

Modelos líneales microeconometria (Cameron & Trivedi)

Santiago Alonso-Diaz, PhD

Profesor Asistente
Departamento de Economía
Universidad Javeriana

ADVERTENCIA: No tiene tildes para evitar problemas con Latex que no pude solucionar

Tutorial R

Hacer tutorial R (<https://rstudio.cloud/learn/primers>)



Capitulo 3 (Stata book) & 4 (Other book): Modelos Lineales

Modelos Lineales

Ejemplo: retornos de educación

$$\ln(\text{salario}_i) = \alpha \text{educacion}_i + \beta \text{controles}_i + u_i$$

- ▶ La transformacion a logaritmo es opcional, pero puede ayudar a producir errores homocedastico (i.e. que no depende de nivel de educacion)
- ▶ α puede interpretarse directamente en ESTA regresion e.g. si $\hat{\alpha} = 0,1$, un agno de educacion aumenta 10% el salario
- ▶ En ESTA regresion educacion es endogena, no es aleatoria, no se puede concluir causalidad.

Modelos Lineales

Objetivo: estimar parametros α y β

$$\ln(\text{salario}_i) = \alpha \text{educacion}_i + \beta \text{controles}_i + u_i$$

Metodo: minimizar perdida (L) esperada

$$\min_{\hat{y}} E[L(y - \hat{y})]$$

L puede cambiar pero la mas conocida es $L = \sum_{i=1}^N (y_i - \beta x)$

Modelos Lineales

Estimador OLS (ordinary least squares)

$$\sum_{i=1}^N u_i^2 = (y - \beta X)'(y - \beta X)$$

Ponemos la derivada con respecto a β igual a zero, obtenemos el estimador OLS,

$$\hat{\beta}_{OLS} = (X'X)^{-1}X'y$$

Modelos Lineales

Estimadores que incluyen la varianza del error (ir a LinearRegression.R).

Table 4.2. Least-Squares Estimators and Their Asymptotic Variance

Estimator ^a	Definition	Estimated Asymptotic Variance
OLS	$\hat{\beta} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}$	$(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\hat{\Omega}\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}$
FGLS	$\hat{\beta} = (\mathbf{X}'\hat{\Omega}^{-1}\mathbf{X})^{-1}\mathbf{X}'\hat{\Omega}^{-1}\mathbf{y}$	$(\mathbf{X}'\hat{\Omega}^{-1}\mathbf{X})^{-1}$
WLS	$\hat{\beta} = (\mathbf{X}'\hat{\Sigma}^{-1}\mathbf{X})^{-1}\mathbf{X}'\hat{\Sigma}^{-1}\mathbf{y}$	$(\mathbf{X}'\hat{\Sigma}^{-1}\mathbf{X})^{-1}\mathbf{X}'\hat{\Sigma}^{-1}\hat{\Omega}\hat{\Sigma}^{-1}\mathbf{X}(\mathbf{X}'\hat{\Sigma}^{-1}\mathbf{X})^{-1}$.

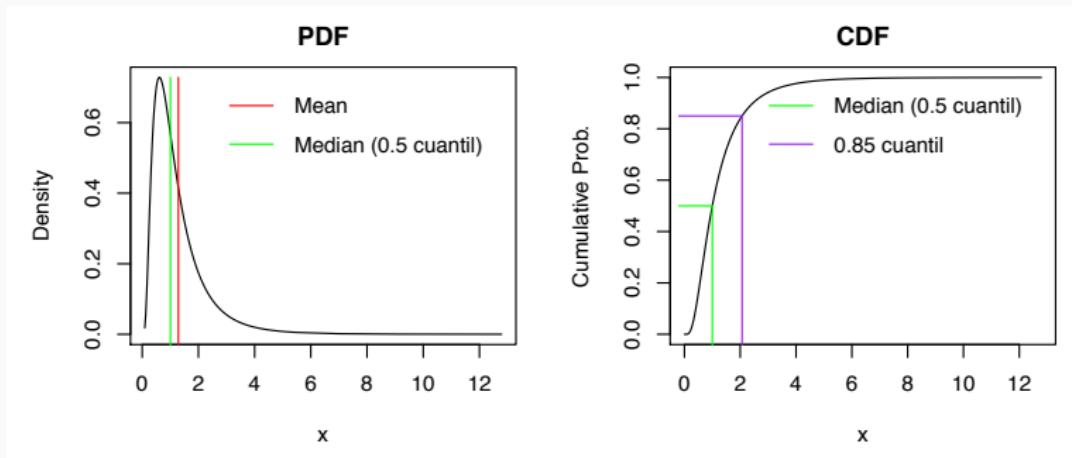
^a Estimators are for linear regression model with error conditional variance matrix Ω . For FGLS it is assumed that $\hat{\Omega}$ is consistent for Ω . For OLS and WLS the heteroskedastic robust variance matrix of $\hat{\beta}$ uses $\hat{\Omega}$ equal to a diagonal matrix with squared residuals on the diagonals.

Lineal: Ejemplo problemas con heteroscedasticidad



Modelos Lineales

NUEVO TEMA: Modelos de medianas y cuantiles



Modelos Lineales

Nos interesa el cuantil q de $y|x$ e.g. el cuantil 0.95 de ingresos dado un nivel de educación.

$$\mu_q(x) = F_{y|x}^{-1}(q)$$

$F_{y|x}(q)$ es la CDF de $y|x$. La CDF nos da una probabilidad pero nos interesa la inversa i.e. el valor de x que corresponde a esa probabilidad-cuantil.

Por ejemplo

$$\mu_{0.95}(\text{pregrado}) = F_{\text{ingresos}| \text{pregrado}}^{-1}(0.95)$$

Modelos Lineales

Ahora definimos un modelo con heterocedasticidad (i.e. con error u que aumenta con x),

$$y = x'\beta + u$$

$$u = x'\alpha \times \epsilon$$

$$\epsilon \sim iid[0, \sigma^2]$$

Modelos Lineales

Ahora obtenemos $\mu_q(x, \beta\alpha)$. El cuantil q por definicion es,

$$q = Pr[y \leq \mu_q(x, \beta\alpha)]$$

Modelos Lineales

Reemplazando y obtenemos

$$\begin{aligned} q &= Pr[y \leq \mu_q(x, \beta\alpha)] \\ &= Pr[u \leq \mu_q(x, \beta\alpha) - x'\beta] \end{aligned}$$

Modelos Lineales

Reemplazando u obtenemos

$$\begin{aligned} q &= \Pr[y \leq \mu_q(x, \beta\alpha)] \\ &= \Pr[u \leq \mu_q(x, \beta\alpha) - x'\beta] \\ &= \Pr[\epsilon \leq (\mu_q(x, \beta\alpha) - x'\beta)/x'\alpha] \end{aligned}$$

Modelos Lineales

Es decir, encontramos que q es la CDF del error ϵ evaluada en una función de α & β

$$\begin{aligned} q &= \Pr[y \leq \mu_q(x, \beta\alpha)] \\ &= \Pr[u \leq \mu_q(x, \beta\alpha) - x'\beta] \\ &= \Pr[\epsilon \leq (\mu_q(x, \beta\alpha) - x'\beta)/x'\alpha] \\ &= F_e[(\mu_q(x, \beta\alpha) - x'\beta)/x'\alpha] \end{aligned}$$

Modelos Lineales

Ahora estamos en posición de encontrar $\mu_q(x, \beta\alpha)$ invirtiendo F_e

$$\begin{aligned} q &= Pr[y \leq \mu_q(x, \beta\alpha)] \\ &= Pr[u \leq \mu_q(x, \beta\alpha) - x'\beta] \\ &= Pr[\epsilon \leq (\mu_q(x, \beta\alpha) - x'\beta)/x'\alpha] \\ &= F_e[(\mu_q(x, \beta\alpha) - x'\beta)/x'\alpha] \end{aligned}$$

$$F_e^{-1}(q) = (\mu_q(x, \beta\alpha) - x'\beta)/x'\alpha$$

$$\mu_q(x, \beta\alpha) = x'(\beta + \alpha \times F_e^{-1}(q))$$

Modelos Lineales

Es decir, los cuantiles son lineales en x con la heterocedasticidad que definimos i.e. $u = x'\alpha \times \epsilon$

$$\mu_q(x, \beta\alpha) = x'(\beta + \alpha \times F_e^{-1}(q))$$

Puede haber heterocedasticidad no lineal pero es usual usar la aproximación lineal

Modelos Lineales

La función objetivo a minimizar en β_q

$$Q_N(\beta_q) = \sum_{i:y_i \geq x_i' \beta}^N q|y_i - x_i' \beta_q| + \sum_{i:y_i \geq x_i' \beta}^N (1-q)|y_i - x_i' \beta_q|$$

Esta función no es intuitiva, hay que obtenerla (ver Koenker and Bassett (1978)). Lo importante es que este es el costo entre data (y) y modelo lineal ($x_i' \beta$) a minimizar.

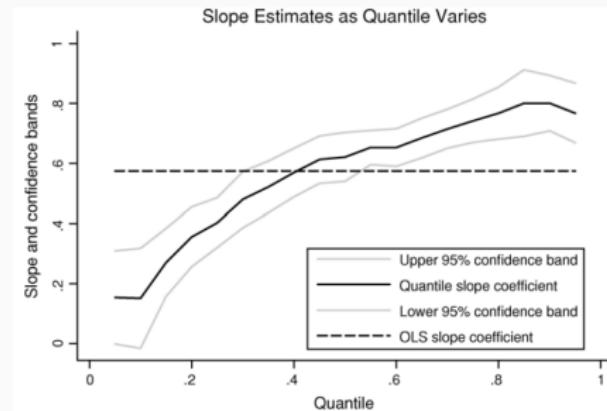
IMPORTANTE: no es diferenciable (valores absolutos), hay que usar métodos numéricos.

