

IVs

Ricardo Pasquini

FCE - Universidad Austral

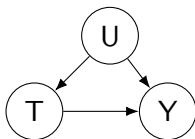
6/6/2023

Intro

- ▶ Seguimos en el contexto de datos observacionales, donde existen factores que pueden confundir (sesgar) un efecto causal de interés.
- ▶ El método de Variables Instrumentales (IVs) propone una nueva técnica para la identificación de un efecto causal de interés.
- ▶ El método se basa en contar con una variable adicional, "*la variable instrumental*", que se utilizará para aproximar el efecto. Esta variable debe cumplir dos condiciones, y estas no son siempre fáciles de conseguir.
- ▶ Veremos que utilizar un DAG es una buena forma de entender cuáles son las condiciones necesarias para esta variable.

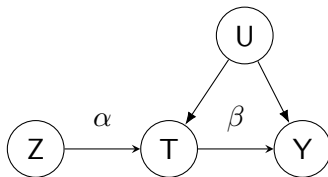
Intro

- ▶ Supongamos una situación en donde nos interesa conocer el efecto de una variable T sobre una variable Y , pero existe un factor U , no observable, que podría generar una correlación entre T e Y , sesgando la medición del efecto de interés (e.g., hay un sesgo de selección).
- ▶ Podemos representar esta relación en el siguiente DAG:



Método de VI

- ▶ Llamemos al efecto de interés de estimación β .
- ▶ El método de VI propone la identificación del efecto de β cuando contamos con una variable Z que tiene las siguientes características:
 1. Z tiene un efecto (no nulo) sobre T . Llamemos a este efecto α .
 2. U y Z no están directamente relacionados.
- ▶ Podemos representar esta relación en el siguiente DAG:

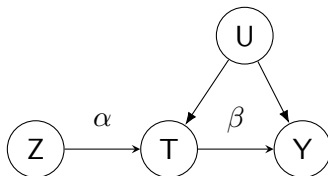


Método de VI

Estimación de la primera etapa

- ▶ Es importante notar que el efecto causal α se puede estimar sin sesgos ya que, en el DAG supuesto, no hay nada extra que afecte esa relación.
- ▶ A este efecto se le suele llamar "first stage".
- ▶ En un contexto donde Z es una dummy, el efecto se podría estimar simplemente como una diferencia:

$$\hat{\alpha} = E[T|Z = 1] - E[T|Z = 0]$$

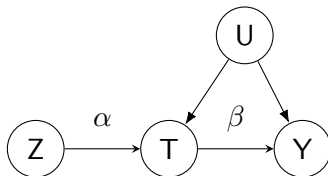


Método de VI

Estimación del efecto reducido

- ▶ También es importante notar que el efecto γ también se puede estimar sin sesgos.
 - ▶ La razón es que no hay ningún sendero alternativo (back-door path) entre Z e Y.
 - ▶ A este efecto se le suele llamar "reduced form".
 - ▶ En un contexto donde T es una dummy, el efecto se podría estimar simplemente como una diferencia:

$$\hat{\gamma} = E[Y|Z = 1] - E[Y|Z = 0]$$



Método de VI

Estimador LATE

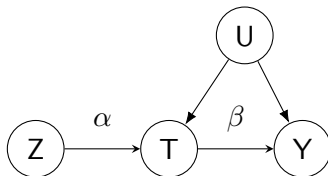
- ▶ Por último notemos que el efecto reducido γ se puede descomponer como la multiplicación de ambos efectos:

$$\gamma = \alpha\beta$$

- ▶ De aquí se sigue que podemos estimar el efecto de interés β como :

$$\hat{\beta}_{\text{LATE}} = \frac{\hat{\gamma}}{\alpha}$$

- ▶ Este estimador se conoce como Local Average Treatment Effect (LATE)



Caso: Efecto de escuelas charter KIPP en USA. (Angrist 2010, MM cap3)

- ▶ Las escuelas charter son escuelas públicas autónomas.

Caso: Efecto de escuelas charter KIPP en USA. (Angrist 2010, MM cap3)

- ▶ Las escuelas charter son escuelas públicas autónomas.
- ▶ Las pertenecientes al programa KIPP (Knowledge is Power Program) son especialmente emblemáticas pues tienen un modelo que enfatiza la disciplina en el estudio, jornadas alargadas, maestros especialmente seleccionados, y foco en las matemáticas y las habilidades de lectura.

