	18,538	14,022	14,022
<i>Panel B. Years 2–5</i>				
Fertility	–29,378 (5,285)	–0.041 (0.012)	1.473 (0.355)	–26.851 (4.453)
Percent impact	–12	–5	5	–13
Observations	18,435	18,435	12,332	12,332
<i>Panel C. Years 6–10</i>				
Fertility	–30,675 (10,546)	–0.015 (0.022)	0.487 (0.634)	–25.301 (8.801)
Percent impact	–11	–2	2	–12
Observations	13,779	13,779	9,627	9,627
Baseline mean	223,038	0.90	28.63	183.01
Pretreatment effect	874 (1,811)	0.010 (0.004)	0.519 (0.375)	–0.061 (1.162)

La Forma Reducida

- La mayor parte de los artículos con MC2E muestran también la forma reducida.
- Esta regresión estima directamente el efecto de Z_i sobre Y_i :

$$Y_i = \gamma_0 + \gamma_1 Z_i + \gamma_2 X_{2i} + \varepsilon_i \quad (3)$$

Pero la interpretación tiene en cuenta que el efecto pasa por X_{1i} a pesar de que X_{1i} no esté en la regresión.

Las ventajas de esta estimación son:

- 1) Permitir una visión más transparente, en una sola regresión (si esta regresión corta no da un γ_1 significativo, es poco probable que β_1 salga significativo).
 - 2) A veces la forma reducida tiene una interpretación que es interesante en sí misma.
- ¿Como aplicar esta estimación a nuestro ejemplo de riego instrumentado por la disponibilidad de agua en subsuelos?
 - ¿Y cómo se aplica en nuestro ejemplo de FIV?

TABLE 2—IVF TREATMENT EFFECTS ON FEMALE LABOR OUTCOMES:
RESULTS FROM FIRST-STAGE AND REDUCED-FORM REGRESSIONS

Independent variable	Fertility (1)	Earnings (2)	Positive earnings (3)	Weekly hours (4)	Wages (5)	log wages (6)	Partner earnings (7)	Depression (8)	Divorce (9)
<i>Panel A. Years 0–1</i>									
IVF success	0.694 (0.004)	−48,633 (1,439)	−0.050 (0.004)	−4.036 (0.131)	2.899 (2.212)	0.009 (0.005)	−5,375 (2,470)	−0.013 (0.003)	−0.009 (0.002)
Observations	18,538	18,538	18,538	14,022	14,022	14,022	16,689	18,538	18,538
F-statistic	38,427								
<i>Panel B. Years 2–5</i>									
IVF success	0.320 (0.004)	−9,402 (1,703)	−0.013 (0.004)	0.476 (0.114)	−8.690 (1.437)	−0.034 (0.005)	−3,523 (3,004)	0.002 (0.003)	−0.009 (0.003)
Observations	18,435	18,435	18,435	12,332	12,332	12,332	16,590	18,435	18,435
F-statistic	6,281								
<i>Panel C. Years 6–10</i>									
IVF success	0.227 (0.005)	−6,960 (2,397)	−0.003 (0.005)	0.103 (0.134)	−5.348 (1.861)	−0.021 (0.006)	−5,082 (4,536)	0.003 (0.005)	0.003 (0.005)
Observations	13,779	13,779	13,779	9,627	9,627	9,627	12,367	13,779	13,779
F-statistic	2,273								
Baseline mean	—	223,038	0.90	28.63	183.01	5.16	301,683	0.05	0.05
Pretreatment effect	—	874 (1,811)	0.010 (0.004)	0.519 (0.375)	−0.061 (1.162)	0.001 (0.005)	−1,298 (2,800)	−0.005 (0.003)	0.001 (0.003)

IV con Variables Binarias

Consideramos el caso donde:

- La variable de interés es una variable dicotómica (e.g. $D_i = 1$ si está afectado por el tratamiento)

$$D_i = \alpha_0 + \alpha_1 Z_i + \alpha_2 X_{2i} + \epsilon_i \quad (1)$$

$$y_i = \beta_0 + \beta_1 \widehat{D}_i + \beta_2 X_i + \varepsilon_i \quad (2)$$

- El instrumento también es una dummy:
tener $Z_i = 1$ aumenta la probabilidad de estar tratado.

Todo lo que hemos mencionado antes aplica.

También nos permite tener grupos y relacionarlo con nuestro vocabulario de contrafactual.