Lineal cuantiles: ejemplo



Modelos Lineales

$Gasto_{total} \sim Intercepto + \beta Gasto_{salud} + \epsilon$ en Vietnam



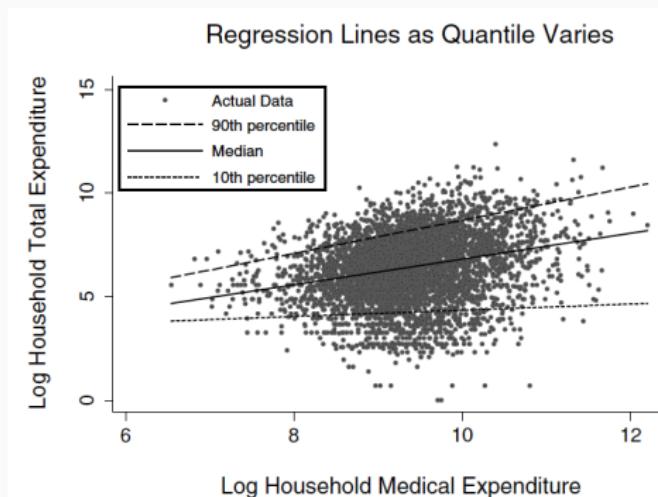
Que quiere decir una pendiente = 0.57. Elastico o inelastico?

Una pendiente menor en $q = 0.1$ y mayor en $q = 0.9$

Modelos Lineales

El gasto en salud es inelastico pero en menor medida en cuantiles con ingreso mayores.

Pendientes diferentes por cuantil es evidencia de heteroscedasticidad o no-linealidad



Modelos Lineales

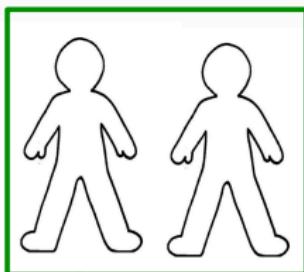
Hacer ejercicio 1 del taller (punto de heterocedasticidad).

Causalidad: Diferencia en diferencias (contrafactuals)

Modelos Lineales

$$y_{i1}^0 - y_{i0}^0$$

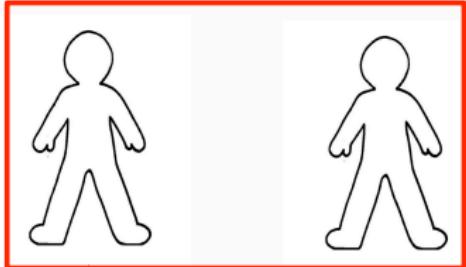
$$\overline{t_0 \quad t_1}$$



j=0

$$y_{i1}^1 - y_{i0}^1$$

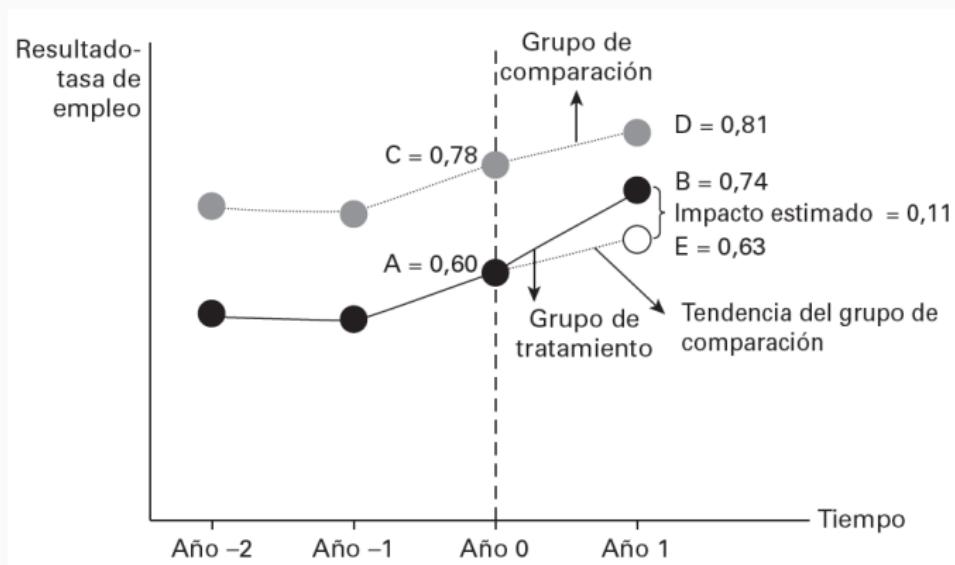
$$\overline{t_0 \quad t_1}$$



j=1

Modelos Lineales

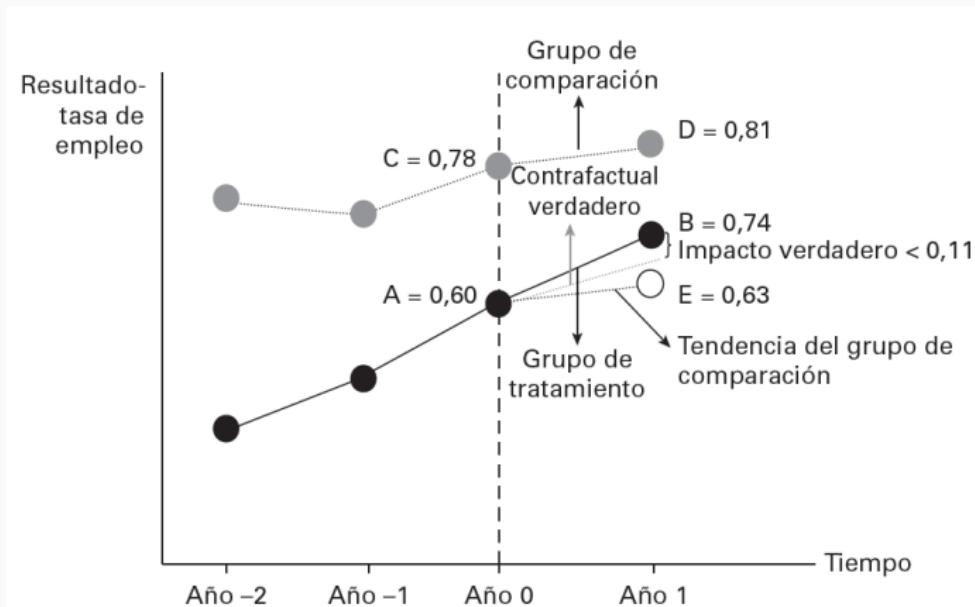
Otra visualización. SUPUESTO IMPORTANTE: antes del tratamiento las tendencias eran iguales



[Gertler et al., 2016]

Modelos Lineales

No se cumple el supuesto.



[Gertler et al., 2016]

Modelos Lineales

Otra visualizacion.

Cuadro 7.1 Cálculo del método de diferencias en diferencias

	Después	Antes	Diferencia
Tratamiento/inscritos	B	A	$B - A$
Comparación/no inscritos	D	C	$D - C$
Diferencia	$B - D$	$A - C$	$DD = (B - A) - (D - C)$

[Gertler et al., 2016]

Modelos Lineales

Caso:

El sector del agua se presta para monopolios (altos costos de entrada), externalidades (transmision/prevencion de enfermedades), y demanda inelastica.

Usualmente publico pero pocos incentivos para mejorar (e.g. infraestructura) y riesgo de politizar el agua.

Argentina 1990s privatiza el suministro del agua, con la mayoria de privatizaciones en 1995

[Galiani et al., 2005]

En los 90s en Argentina se privatizo el suministro en muchos municipios
(la decision era local)

TABLE 1
CHANGE IN OWNERSHIP OF WATER SYSTEMS, 1990–99

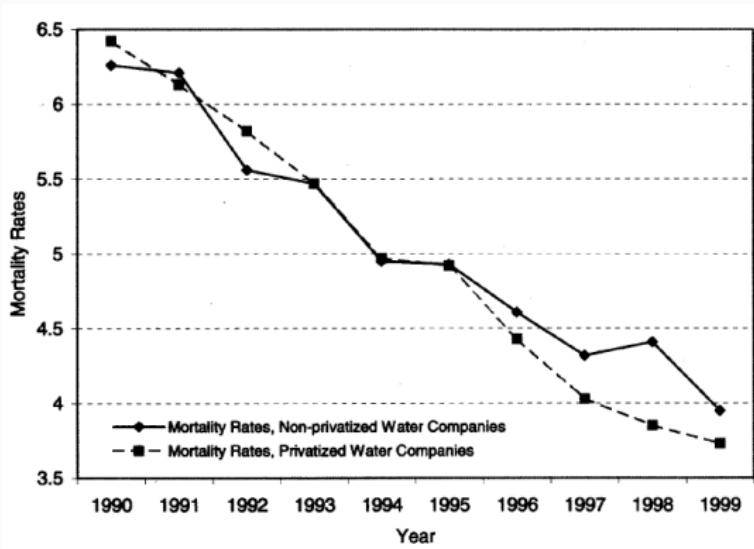
Ownership	Number of Municipalities	Percentage
Always public	196	39.7
Always private not-for-profit cooperative	143	28.9
Transferred from public to private for-profit	137	27.7
Always private for-profit	1	.2
No service or missing information	17	3.4
Total	494	100.0

NOTE.—In municipalities in which more than one company provides water services, we defined the ownership status of the municipality as the ownership of the company supplying the largest fraction of the population. Source: SPIDES, ENOHSA.