Caso: Efecto de escuelas charter KIPP en USA. (Angrist 2010, MM cap3)

- ▶ Las escuelas charter son escuelas públicas autónomas.
- ▶ Las pertenecientes al programa KIPP (Knowledge is Power Program) son especialmente emblemáticas pues tienen un modelo que enfatiza la disciplina en el estudio, jornadas alargadas, maestros especialmente seleccionados, y foco en las matemáticas y las habilidades de lectura.
- ▶ El alumnado es 95% son de raza negra o latinos.

Caso: Efecto de escuelas charter KIPP en USA. (Angrist 2010, MM cap3)

- ▶ No hay posibilidad de selección aleatoria.

Caso: Efecto de escuelas charter KIPP en USA. (Angrist 2010, MM cap3)

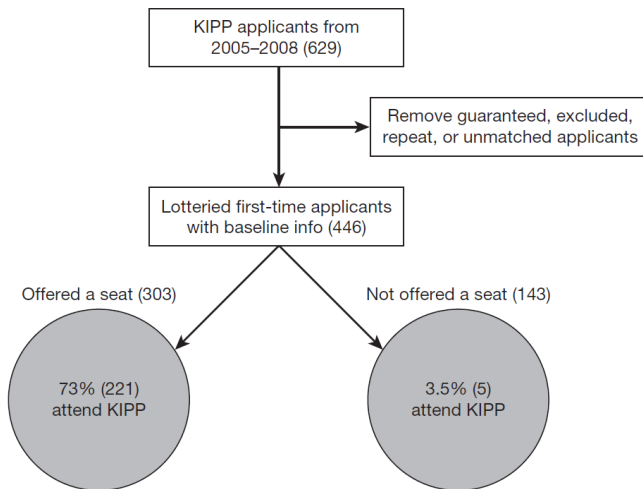
- ▶ No hay posibilidad de selección aleatoria.
- ▶ Pero un experimento natural ayuda a identificar el efecto:
Desde 2005 la demanda de asientos superó la oferta, y se comenzó a sortear el cupo.

Caso: Efecto de escuelas charter KIPP en USA. (Angrist 2010, MM cap3)

- ▶ No hay posibilidad de selección aleatoria.
- ▶ Pero un experimento natural ayuda a identificar el efecto: Desde 2005 la demanda de asientos superó la oferta, y se comenzó a sortear el cupo.
- ▶ El mecanismo de asignación, no obstante no es perfecto, como se muestra más abajo.

Caso: Efecto de escuelas charter. KIPP

FIGURE 3.1
Application and enrollment data from KIPP Lynn lotteries



Note: Numbers of Knowledge Is Power Program (KIPP) applicants are shown in parentheses.

Están balanceados los grupos? Analisis pre-intervención

TABLE 3.1
Analysis of KIPP lotteries

	KIPP applicants				
	Lynn public fifth graders	KIPP Lynn lottery winners	Winners vs. losers	Attended KIPP	Attended KIPP vs. others
	(1)	(2)	(3)	(4)	(5)
Panel A. Baseline characteristics					
Hispanic	.418	.510	−.058 (.058)	.539	.012 (.054)
Black	.173	.257	.026 (.047)	.240	−.001 (.043)
Female	.480	.494	−.008 (.059)	.495	−.009 (.055)
Free/Reduced price lunch	.770	.814	−.032 (.046)	.828	.011 (.042)
Baseline (4th grade) math score	−.307	−.290	.102 (.120)	−.289	.069 (.109)
Baseline (4th grade) verbal score	−.356	−.386	.063 (.125)	−.368	.088 (.114)

Post-intervención

TABLE 3.1
Analysis of KIPP lotteries

		KIPP applicants			
	Lynn public fifth graders (1)	KIPP Lynn lottery winners (2)	Winners vs. losers (3)	Attended KIPP (4)	Attended KIPP vs. others (5)
Panel B. Outcomes					
Attended KIPP	.000	.787	.741 (.037)	1.000	1.000 —
Math score	−.363	−.003	.355 (.115)	.095	.467 (.103)
Verbal score	−.417	−.262	.113 (.122)	−.211	.211 (.109)
Sample size	3,964	253	371	204	371

Caso: Efecto de escuelas charter. KIPP

- ▶ Lo que estamos seguros es que la comparación de la lotería es manzanas con manzanas. Pero no necesariamente la comparación entre quienes finalmente cursaron y quienes no.