Los “Cumplidos” y otros grupos

Ahora podemos definir los grupos siguientes:

- 1) Los siempre tratados: $D_i = 1|Z_i = 1$ y $D_i = 1|Z_i = 0$
- 2) Los nunca tratados: $D_i = 0|Z_i = 1$ y $D_i = 0|Z_i = 0$
- 3) Los “Cumplidos”, solo participan si $Z_i = 1$: $D_i = 1|Z_i = 1$ y $D_i = 0|Z_i = 0$
- 4) Los “desafiantes”, solo participan si $Z_i = 0$: $D_i = 0|Z_i = 1$ y $D_i = 1|Z_i = 0$

En nuestro ejemplo los Cumplidos son los que tienen hijos solamente si la 1ª FIV fue exitosa (que cambian su comportamiento con el instrumento)

	$Z_i = 0$	$Z_i = 1$
$D_i = 0$	Los nunca tratados Los Cumplidos	Los nunca tratados Los Desafiantes
$D_i = 1$	Los siempre tratados Los Desafiantes	Los siempre tratados Los Cumplidos

La primera etapa nos dice que pasar de $Z_i = 0$ a $Z_i = 1$ aumenta la proporción de tratados de α_1

Entonces α_1 representa la proporción de Cumplidos

Los “Cumplidos” y otros grupos

Ahora podemos definir los grupos siguientes:

1) Los siempre tratados: $D_i = 1|Z_i = 1$ y $D_i = 1|Z_i = 0$

2) Los nunca tratados: $D_i = 0|Z_i = 1$ y $D_i = 0|Z_i = 0$

3) Los “Cumplidos”, solo participan si $Z_i = 1$: $D_i = 1|Z_i = 1$ y $D_i = 0|Z_i = 0$

4) Los “desafiantes”, solo participan si $Z_i = 0$: $D_i = 0|Z_i = 1$ y $D_i = 1|Z_i = 0$

En nuestro ejemplo los Cumplidos son los que tienen hijos solamente si la 1ª FIV fue exitosa (que cambian su comportamiento con el instrumento)

	$Z_i = 0$	$Z_i = 1$
$D_i = 0$	Los nunca tratados Los Cumplidos	Los nunca tratados Los Desafiantes
$D_i = 1$	Los siempre tratados Los Desafiantes	Los siempre tratados Los Cumplidos

La primera etapa nos dice que pasar de $Z_i = 0$ a $Z_i = 1$ aumenta la proporción de tratados de α_1

Entonces α_1 representa la proporción de Cumplidos

Los “Cumplidos” y otros grupos

Ahora podemos definir los grupos siguientes:

- 1) Los siempre tratados: $D_i = 1|Z_i = 1$ y $D_i = 1|Z_i = 0$
- 2) Los nunca tratados: $D_i = 0|Z_i = 1$ y $D_i = 0|Z_i = 0$
- 3) Los “Cumplidos”, solamente participan si $Z_i = 1$: $D_i = 1|Z_i = 1$ y $D_i = 0|Z_i = 0$
- 4) Los “desafiantes”, solo participan si $Z_i = 0$: $D_i = 0|Z_i = 1$ y $D_i = 1|Z_i = 0$

En nuestro ejemplo los Cumplidos son los que tienen hijos solamente si la 1ª FIV fue exitosa (que cambian su comportamiento con el instrumento)

	$Z_i = 0$	$Z_i = 1$
$D_i = 0$	Los nunca tratados Los Cumplidos	Los nunca tratados Los Desafiantes
$D_i = 1$	Los siempre tratados Los Desafiantes	Los siempre tratados Los Cumplidos

La primera etapa nos dice que pasar de $Z_i = 0$ a $Z_i = 1$ aumenta la proporción de tratados de α_1

Entonces α_1 representa la proporción de Cumplidos

Los “Cumplidos” y otros grupos

Ahora podemos definir los grupos siguientes:

- 1) Los siempre tratados: $D_i = 1|Z_i = 1$ y $D_i = 1|Z_i = 0$
- 2) Los nunca tratados: $D_i = 0|Z_i = 1$ y $D_i = 0|Z_i = 0$
- 3) Los “Cumplidos”, solo participan si $Z_i = 1$: $D_i = 1|Z_i = 1$ y $D_i = 0|Z_i = 0$
- 4) Los “desafiantes”, solamente participan si $Z_i = 0$: $D_i = 0|Z_i = 1$ y $D_i = 1|Z_i = 0$