[Galiani et al., 2005]

Modelos Lineales

Tasas de mortalidad parecen bajar mas rapido en municipios donde se privatizo



[Galiani et al., 2005]

El modelo formal de diferencia en diferencias usado por Galiani et al 2005
(panel)

$$y_{it} = \alpha dI_{it} + \beta x_{it} + \lambda_t + \mu_i + \epsilon_{it}$$

y_{it} : tasa de mortalidad en municipalidad i en agno t

dI_{it} : dummy; 1 si el servicio de agua en la municipalidad i en el agno t era privado; 0 si publico (este es el D&D)

x_{it} : variables de control que cambian en tiempo y municipalidad

λ_t : cambios en el tiempo comunes a todas las municipalidades

μ_i : efectos fijos que cambian en cada municipalidad

[Galiani et al., 2005]

Modelos Lineales

Tener servicios de agua privados (=1) redujo tasas de mortalidad (Coef. neg.).

	FULL SAMPLE		
	(1)	(2)	(3)
Private water services (=1)	-.334 (.169)** [.157]** .195)*	-.320 (.170)* [.163]** .203]	-.283 (.170)* [.162]* .194]
%Δ in mortality rate	-5.3	-5.1	-4.5
Other covariates:			
Real GDP per capita	.007 (.005) [.006] .007]	.009 (.006) [.006] .007]	
Unemployment rate	-.555 (1.757) [2.161] [2.862]	-.636 (1.758) [2.166] [2.846]	
Income inequality	5.171 (2.868)* [3.468] [3.696]	5.085 (2.880)* [3.445] [3.691]	
Public spending per capita	-.028 (.038) [.055] .054]	-.035 (.038) [.055] .055]	
Local government by Radical party (=1)		.482 (.267)* [.281]* [.288]*	
Local government by Peronist party (=1)		-.202 (.191) [.202] .254]	
R ²	.1227	.1256	.1272
Observations	4,792	4,597	4,597

Modelos Lineales

Bajaron todas pero en especial las parasiticas

TABLE 4
IMPACT OF PRIVATIZATION ON CHILD MORTALITY BY CAUSE OF DEATH

	1990 Mean Mortality Rate	Estimated Impact Coefficients	%Δ in Mortality Rate
Infectious and parasitic diseases	.565	-.103 (.048)** [.055]* .068}	-18.2
Perinatal deaths	2.316	-.266 (.105)** [.107]** .123)**	-11.5
All other causes in aggregate	2.565	-.082 (.114) [.101] .109}	-3.2

[Galiani et al., 2005]

Modelos Lineales

El efecto fue detectable en municipalidades mas pobres

TABLE 5
IMPACT OF PRIVATIZATION ON CHILD MORTALITY BY POVERTY LEVEL

	1990 Mean Mortality Rate	Estimated Impact Coefficients	%Δ in Mortality Rate
Nonpoor municipalities	5.07	.114 (.233) [.165] .159	...
Poor municipalities	6.97	-1.004 (.279)*** [.297]*** .278***	-14.4
Extremely poor municipalities	9.11	-2.415 (.544)*** [1.051]** .605***	-26.5

[Galiani et al., 2005]

Modelos Lineales

En este caso, se mejoró la eficiencia ... pero si se bajaron los trabajadores de 9300 a 4000

TABLE 6
COMPARISON OF OSN (Public) VS. AGUAS ARGENTINAS (Private) PERFORMANCE, 1980–99

	OSN ^a (before Privatization)	Aguas Argentinas ^b (after Privatization)	Δ after Privatization (%)
Water production (1) (millions of m ³ per day)	3.56	3.89	9.3
Spilled water (2) (millions of m ³ per day)	1.49 ^c	1.27	-14.8
Water supply (1–2) (millions of m ³ per day)	2.07 ^c	2.62	26.6
Sewage drainage volume (millions of m ³ per day)	2.18	2.45	12.4
Water network extension (km of network)	10,148	13,287	30.9
Sewerage network extension (km of network)	6,875	8,312	20.9
Average delay in attending repair requests (days)	180 ^d	32 ^e	-82.2
Water leakages repaired per year	42,000 ^f	96,383	129.5
Sewerage blockages repaired per year	100,000 ^f	148,500	48.5
Percentage of clients with appropriate water pressure	17 ^c	54 ^f	217.6
Water turbidity (turbidity units)	7.5	2.3	-70
Usage fee index ^g	100	84	-16
Employees	9,300	4,000	-57

[Galiani et al., 2005]

Modelos Lineales

z-test: otra forma de hacer diferencia en diferencias sin regresiones

TABLE 8
DIFFERENCE-IN-DIFFERENCES ESTIMATES OF THE IMPACT OF PRIVATIZATION ON THE PROPORTION OF HOUSEHOLDS CONNECTED TO THE WATER NETWORK, 1991–97

	All Municipalities	Excluding Buenos Aires
	Municipalities That Were Not Privatized before 1997	
Proportion of households connected in 1991 (p_{91}^{public})	.866	.866
Proportion of households connected in 1997 (p_{97}^{public})	.898	.898
Difference 1997 – 1991 ($p_{97}^{\text{public}} - p_{91}^{\text{public}}$)	.032	.032
	Municipalities That Were Privatized before 1997	
Proportion of households connected in 1991 (p_{91}^{private})	.730	.640
Proportion of households connected in 1997 (p_{97}^{private})	.780	.714
Difference 1997 – 1991 ($p_{97}^{\text{private}} - p_{91}^{\text{private}}$)	.050	.074
Difference-in-differences ($p_{97}^{\text{private}} - p_{91}^{\text{private}}$ – $(p_{97}^{\text{public}} - p_{91}^{\text{public}})$)	.018	.042
Ztest for difference-in-differences estimate ^a	2.83***	5.78***
$z = \frac{(p_{97}^{\text{private}} - p_{91}^{\text{private}}) - (p_{97}^{\text{public}} - p_{91}^{\text{public}})}{\sqrt{[p_{97}^{\text{private}}(1 - p_{97}^{\text{private}})/n_{97}^{\text{private}}] + [p_{97}^{\text{public}}(1 - p_{97}^{\text{public}})/n_{97}^{\text{public}}]}},$		

[Galiani et al., 2005]

Modelos Lineales

Es un buen caso de diferencia en diferencias por que:

1. Las decisiones de privatizar son locales y no afectan mortalidad (no endogeneidad).
2. Control y tratamiento tenian la misma tendencia antes de privatizar (ver grafica).
3. La mortalidad infantil mejoró principalmente por enfermedades parasiticas
4. Los efectos se presentaron en estratos bajos

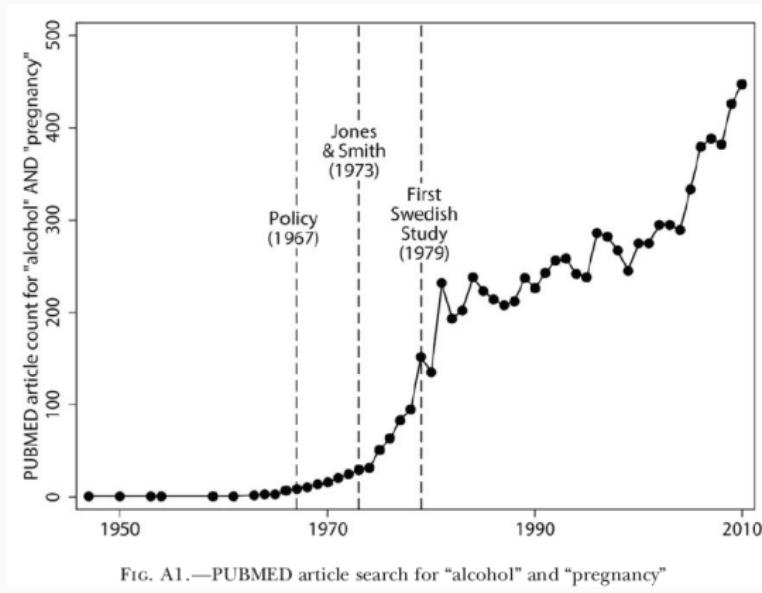


[Galiani et al., 2005]

Modelos Lineales

Caso 2:

Hoy sabemos que el alcohol afecta el desarrollo fetal y cognitivo ... en 1967 no



[Nilsson, 2017]



- ▶ La mayoría de estudios, sin embargo, son correlacionales.
- ▶ En Suecia solo hay una tienda estatal que vende trago fuera de bares y restaurantes: Systembolaget
- ▶ Experimento Natural: 1967 (Nov) - 1968 (Julio) en 2 regiones se permite vender cerveza fuerte en supermercados regulares

[Nilsson, 2017]

Modelos Lineales

La política terminó por el aumento exagerado de consumo de cerveza.

Más importante, se cumple el supuesto de tendencias iguales antes de la política.

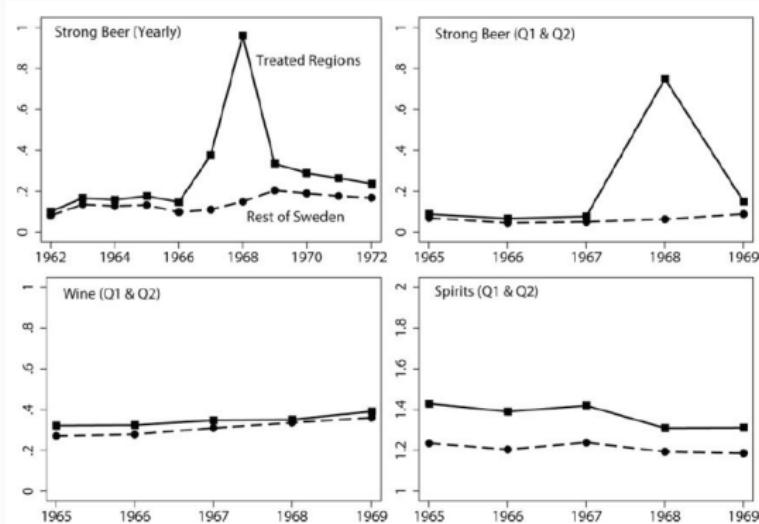


FIG. 1.—Strong beer, spirits, and wine sales, in liters of 100 percent alcohol per capita.
Sources: Yearly data from Statistics Sweden (1962–72); quarterly data from SOU (1971:77).
Data are not available for the quarters individually.