Caso: Efecto de escuelas charter. KIPP

- ▶ Lo que estamos seguros es que la comparación de la lotería es manzanas con manzanas. Pero no necesariamente la comparación entre quienes finalmente cursaron y quienes no.
- ▶ Solución: Proponemos un estimador que permite computar el efecto de la asistencia (\leq lo que nos interesa), corrigiendo por el hecho de que el sorteo funciona como un mecanismo que:

Caso: Efecto de escuelas charter. KIPP

- ▶ Lo que estamos seguros es que la comparación de la lotería es manzanas con manzanas. Pero no necesariamente la comparación entre quienes finalmente cursaron y quienes no.
- ▶ Solución: Proponemos un estimador que permite computar el efecto de la asistencia (\leq lo que nos interesa), corrigiendo por el hecho de que el sorteo funciona como un mecanismo que:
 - ▶ separa a los grupos de potenciales tratados y no tratados de la manera que queremos (sin sesgo de selección)

Caso: Efecto de escuelas charter. KIPP

- ▶ Lo que estamos seguros es que la comparación de la lotería es manzanas con manzanas. Pero no necesariamente la comparación entre quienes finalmente cursaron y quienes no.
- ▶ Solución: Proponemos un estimador que permite computar el efecto de la asistencia (\leq lo que nos interesa), corrigiendo por el hecho de que el sorteo funciona como un mecanismo que:
 - ▶ separa a los grupos de potenciales tratados y no tratados de la manera que queremos (sin sesgo de selección)
 - ▶ si hacemos un ajuste por aquellos que no siguieron el camino indicado, podemos recuperar el efecto.

Caso: Efecto de escuelas charter. KIPP

- ▶ Lo que estamos seguros es que la comparación de la lotería es manzanas con manzanas. Pero no necesariamente la comparación entre quienes finalmente cursaron y quienes no.
- ▶ Solución: Proponemos un estimador que permite computar el efecto de la asistencia (\leq lo que nos interesa), corrigiendo por el hecho de que el sorteo funciona como un mecanismo que:
 - ▶ separa a los grupos de potenciales tratados y no tratados de la manera que queremos (sin sesgo de selección)
 - ▶ si hacemos un ajuste por aquellos que no siguieron el camino indicado, podemos recuperar el efecto.
- ▶ A este estimado lo llamamos **Método de Variables Instrumentales**.

Caso: Efecto de escuelas charter. KIPP

Intuitivamente el MVI propone una estrategia de identificación basada en una cadena causal:

$$\begin{aligned} \text{Efecto lotería en calificaciones} &= \text{Efecto lotería en asistencia} \\ &\quad * \text{Efecto asistencia en calificaciones} \end{aligned}$$

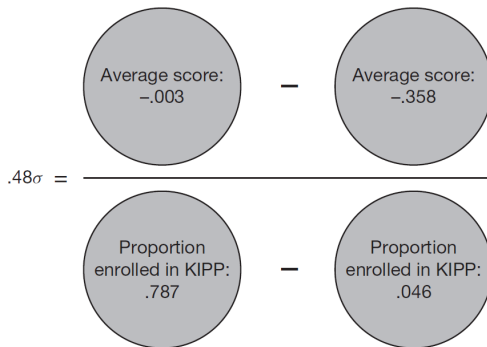
Puesto que nosotros estamos interesados en *Efecto asistencia en calificaciones* :

$$\text{Efecto asistencia en calificaciones} = \frac{\text{Efecto lotería en calificaciones}}{\text{Efecto lotería en asistencia}}$$

Caso: Efecto de escuelas charter. KIPP

FIGURE 3.2
IV in school: the effect of KIPP attendance on math scores

Offered a seat (253)	Not offered a seat (118)
----------------------	--------------------------



Note: The effect of Knowledge Is Power Program (KIPP) enrollment described by this figure is $.48\sigma = .355\sigma / .741$.

From Mastering Metrics: The Path from Cause to Effect. © 2015 Princeton University Press. Used by permission. All rights reserved.

Datos necesarios y definiciones para la estimación IV

- ▶ El instrumento Z : Aquí una variable que identifica a los que ofrecieron un asiento en las KIPP
- ▶ La variable de Tratamiento: D_i También una dummy que toma el valor $=1$ si asistió a una KIPP
- ▶ El resultado Y_i En este caso la nota de 5to grado

Definiciones IV

- ▶ *La primera etapa:*

$$E [D_i | Z_i = 1] - E [D_i | Z_i = 0] \equiv \phi$$

- ▶ *La forma reducida*

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

- ▶ *El efecto del tratamiento promedio local (Local Average Treatment Effect, LATE)*

$$\lambda = \frac{\rho}{\phi} = \frac{E [Y_i | Z_i = 1] - E [Y_i | Z_i = 0]}{E [D_i | Z_i = 1] - E [D_i | Z_i = 0]}$$

A quienes es aplicable el efecto LATE?