	$Z_i = 0$	$Z_i = 1$
$D_i = 0$	Los nunca tratados Los Cumplidos	Los nunca tratados Los Desafiantes
$D_i = 1$	Los siempre tratados Los Desafiantes	Los siempre tratados Los Cumplidos

**Pero en general
suponemos que
no hay desafiantes**

Los “Cumplidos” y otros grupos

Ahora podemos definir los grupos siguientes:

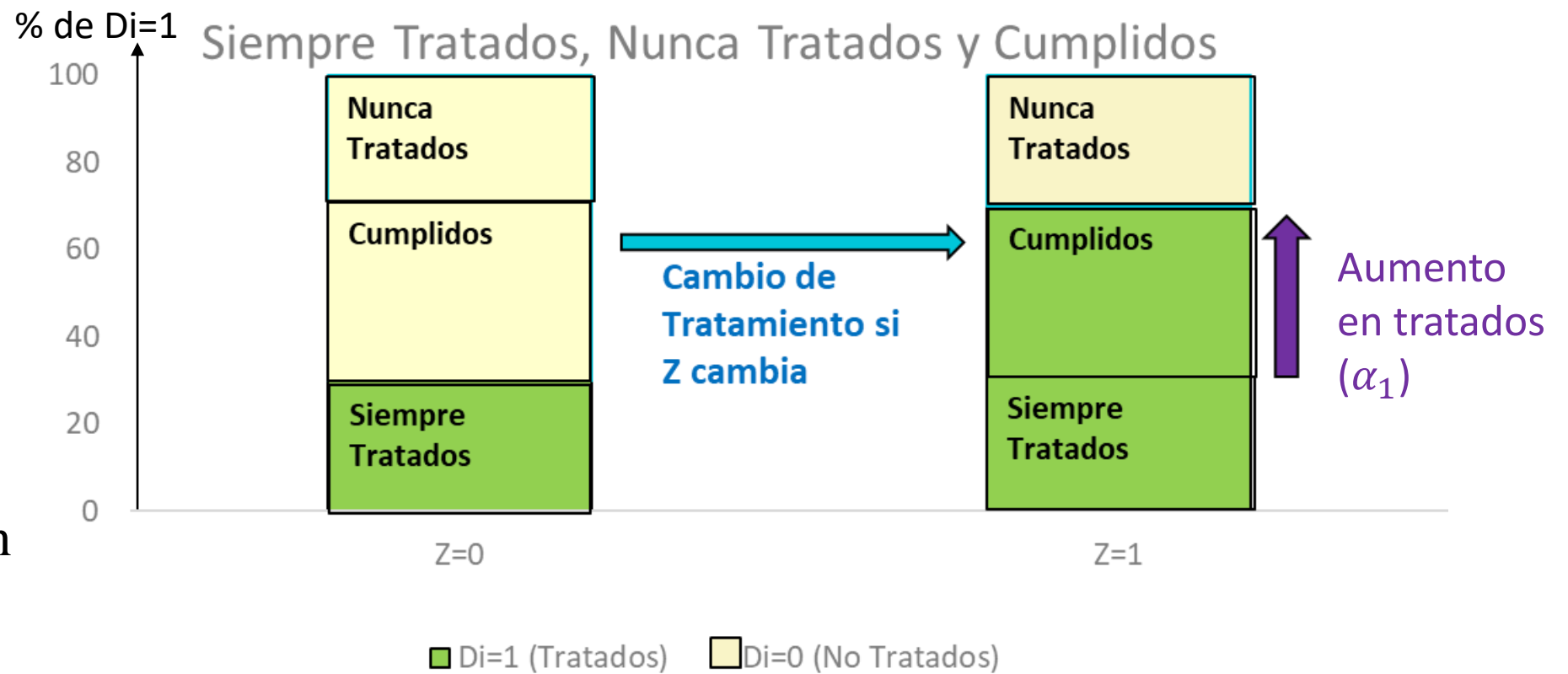
- 1) Los siempre tratados: $D_i = 1|Z_i = 1$ y $D_i = 1|Z_i = 0$
- 2) Los nunca tratados: $D_i = 0|Z_i = 1$ y $D_i = 0|Z_i = 0$
- 3) Los “Cumplidos”, solamente participan si $Z_i = 1$: $D_i = 1|Z_i = 1$ y $D_i = 0|Z_i = 0$
- 4) Los “desafiantes”, solamente participan si $Z_i = 0$: $D_i = 0|Z_i = 1$ y $D_i = 1|Z_i = 0$