Modelos Lineales

Este es el modelo de diferencia en diferencias usado por Nilsson. Nos interesa β_3 : impacto de la política en niños que estaban temprano en el embarazo en madres menores de 21 (*>21 ya podían comprar en Systembolaget, con la política >16 en los supermercados*)

$$Y_{r,t,m<21} = \alpha_0 + \beta_3 EXPOSURE_{r,t,m<21} + \eta_{r,t} + \eta_{r,m<21} + \eta_{t,m<21} + \epsilon_{r,t,m<21}$$

t: mes; $m_{<21}$: edad de la madre en binario (<21); r: región de nacimiento

$Y_{r,t,m<21}$: outcome de interés (e.g. ingreso promedio)

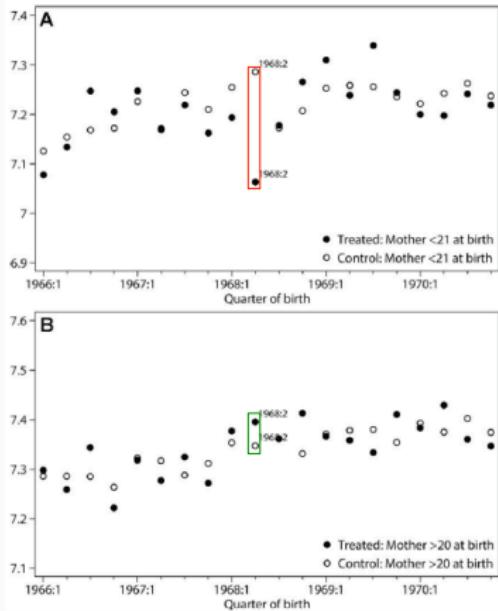
$EXPOSURE_{r,t,m<21}$: 1 si mama <21 & nacimiento en región de tratamiento y concebido entre jul. y oct. de 1967; de lo contrario 0

Y los η ? Quiz o discusión. Tip: ver los subíndices.

[Nilsson, 2017]

Modelos Lineales

La diferencia en diferencias es notable: los ninos concebidos en los 8 meses de la politica tienen menores log(ingresos) a los 32 (**<21** vs. **>21**).



[Nilsson, 2017]

Modelos Lineales

Los niños concebidos en los 8 meses de la política tienen peores habilidades cognitivas y no cognitivas (determinadas por un examen del ejercito)

TABLE 6
EFFECTS ON COGNITIVE AND NONCOGNITIVE ABILITY

	Average Score (1)	P(Low Score) (2)	P(Medium Score) (3)	P(High Score) (4)
A. Cognitive Ability				
Exposed from 1st half of pregnancy	-.053 (.070)	.069*** (.025)	-.052*** (.020)	-.002 (.029)
R ²	.88	.69	.29	.85
Mean of dependent variable (young mothers)	-.224	.257	.585	.167
B. Noncognitive Ability				
Exposed from 1st half of pregnancy	-.099 (.096)	.030** (.013)	-.013 (.050)	-.016 (.066)
Adjusted R ²	.74	.56	.42	.68
Mean of dependent variable (young mothers)	-.113	.188	.640	.207

NOTE.—The table reports eq. (4) estimated effects on cognitive and noncognitive skills score as graded by trained psychologists at military enlistment procedure at age 18 for males. For ease of comparison, all scores have been standardized, with mean zero and

Modelos Lineales

Los niños concebidos en los 8 meses de la política tienen menos educación (tabla superior), y esto no se ve en regiones vecinas donde no se implementó la política (tabla inferior)

TABLE 7
LONG-TERM EFFECTS BY FAMILY INCOME

	OUTCOME VARIABLES					
	(ln) Earnings (1)	Zero Earnings (2)	Welfare (3)	Years of Schooling (4)	High School (5)	University (6)
A. Baseline Estimates for the Family Income Sample						
Exposed from 1st half of pregnancy	-.269*** (.105)	.074*** (.024)	.041*** (.012)	-.464*** (.111)	-.081*** (.020)	.057*** (.018)
Mean of dependent variable	7.14	.11	.06	11.79	.837	.194

TABLE 8
IMPACT ON NEIGHBORING REGIONS: FULL SAMPLE

	DEPENDENT VARIABLE					
	Earnings	Zero Earnings	Welfare	Years of Schooling	High School Graduates	Higher Education
Exposed from 1st half of pregnancy	.047 (.060)	.023 (.017)	.005 (.010)	-.071 (.117)	-.006 (.019)	-.015 (.027)
Observations	4,752	4,752	4,752	4,752	4,752	4,752

Nilsson encontro con un modelo de diferencia en diferencias

1. La politica tuvo un efecto en el desarrollo fetal, cognitivo, y de vida (ingresos, educacion).
2. No se mostro en estas diapositivas, pero el efecto no fue intergeneracional, no afecto a los ninos de los nacidos durante la politica.

[Nilsson, 2017]

Modelos Lineales

En general, tips prácticos si hacen diferencia en diferencias

1. Revisar que tendencias antes de la intervencion son similares
2. Buscar en lo posible varios grupos de comparacion, por ejemplo, dos grupos adicionales que no les afecto el programa.

Diff. in Diff. ejemplo



Modelos Lineales

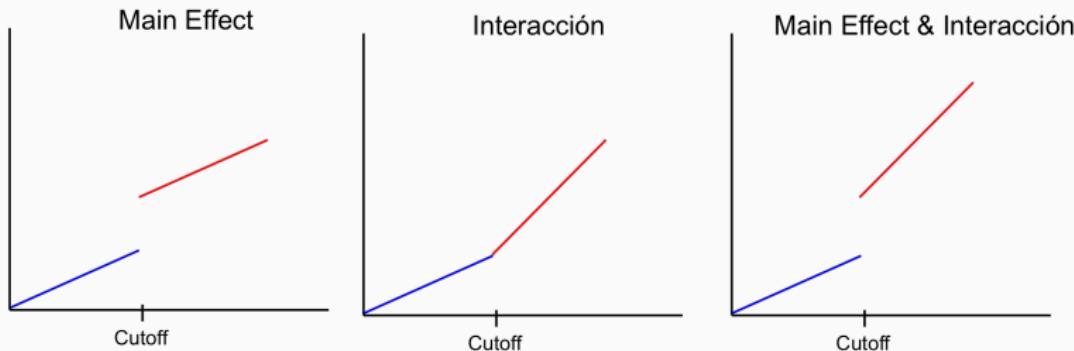
Hacer ejercicio 2 del taller (punto de diff in diff).

Modelos Lineales

Causalidad: Regresion discontinua (contrafactuals)

Modelos Lineales

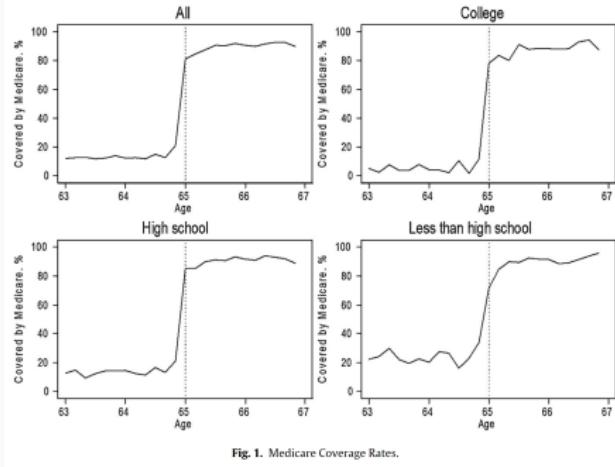
Despues de un umbral o cutoff en una variable de seleccion S el outcome cambia. La figura muestra una discontinuidad que se llama sharp. El otro tipo es fuzzy, (e.g. prob. acato a la intervencion < 100%). Solo veremos el tipo sharp.



Modelos Lineales

Algunos supuestos importantes

- ▶ Asignar tratamiento con umbrales se cumple (e.g. no hay sobornos para entrar a Medicare)
- ▶ El efecto solo está en el umbral (e.g. hacer regresiones placebo en diferentes umbrales)
- ▶ El umbral no afecta ninguna variable independiente



Modelos Lineales

Ejemplo del supuesto que el umbral no afecta ninguna variable independiente de manera considerable. BAC: niveles de alcohol en la sangre. Líneas verticales dos regulaciones (pasar umbral: castigo). Eje y: reincidencia de exceder de límite.

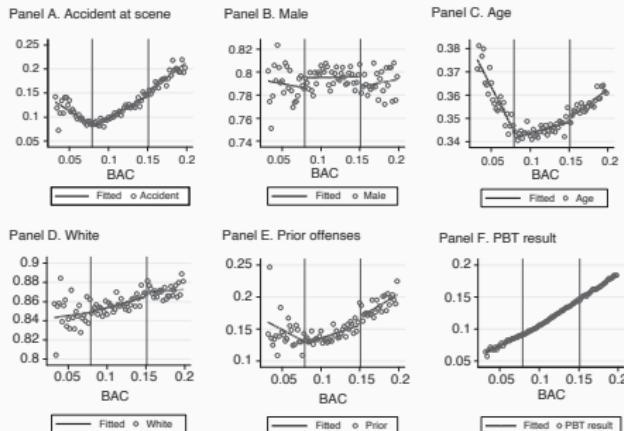


Figure 2. BAC and Characteristics

Notes: Based on administrative records from the Washington State Impaired Driver Testing Program, 1999–2007. Points represent the averages, with fitted values based on local linear models in black lines. The vertical black lines represent the two legal thresholds at 0.08 and 0.15. The bin width is 0.002.