TABLE 3.2
The four types of children

		Lottery losers $Z_i = 0$	
		Doesn't attend KIPP $D_i = 0$	Attends KIPP $D_i = 1$
Lottery winners $Z_i = 1$	Doesn't attend KIPP $D_i = 0$	Never-takers (<i>Normando</i>)	Defiers
	Attends KIPP $D_i = 1$	Compliers (<i>Camila</i>)	Always-takers (<i>Alvaro</i>)

Note: KIPP = Knowledge Is Power Program.

A quienes es aplicable el efecto LATE?

- ▶ Con un instrumento basado en una lotería como en este caso, el LATE refleja el efecto del tratamiento para personas como Camila (compliers), pero no tiene nada de información sobre always-takers y never-takers, porque el instrumento no dice nada sobre su estatus de tratamiento.
- ▶ No pensamos que haya defiers en este caso. Si hubiere estaríamos en problemas. (este supuesto se llama monotonicidad)

En general TOT y LATE no van a coincidir

- ▶ TOT El efecto del Tratamiento sobre los Tratados está definido como:

$$E[Y_{1i} - Y_{0i} | D_i = 1]$$

- ▶ Pero en este caso, el grupo de los que toman el tratamiento son los *compliers* y también los *always takers*.
- ▶ De hecho en este caso podemos suponer que LATE subestima el TOT porque probablemente los *always takers* esperen un efecto mayor del tratamiento. Lo cierto es que no sabemos nada sobre este último grupo, y en general entonces no obtenemos TOT, sino solo LATE.
- ▶ Nota: En algunos casos particulares, puede pasar que por la naturaleza de la aplicación no hayan *always takers*. En ese caso excepcional tendremos TOT=LATE (ejm ver caso de Sherman y Beck sobre violencia doméstica)

Implementación econométrica usando 2SLS

En un primer paso explicamos el tratamiento en función del instrumento

$$D_i = \gamma Z_i + e_{0i}$$

En segundo lugar usamos los valores ajustados (los valores de la predicción: la variación que se debe a lo exógeno), estimamos el efecto sobre el resultado.

$$Y_i = \alpha X_i + \beta(\hat{\gamma}Z_i) + \varepsilon_i$$

Case: Quantity Quality Trade Off in Family Size

- ▶ Evidence usually finds a negative correlation between family size and education
- ▶ Comparison of families will probably entail selection bias.
- ▶ Desired experiment: random assignment of kids.
- ▶ Any ideas for a (natural) experiment?

Case: Quantity Quality Trade Off in Family Size

- ▶ Twin births, might exogenously affect family size.
- ▶ When twins are second born, statistically changes the expected size of the family (when first-born, families adjust back)
- ▶ Another idea: when the second born is same sex as the first born, families expected size is 3.68. When they are different, family size is 3.6: Another instrument!

TABLE 3.4
Quantity-quality first stages

	Twins instruments		Same-sex instruments		Twins and same- sex instruments
	(1)	(2)	(3)	(4)	(5)
Second-born twins	.320 (.052)	.437 (.050)			.449 (.050)
Same-sex sibships			.079 (.012)	.073 (.010)	.076 (.010)
Male		-.018 (.010)		-.020 (.010)	-.020 (.010)
Controls	No	Yes	No	Yes	Yes

Notes: This table reports coefficients from a regression of the number of children on instruments and covariates. The sample size is 89,445. Standard errors are reported in parentheses.

From *Mastering 'Metrics: The Path from Cause to Effect*. © 2015 Princeton University Press. Used by permission.
All rights reserved.

TABLE 3.5
OLS and 2SLS estimates of the quantity-quality trade-off

Dependent variable	OLS estimates (1)	2SLS estimates		
		Twins instruments (2)	Same-sex instruments (3)	Twins and same- sex instruments (4)
Years of schooling	-.145 (.005)	.174 (.166)	.318 (.210)	.237 (.128)
High school graduate	-.029 (.001)	.030 (.028)	.001 (.033)	.017 (.021)
Some college (for age ≥ 24)	-.023 (.001)	.017 (.052)	.078 (.054)	.048 (.037)
College graduate (for age ≥ 24)	-.015 (.001)	-.021 (.045)	.125 (.053)	.052 (.032)

Notes: This table reports OLS and 2SLS estimates of the effect of family size on schooling. OLS estimates appear in column (1). Columns (2), (3), and (4) show 2SLS estimates constructed using the instruments indicated in column headings. Sample sizes are 89,445 for rows (1) and (2); 50,561 for row (3); and 50,535 for row (4). Standard errors are reported in parentheses.