	$Z_i = 0$	$Z_i = 1$
$D_i = 0$	Los nunca tratados Los Cumplidos	Los nunca tratados
$D_i = 1$	Los siempre tratados	Los siempre tratados Los Cumplidos

Suponemos que no
hay desafiantes

- Para practicar: en el ejemplo de FIV, ¿quiénes son los cumplidos?
¿ los nunca tratados? ¿ Los siempre tratados?
- En el ejemplo de agua subterránea, ¿quiénes son los cumplidos?
¿ los nunca tratados? ¿ Los siempre tratados?

Los únicos que pueden tener un contrafactual (observable) son los Cumplidos ya que son los únicos que cambian de comportamiento con el instrumento.



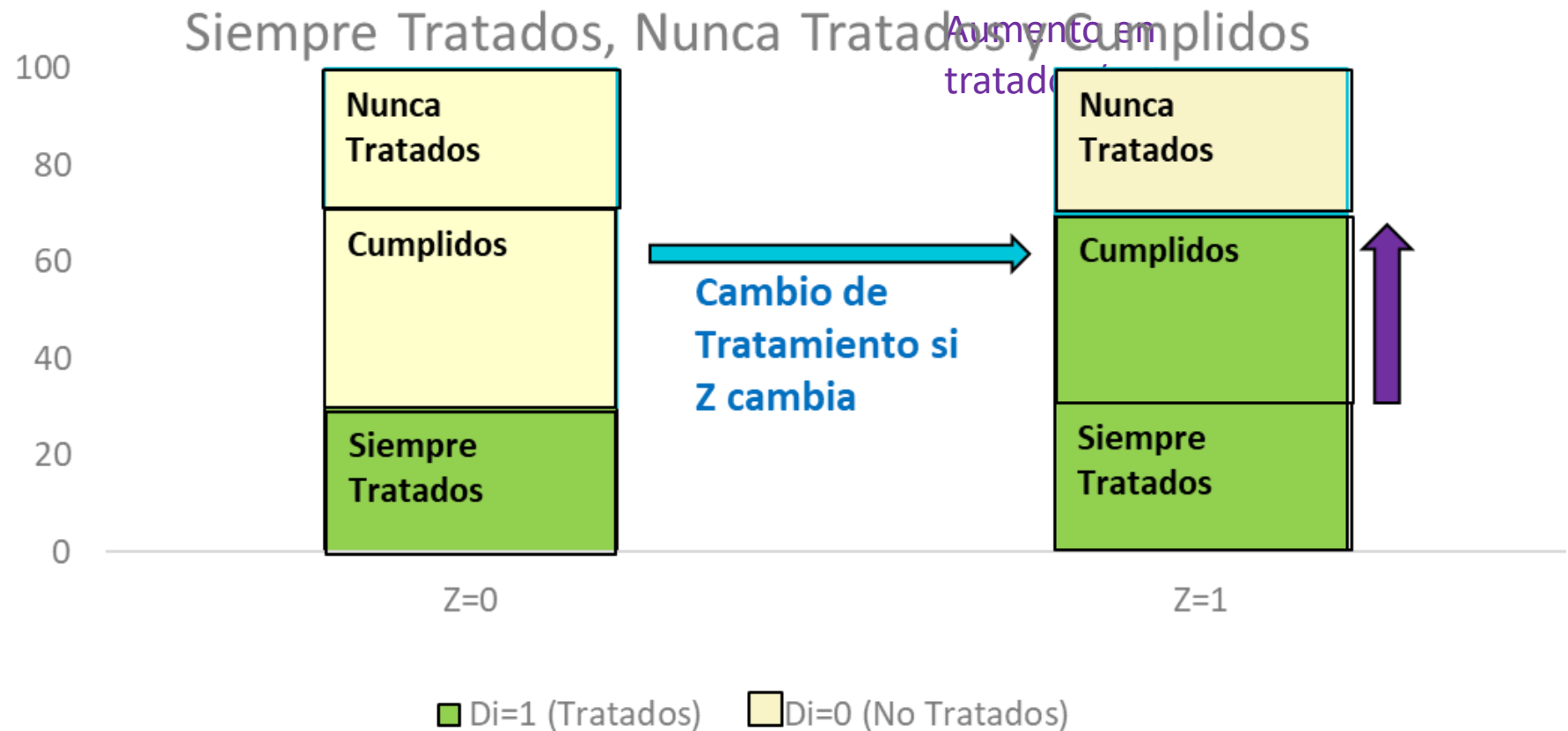
En la primera etapa :

$$D_i = \alpha_0 + \alpha_1 Z_i + \alpha_2 X_{2i} + \epsilon_i$$

α_1 nos dice el efecto de pasar de $Z_i = 0$ a $Z_i = 1$ sobre la proporción de tratados de α_1