Modelos Lineales

Ejemplo de umbral: En la figura vemos dos umbrales con agno de nacimiento: caen las personas con menos de 9 agnos de educacion por leyes que obligaban a educarse

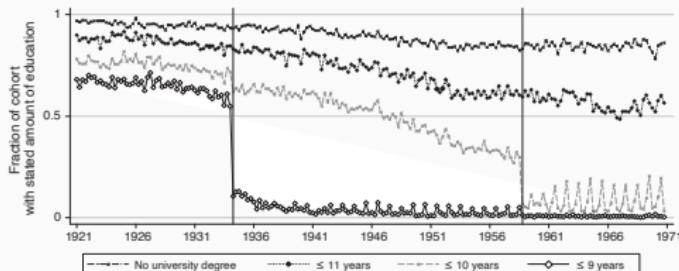


Figure 1. Years of Full-Time Education by Quarter of Birth

Notes: Sample includes individuals in the Health Survey for England: 1991–2004. Points represent means among people in each quarter-year of birth cell (all later graphs present data by month-year of birth). The vertical lines are cutoffs corresponding to the first cohorts subject to the new compulsory schooling laws. The first of these took effect on April 1, 1947 and the second took effect on September 1, 1972. Thus, since the two compulsory schooling reforms affected 14 (first reform) and 15 (second reform) year olds, the first cohorts impacted are those born in April 1933 for the first reform and September 1957 for the second reform.

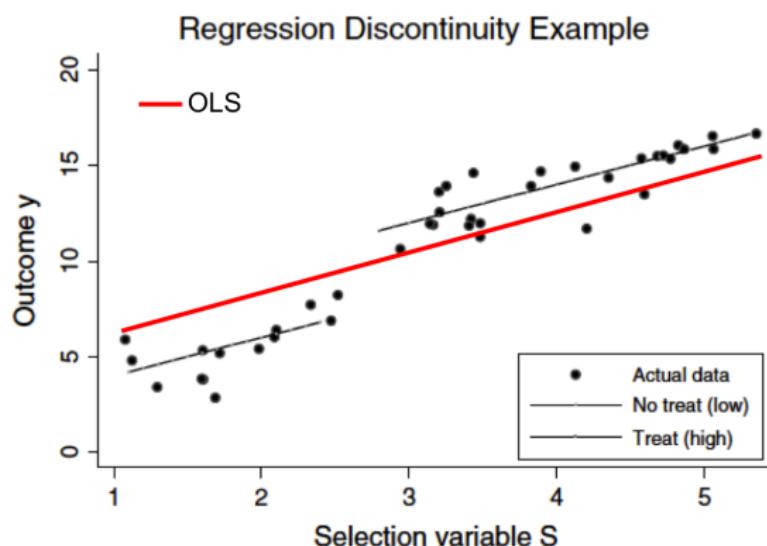
Ejemplos de umbrales: **Edad de pension; Dia de nacimiento; Distancia a frontera geografica; Severidad de enfermedad; Puntaje en prueba academic;a; limites legales**

[Clark and Royer, 2013]

Modelos Lineales

$y_i = \beta_0 + \beta_1 Treat. + \beta_2 Sel. + \epsilon$ no es apropiada

β_1 es inconsistente i.e. incluso con mucha data no obtendriamos el verdadero efecto del tratamiento



Una solucion es introducir una funcion de control a la regresion,

$$y_i = \beta_0 + \beta_1 Treat. + k(Sel_i) + \epsilon_i$$

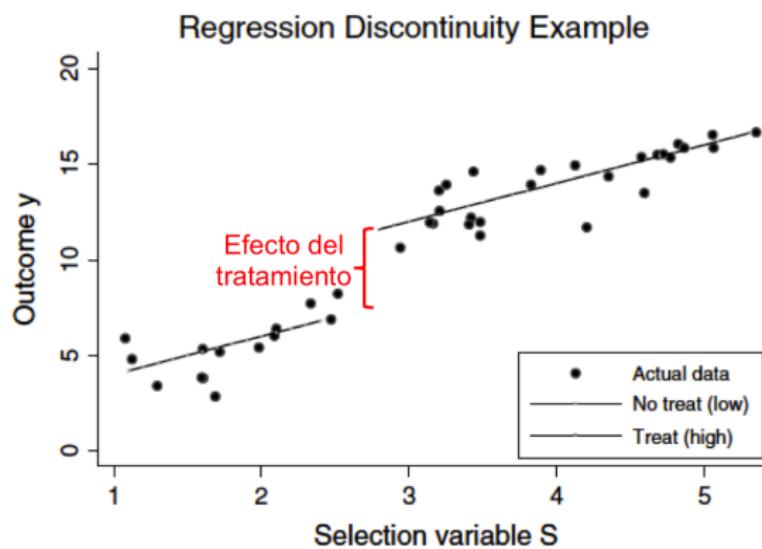
$k(Sel_i)$ es la funcion de control. $\epsilon_i = y_i - E[y_i|D_i, Sel_i]$, es decir el error se desplaza dependiendo de donde esta la funcion dada la discontinuidad.

Ejemplo de $k()$: distancia al umbral

Modelos Lineales

$$y_i = \beta_0 + \beta_1 \text{ Treat.} + k(Sel_i) + \epsilon_i$$

Si $k(Sel_i)$ es lineal, y dado que Treat. es una dummy [0,1], ahora podemos interpretar β_1 como el 'salto' i.e. efecto de tratamiento



Ademas de usarse en regresiones discontinuas, las funciones de control, cuando son lineales, llevan a controlar por endogeneidad pues es similar a hacer 2SLS (ver seccion de variables instrumentales).

Cuando no son lineales tambien ayudan a controlar por endogeneidad pero la conexion con 2SLS no es evidente.

Ejemplo regresion discontinua: transferencias de dinero condicionadas (CCT, siglas en ingles)

- ▶ Las CCT son instrumentos de politica publica para la pobreza e.g. se transfiere dinero si los padres llevan al nino a la escuela
- ▶ Se ha mostrado que aumentan el uso de servicios de salud y educacion
- ▶ Pero poco se sabe si mejoran resultados e.g. en notas de los ninos o estado de salud

[Levy and Ohls, 2010]

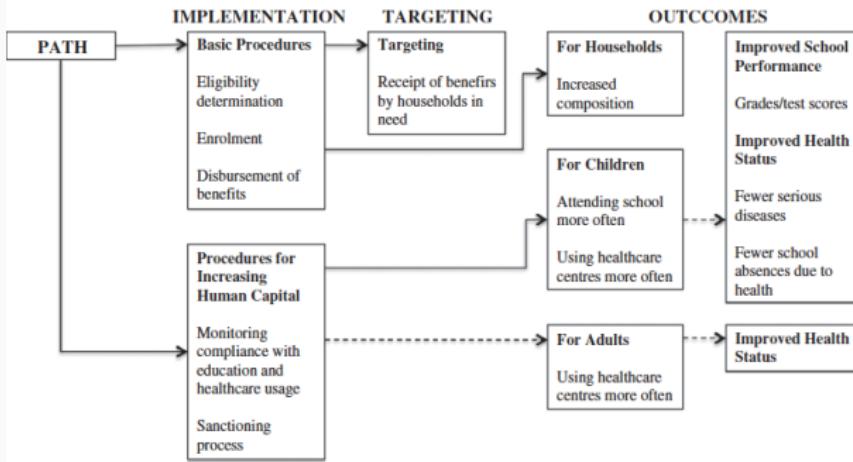


- ▶ Padres no mandan a la escuela a ninos por que no hay dinero y el acceso a salud de los pobres es malo.
- ▶ La red de apoyo social del gobierno tenia 54 programas en 12 ministerios
- ▶ Se diseno PATH (Programme of Advancement through Health and Education), un CTT, para reemplazar 3 de esos programas

[Levy and Ohls, 2010]

Modelos Lineales

PATH tenia dos componentes: 1) Asistencia por nino para salud; 2) Asistencia para educacion condicionada que el nino fuera a centros de salud cada 6 meses.



[Levy and Ohls, 2010]

Data: Encuesta en periodo 0 (baseline) y 18 meses despues

Discontinuidad: El ministerio de seguridad social usaba un threshold fijo sobre un puntaje de elegibilidad (e.g. ingresos*numero de hijos). Debajo se era beneficiario de CTT, arriba no.

[Levy and Ohls, 2010]

Modelos Lineales

La regresion discontinua es

$$Outcome_{ij} = \beta_0 + \beta_1 PART_j + \beta_2 f(Score_j) + \beta_3 X_{ij} + \beta_4 APPDATE_j + u_{ij}$$

$PART_j$ Dummy de participacion o no en PATH para household j

$f(Score_j)$ Funcion lineal del puntaje de elegibilidad

X_{ij} Caracteristicas del individuo i en household j (e.g. demograficos)

$APPDATE$ Fecha de aplicacion al PATH

[Levy and Ohls, 2010]

Modelos Lineales

En baseline el grupo tratado y control tienen outcomes similares

Characteristics	Participant group	Comparison group
Education^a		
Enrolled in school	97	96
In a typical 4-week period, how many days was child sent to school? ^b		
Mean (days)	18	18
Median (days)	20	20
Attending 0–14 days	10	10
Attending 15–16 days	16	16
Attending 17–20 days	73	74
During the reference period, how many days was child sent to school?		
Mean (days)	18	18
Median (days)	20	20
Attending 0–14 days	7	7
Attending 15–16 days	16	16
Attending 17–20 days	77	77
Health		
Visited health practitioner last 12 months		
Children 0–6 years old	70	71
Elderly 60 and over	75	77
Visited health practitioner last 12 months for check-up		
Children 0–6 years old	57	59
Elderly 60 and over	51	58
If currently pregnant, visiting health practitioner for antenatal care?	73	81
<i>N</i>	12,175	11,886

^aAmong children 7–17 years old.