Entonces α_1 representa la proporción de Cumplidos

Independent variable	Fertility (1)
<i>Panel A. Years 0–1</i>	
IVF success	0.694 (0.004)
Observations	18,538
F-statistic	38,427
<i>Panel B. Years 2–5</i>	
IVF success	0.320 (0.004)
Observations	18,435
F-statistic	6,281
<i>Panel C. Years 6–10</i>	
IVF success	0.227 (0.005)
Observations	13,779
F-statistic	2,273



➤ En nuestro ejemplo de FIV,
¿Cuál es la proporción de Cumplidos en cada periodo?

¿ Como evoluciona y por qué?

El Estimador de Wald (y intuición del método de VI)

- Si pudiéramos observar quienes son los Cumplidos, nos gustaría comparar $E[Y_i|D_i = 1] - E[Y_i|D_i = 0]$ solo dentro de los Cumplidos
- Pero, observando solo D_i y Z_i , no sabemos cuáles son los Cumplidos .
- ¡La magia del método de VI es que permite estimar el efecto sobre los Cumplidos sin saber quiénes son los Cumplidos!
- La lógica es comparar $E[Y_i|Z_i = 1] - E[Y_i|Z_i = 0]$ pero ajustando por la proporción de Cumplidos (que cambió su tratamiento debido al instrumento).

El Estimador de Wald (y intuición del método de VI)

■ Llamamos Y_{i1} al resultado potencial esperado de un individuo tratado ($D_i = 1$), y Y_{i0} el resultado potencial esperado de un individuo no tratado ($D_i = 0$).

■ Cuando el instrumento pasa de $Z_i = 0$ a $Z_i = 1$, dos cosas ocurren:

- La 1ª etapa ($D_i = \alpha_0 + \alpha_1 Z_i + \alpha_2 X_{2i} + \epsilon_i$) nos dice que la proporción de tratados aumenta en α_1 (la proporción de Cumplidos)

- La forma reducida ($Y_i = \gamma_0 + \gamma_1 Z_i + \gamma_2 X_{2i} + \epsilon_i$) nos dice que

$$E[Y_i | Z_i = 1] - E[Y_i | Z_i = 0] = \gamma_1$$

■ El estimador de Wald del efecto del tratamiento esta dado por:

$$\beta_{1wald} = \frac{\gamma_1}{\alpha_1}$$

(β_{1wald} es exactamente igual a β_1 de la 2da etapa sin controles, pero con controles, nos da una aproximación)

➤ Aplicando la formula a nuestro ejemplo de FIV, el efecto sobre *Earnings* en 2-5 años:

$$\hat{\alpha}_1 = \underline{\hspace{1cm}} \quad ; \quad \hat{\gamma}_1 = \underline{\hspace{1cm}} \quad ; \quad \widehat{\beta_{1wald}} = \underline{\hspace{1cm}} = \underline{\hspace{1cm}}$$

➤ Y como esto compara con el $\hat{\beta}_1$ estimado por MC2E de la primera tabla?

TABLE 2—IVF TREATMENT EFFECTS ON FEMALE LABOR OUTCOMES:
RESULTS FROM FIRST-STAGE AND REDUCED-FORM REGRESSIONS

Independent variable	Fertility (1)	Earnings (2)	Positive earnings (3)	Weekly hours (4)	Wages (5)	log wages (6)	Partner earnings (7)	Depression (8)	Divorce (9)
<i>Panel A. Years 0–1</i>									
IVF success	0.694 (0.004)	−48,633 (1,439)	−0.050 (0.004)	−4.036 (0.131)	2.899 (2.212)	0.009 (0.005)	−5,375 (2,470)	−0.013 (0.003)	−0.009 (0.002)
Observations	18,538	18,538	18,538	14,022	14,022	14,022	16,689	18,538	18,538
F-statistic	38,427								
<i>Panel B. Years 2–5</i>									
IVF success	<u>0.320</u> (0.004)	<u>−9,402</u> (1,703)	−0.013 (0.004)	0.476 (0.114)	−8.690 (1.437)	−0.034 (0.005)	−3,523 (3,004)	0.002 (0.003)	−0.009 (0.003)
Observations	18,435	18,435	18,435	12,332	12,332	12,332	16,590	18,435	18,435
F-statistic	6,281								

- Aplicando la formula a nuestro ejemplo de FIV, el efecto sobre *earnings* en 2-5 años:

$$\widehat{\alpha}_1 = 0.32 \quad ; \quad \widehat{\gamma}_1 = -9,402 \quad ; \quad \widehat{\beta}_{1wald} = \frac{-9,402}{0.32} = -29,381$$

- En la primera tabla, podríamos comprobar que $\widehat{\beta}_1 = -29,378$, está muy cercano.

TABLE 3—FERTILITY EFFECTS ON FEMALE LABOR MARKET OUTCOMES:
RESULTS FROM INSTRUMENTAL VARIABLE REGRESSIONS

	Earnings (1)	Positive earnings (2)	Weekly hours (3)	Wages (4)
<i>Panel A. Years 0–1</i>				
Fertility	−70,088 (2,054)	−0.072 (0.006)	−5.911 (0.190)	4.244 (3.235)
Percent impact	−31	−8	−21	2
Observations	18,538	18,538	14,022	14,022
<i>Panel B. Years 2–5</i>				
Fertility	<u>−29,378</u> (5,285)	−0.041 (0.012)	1.473 (0.355)	−26.851 (4.453)
Percent impact	−12	−5	5	−13
Observations	18,435	18,435	12,332	12,332

Intuición atrás del Estimador de Wald (y intuición del método de VI)

■ Simplificando nuestro ejemplo de FIV, el efecto sobre ingresos en 2-5 años:

$$\widehat{\alpha}_1 = 1/3 \quad ; \quad \widehat{\beta}_1 = -3000 \quad ; \quad \widehat{\gamma}_1 = -1000 \quad ; \quad \widehat{\beta}_{1wald} = \frac{-1000}{1/3} = -3000$$

■ Si solo 1/3 de la población cambio su fertilidad gracias al éxito de la 1ª FIV, para observar un efecto sobre el salario de -1000 en el grupo completo, se requiere un cambio 3 veces más fuerte dentro de los Cumplidos

■ $\beta_{1wald} = \frac{\gamma_1}{\alpha_1}$ se puede reescribir: $\gamma_1 = \alpha_1 * \beta_{1wald}$: $-1000 = 1/3 * -3000 (+ 2/3 * 0)$

■ El efecto promedio de toda la población es igual a la proporción de Cumplidos multiplicado por el efecto sobre los Cumplidos.

■ Cuando la relevancia es baja, el efecto promedio de toda la población se vuelve muy bajo y difícil de detectar (poder insuficiente)

Con nuestro primer ejemplo, supone que:

- donde no hay agua subterránea, 10% tienen riego y la producción es de 1,000.
- donde hay agua subterránea, 60% tiene riego y la producción es de 1,300.

➤ ¿Qué valor nos da el estimador de Wald?

$$\widehat{\alpha}_1 = \underline{\hspace{1cm}} \quad ; \quad \widehat{\gamma}_1 = \underline{\hspace{1cm}} \quad ; \quad \widehat{\beta}_{1wald} = \underline{\hspace{1cm}} = \underline{\hspace{1cm}}$$

➤ Explique la intuición atrás del resultado

Efectos Heterogéneos y LATE

Ahora consideremos explícitamente que el efecto del tratamiento sea heterogéneo, es decir que $\tau_i = (Y_i|D_i = 1) - (Y_i|D_i = 0)$ sea diferente entre los individuos.

Dado que el método de VI aprovecha una variación que viene del cambio en los Cumplidos, el estimador β_1 que obtenemos proviene de:

$$\tau_{LATE} = E[(Y_i|D_i = 1) - (Y_i|D_i = 0)] \mid i \in \text{Cumplidos}$$

Este LATE se refiere a Local Average Treatment Effect. Se trata de un efecto “local” en el sentido que afecta una población en particular.

Nada nos garantiza que τ_{ATE} (el efecto promedio en toda la población) sea $= \tau_{LATE}$

Qué tan similares pueden ser los estimadores depende de qué tan diferentes pueden ser los Cumplidos del resto de la población, de una manera que puede afectar el efecto del tratamiento (muy específico a cada contexto).