^bFor those enrolled in school.

Note: Data are percentages, unless otherwise noted.

Modelos Lineales

El outcome es días de asistencia a la escuela. Usando diferentes especificaciones no se pierde la significancia de participar en el CTT (Participant group): aumenta 0.506 días la asistencia con todos los controles.

Table 4. Impact estimates on attendance.

	Specification			
	1	2	3	4
<i>Reference period</i>				
Participant group	0.483*** (0.14)	0.385* (0.22)	0.443* (0.23)	0.506** (0.22)
Eligibility score		-0.01 (0.02)	-0.008 (0.02)	-0.001 (0.02)
Baseline school attendance			0.424*** (0.02)	0.406*** (0.02)
Household controls	No	No	No	Yes
N	7,751	7,751	6,850	6,819

[Levy and Ohls, 2010]

Sin embargo, los autores no encontraron efectos significativos en otros outcomes como notas en la escuela o pasar el agno.

[Levy and Ohls, 2010]

Regression discontinuity



Causalidad: Controles sintéticos

Basado en: The mixtape de Scott Cunningham

(<https://mixtape.scunning.com/synthetic-control.html>)

- ▶ Estudios comparativos: Contrafactual
- ▶ Un control sintético son pesos aplicados a unidades que no fueron tratadas para simular contrafactuales artificiales

Por ejemplo, ¿la migración baja salarios y empleos? ¿qué control usar?
¿otra ciudad que no recibió migración? ¿otro país? ¿otro barrio?

La técnica de control sintético usa una combinación ponderada de varias posibles unidades no tratadas (más adelante detalles). Esto ayuda a:

1. Evitar selección subjetiva y ad hoc del control
2. Que los efectos no sean del todo por varianza en el sampleo

Algunas ventajas

1. Los tratados y el control sintético son del mismo lugar y/o tiempo.
e.g. año (no hay extrapolación).
2. Se puede construir el contrafactual sin tener que ver outcomes futuros.
3. La ponderación a cada unidad de control es transparente.
4. Acerca los estudios cualitativos a ser cuantitativos al poder escoger controles sintéticos.

Nos interesa el outcome Y_{jt} ; j: unidad (1 tratada; demás no); t: tiempo.
El efecto causal en el grupo tratado es

$$Y_{1t} - \sum_{j=2}^{J+1} w_j^* Y_{jt}$$

Es la distancia en el outcome (e.g. desempleo) que tuvo el grupo tratado al control sintético.

Pero ¿cómo escoger los pesos w_j^* ?

Los pesos W se escogen minimizando la distancia en características post-intervención X de los grupos, que no deben afectarse por la intervención (e.g. %industria en la ciudad; tecnología; demografía; etc).

$$\|X_1 - X_0 W\|$$

No hay pesos negativos (pueden ser cero) y deben sumar uno.

¿Qué distancia usar? Esta es usual

$$\|X_1 - X_0 W\| = \sqrt{(X_1 - X_0 W)' V (X_1 - X_0 W)}$$

Donde V es una matriz diagonal $k \times k$; k = número de características. La diagonal es v_1, \dots, v_k ; v_m es el peso dado a la característica m .

W y v deben reproducir el comportamiento del outcome para la unidad tratada cuando no se ha tratado.

Modelos Lineales

Luego de algo de álgebra y en versión sumatoria, buscamos minimizar

$$\sum_{m=1}^k v_m \left(X_{1m} - \sum_{j=2}^{J+1} w_j^* X_{jm} \right)^2$$

Y escogemos v_m que minimice la distancia al outcome (MSE) (T_0 : tiempo de intervención)

$$\sum_{t=1}^{T_0} \left(Y_{1t} - \sum_{j=2}^{J+1} w_j^*(V) Y_{jt} \right)^2$$

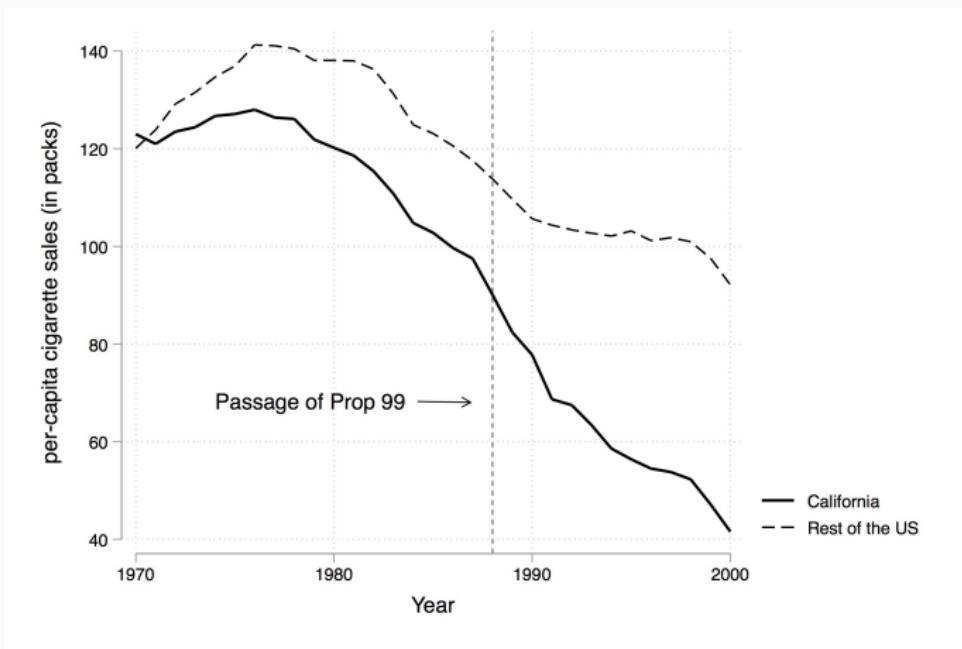
Modelos Lineales

Ejemplo: Propuesta 99 en California.

La propuesta 99 aumentó impuestos a los cigarrillos, ordenanzas de aire limpio, campañas anti-fumar, se dio más ingresos públicos a salud, entre otros.

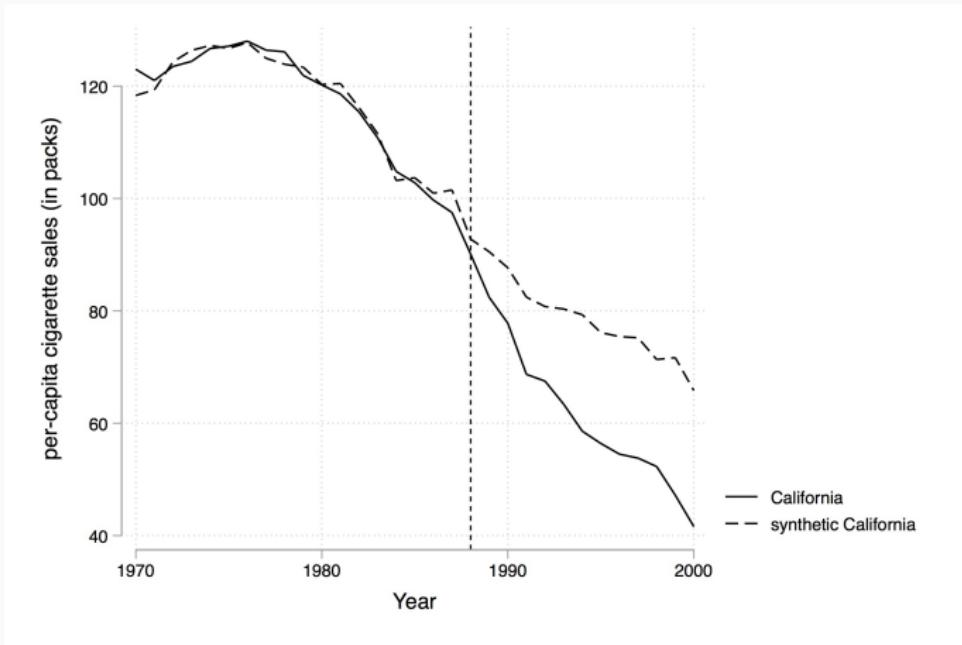
Modelos Lineales

Es difícil saber si la propuesta funcionó pues venían bajando las ventas en todo EE.UU.



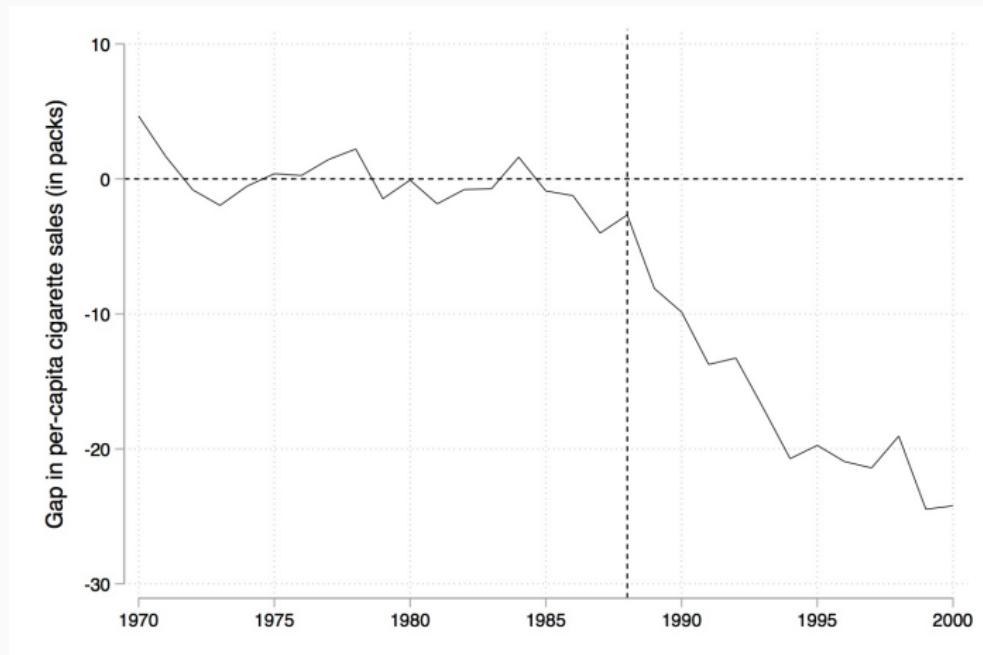
Modelos Lineales

Abadie, Diamond, and Hainmueller (2010) crearon controles sintéticos y es más claro que la propuesta 99 sí afectó las ventas de cigarrillo.