Si usa un método de VI, se recomienda discutir este tema de validez externa

En nuestro ejemplo de FIV, ¿el τ_{LATE} podría ser diferente del efecto sobre el resto de la población? ¿Por qué?

Los MC2E con más de 1 Instrumento

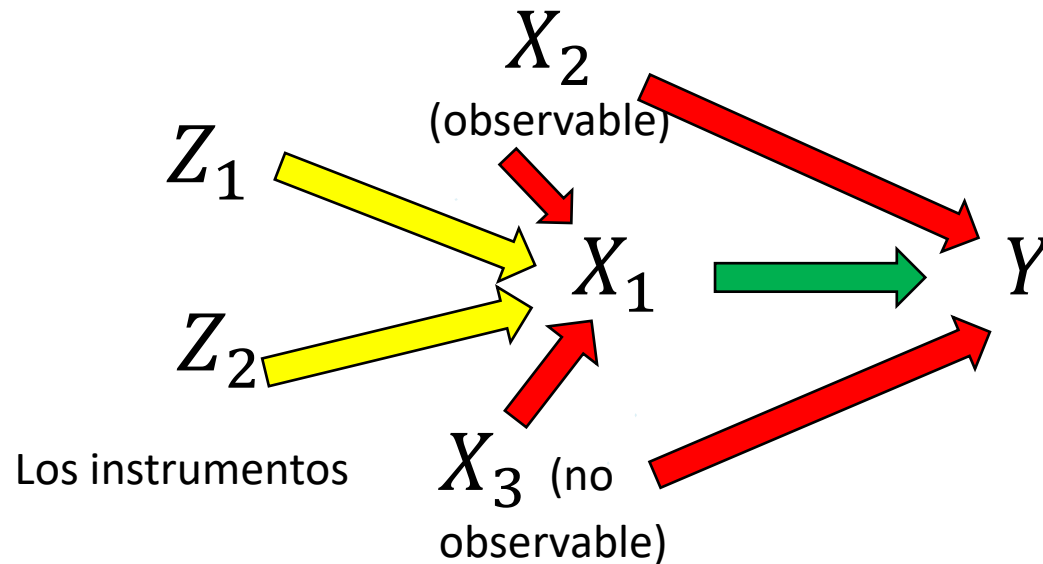
Por ejemplo, con 2 instrumentos:

La primera etapa:

$$X_{1i} = \alpha_0 + \alpha_1 Z_{1i} + \alpha_2 Z_{2i} + \alpha_3 X_{2i} + \varepsilon_i$$

Permite obtener la predicción de X_{1i} e insertarlo en la 2da etapa:

$$Y_i = \beta_0 + \beta_1 \widehat{X}_{1i} + \beta_2 X_{2i} + \varepsilon_i$$



Los MC2E con más de 1 Instrumento y más de una variable endógena

Por ejemplo, con 2 instrumentos y 2 variables endógenas:

La primera etapa:

$$X_{1i} = \alpha_0 + \alpha_1 Z_{1i} + \alpha_2 Z_{2i} + \alpha_3 X_{3i} + \varepsilon_i$$

$$X_{2i} = \delta_0 + \delta_1 Z_{1i} + \delta_2 Z_{2i} + \delta_3 X_{3i} + \varepsilon_i$$

Permite obtener las predicciones de X_{1i} y X_{2i} e insertarlo en la 2da etapa:

$$Y_i = \beta_0 + \beta_1 \widehat{X}_{1i} + \beta_2 \widehat{X}_{2i} + \beta_3 X_{3i} + \varepsilon_i$$

Debe regresar cada variable endógena sobre todos los instrumentos y controles.