Modelos Lineales

Otra forma de verlo. Acá se ve el gap entre ambas líneas.



Modelos Lineales

Podemos ver las características del control sintético (con pesos óptimos), California real, y el promedio de los estados para el control sintético (sin pesos optimos). No hay estadísticas por que son solo dos grupos.

Variables	Real California	Synthetic Calif.	Avg. of 38 Control States
Ln(GDP per capita)	10.08	9.86	9.86
Percent aged 15–24	17.40	17.40	17.29
Retail price	89.42	89.41	87.27
Beer consumption per capita	24.28	24.20	23.75
Cigarette sales per capita 1988	90.10	91.62	114.20
Cigarette sales per capita 1980	120.20	120.43	136.58
Cigarette sales per capita 1975	127.10	126.99	132.81

Table 10.1: Balance table

All variables except lagged cigarette sales are averaged for the 1980–1988 period. Beer consumption is averaged 1984–1988.

Si son dos grupos ¿Cómo calcular estadísticas significativas?

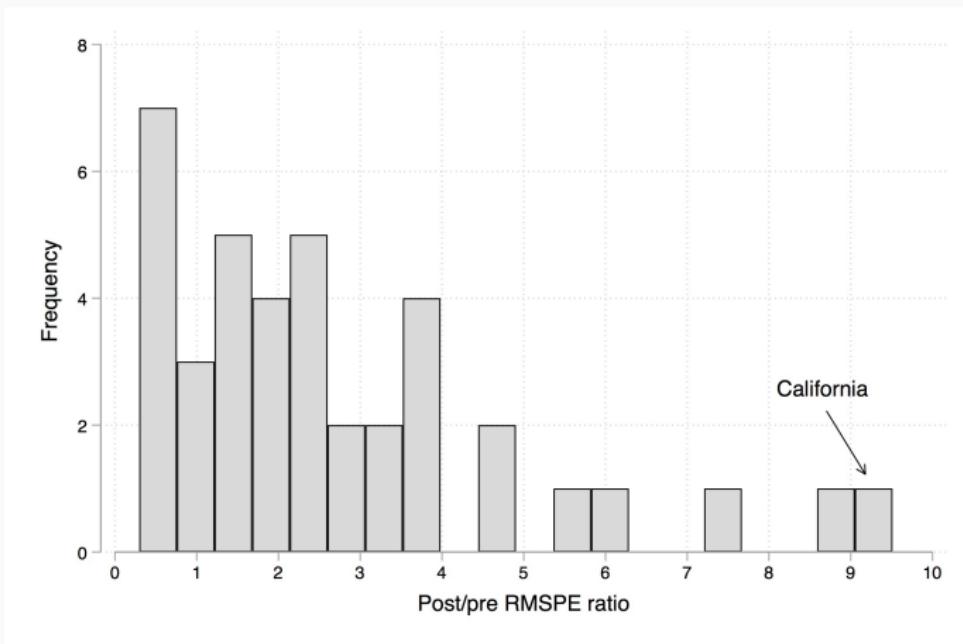
1. Aplicar el método de control sintético a cada uno de los estados donantes del control sintético original. Llámelos placebo.
2. Calcular root mean squared prediction error (RMSPE) para los placebo antes y después del tratamiento:

$$\left(\frac{1}{T - T_0} \sum_{t=T_0+t}^T \left(Y_{1t} - \sum_{j=2}^{J+1} w_j^* Y_{jt} \right)^2 \right)^{\frac{1}{2}}$$

3. Analicé el ratio de RMSPE (pre-post) e.g. de mayor a menor o un histograma.

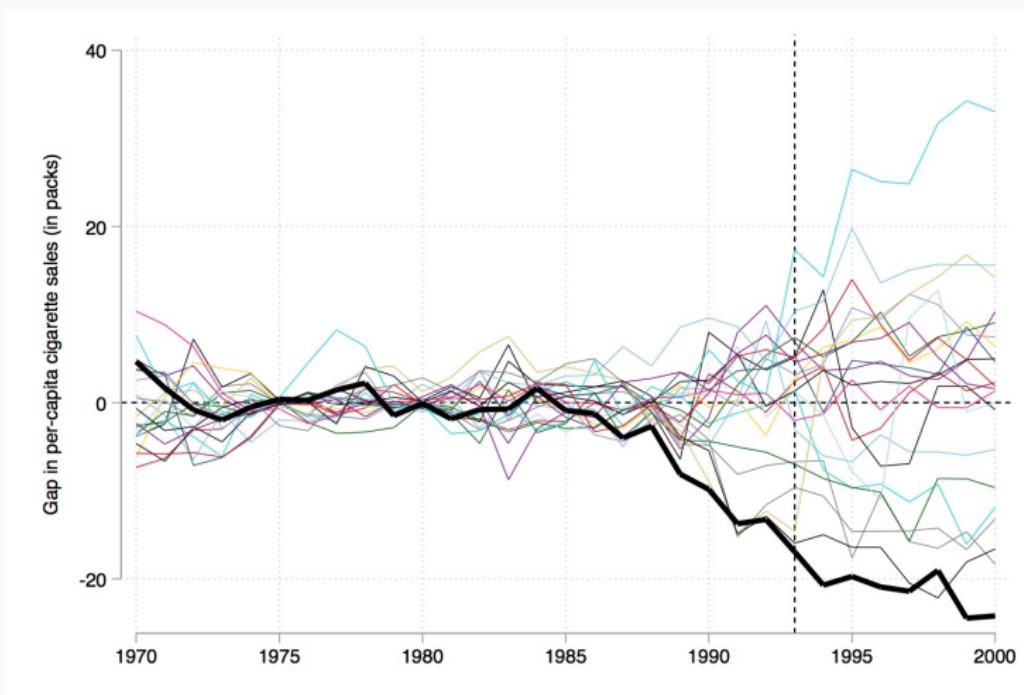
Modelos Lineales

En efecto, los placebo tienen mucho menor RMSPE ratio. Es decir, post-treatment es más notable en California real que en los placebo ($p=0.026$).



Modelos Lineales

Otra forma de verlo. Dado que tenemos placebos, podemos hacer la gráfica del gap. La línea negra gruesa es California original.

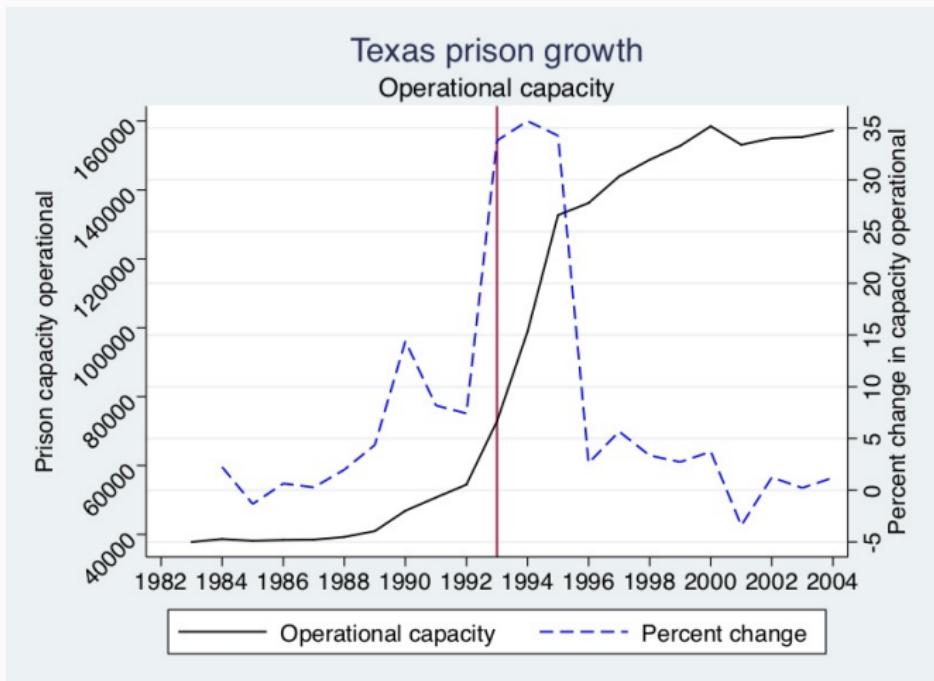


Modelos Lineales

Ejemplo con código: Construcción de prisiones y encarcelamiento de hombres negros.