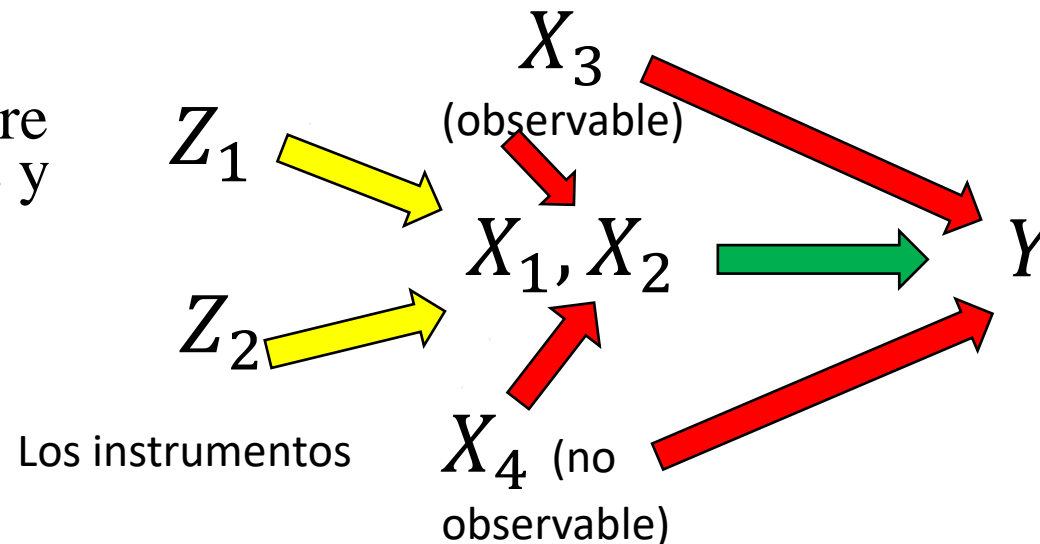


TABLE 4—FERTILITY EFFECTS ON JOB CHANGES AND JOB CHARACTERISTICS:
RESULTS FROM INSTRUMENTAL VARIABLE REGRESSIONS

	Occ. change (1)	Firm change (2)	Occ. earnings (3)	Occ. (log) wages (4)	Occ. hours (5)	Firm earnings (6)	Firm (log) wages (7)	Firm hours (8)	Firm gender ratio (9)	Public sector (10)	Distance (in km) (11)
<i>Panel A. Years 0–1</i>											
Fertility	0.017 (0.010)	0.040 (0.010)	−3,687 (1,090)	0.001 (0.002)	−0.346 (0.080)	−1,197 (1,003)	−0.002 (0.003)	−0.050 (0.061)	−0.008 (0.003)	−0.005 (0.005)	1.379 (0.568)
Observations	17,941	18,194	18,019	18,019	18,019	18,264	18,195	18,194	18,280	17,914	14,440
<i>Panel B. Years 2–5</i>											
Fertility	0.046 <u>(0.028)</u>	0.038 (0.025)	−4,881 (2,679)	−0.005 (0.006)	−0.262 (0.179)	−5,761 (2,528)	−0.017 (0.008)	0.213 (0.034)	−0.004 (0.008)	0.010 (0.013)	−3.667 <u>(1.083)</u>
Observations	14,166	18,058	17,934	17,934	17,934	18,169	18,072	18,084	18,188	17,824	10,618
<i>Panel C. Years 6–10</i>											
Fertility	0.010 (0.044)	−0.013 (0.043)	−5,030 (4,962)	−0.005 (0.011)	−0.383 (0.315)	−7,063 (4,839)	−0.022 (0.014)	−0.301 (0.366)	0.001 (0.016)	0.023 (0.026)	−5.530 (2.145)
Observations	11,680	11,782	13,470	13,470	13,470	13,595	13,522	10,974	13,616	13,337	5,173
Baseline mean	0.54	0.57	221,346	5.24	22.33	244,288	5.20	25.58	0.63	0.55	12.58
Pretreatment effect	−0.005 (0.008)	−0.007 (0.008)	−270 (990)	−0.001 (0.002)	0.125 (0.070)	955 (993)	−0.004 (0.003)	0.022 (0.064)	−0.002 (0.003)	−0.009 (0.008)	0.639 (0.360)

Examinar
los
Canales

Checklist para su Artículo

- 1) Aclarar cuál es el problema de endogeneidad inicial que quiere resolver
- 2) Mostrar la primera etapa y evaluar si cumple con la condición de relevancia
- 3) Discutir si el instrumento realmente cumple con la cláusula de exclusión (e independencia):
No → cambiar de instrumento o de metodología
Si >> debe explicar por qué cumple (y en que caso podría no cumplir)
- 4) Si se puede (y si Z_i es binaria), demuestre la independencia comparando las características iniciales de los $Z_i = 0$ con los $Z_i = 1$.
- 5) Discutir y examinar los canales (remplazando Y_i por cada canal)
- 6) Es bueno discutir cómo el LATE puede diferir del efecto en toda la población (validez externa)

¡Gracias!