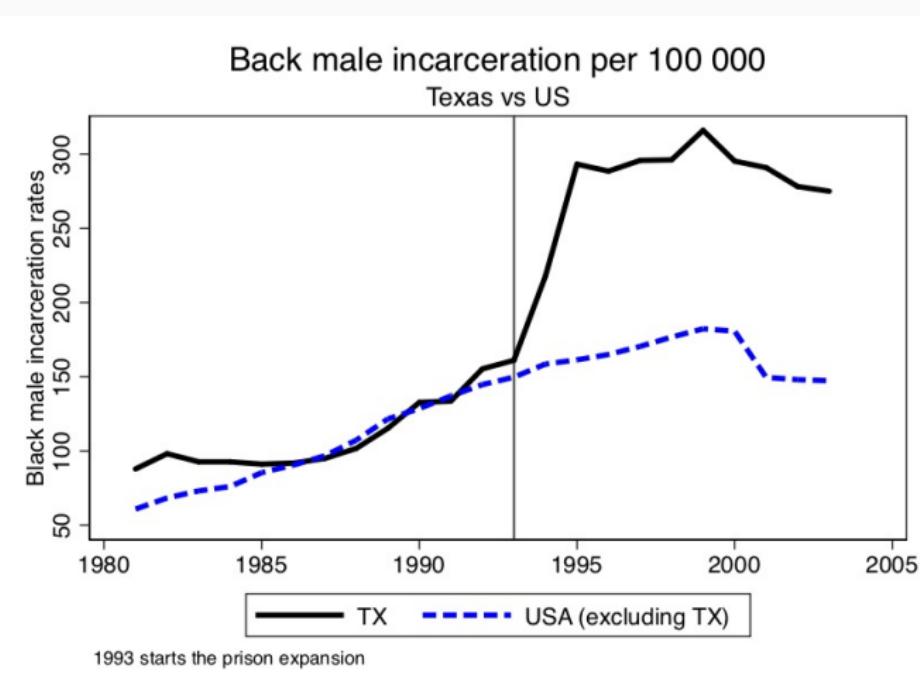
Modelos Lineales

Para reducir hacinamiento Texas construyó muchas cárceles (capacidad operacional = camas)



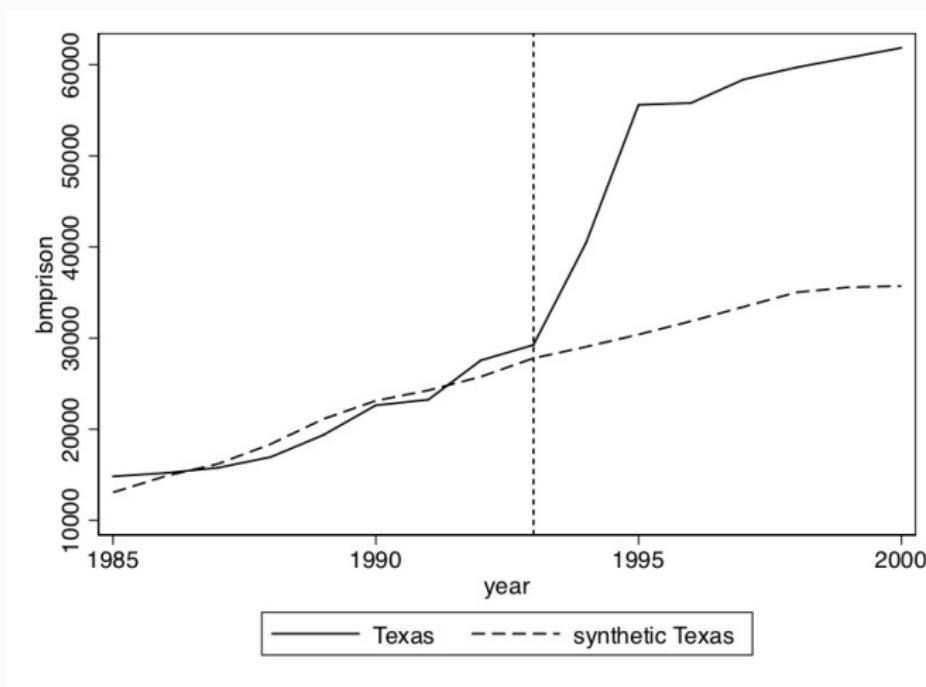
Modelos Lineales

Para la misma época creció mucho el encarcelamiento de hombres negros.



Modelos Lineales

Si construimos un control sintético se ve el mismo patrón (ver R, sección Synthetic control).



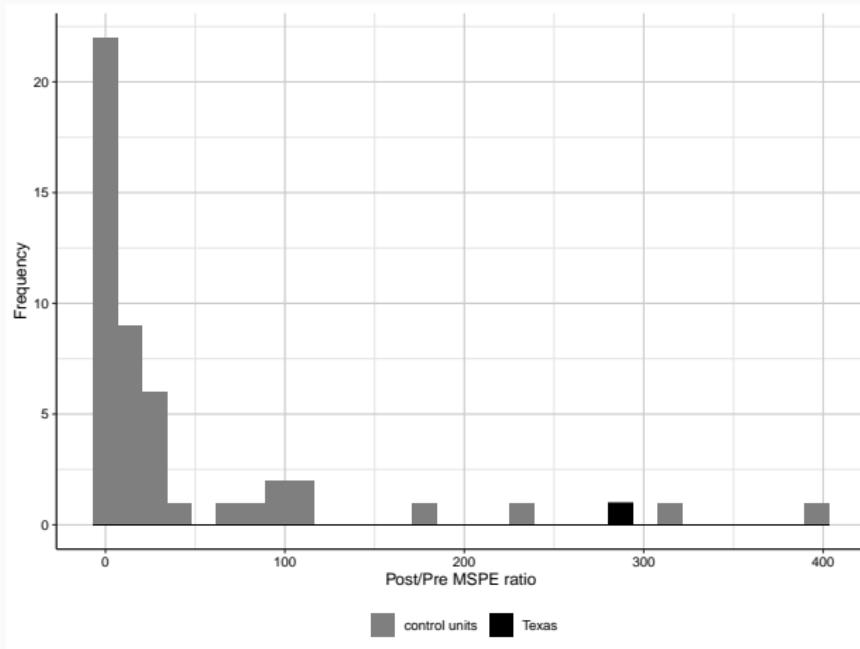
Modelos Lineales

Otra visualización: gap de la anterior gráfica (ver R, sección Synthetic control).



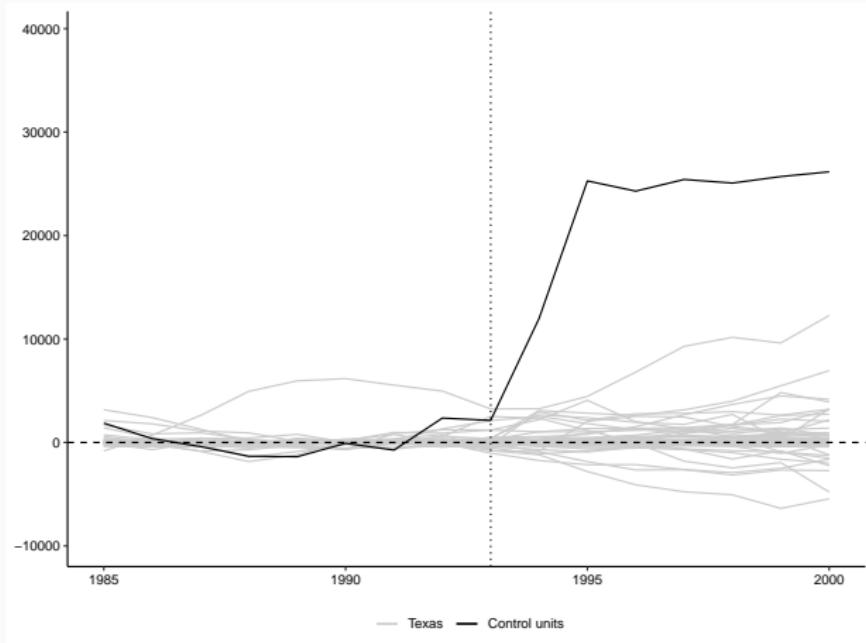
Modelos Lineales

Otra visualización: ratio de RMSPE (ver R, sección Synthetic control).
Texas tuvo un post-treatment outcome extremo.



Modelos Lineales

Otra visualización: Distribución Gaps (ver R, sección Synthetic control).
Texas tuvo un post-treatment outcome extremo.



Causalidad: Variables instrumentales

Aplicaciones:

- ▶ Simultaneidad (y causa x; ademas x causa y)
- ▶ Errores de medicion
- ▶ Variables omitidas.

Variable omitida / Causalidad

Problemas de especificación i.e. el modelo no refleja el proceso causal que generó la data

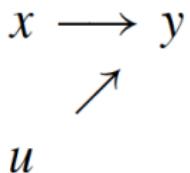
- ▶ La forma funcional de y no es lineal
- ▶ Endogeneidad (correlacion no causalidad)
- ▶ Variable omitida
- ▶ Heterogeneidad de parametros (random effects)

Modelos Lineales

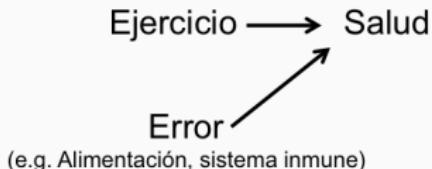
¿Qué asume este modelo? x causa y en proporción dada por β

$$y = \beta x + u$$

¿Están de acuerdo?



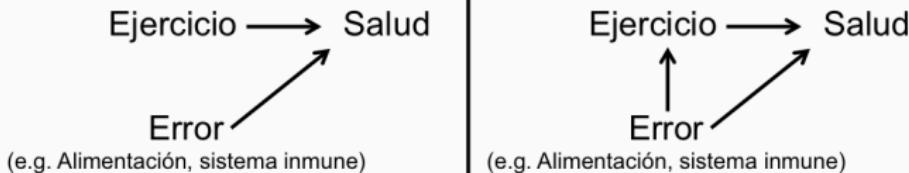
¿Están de acuerdo?



Modelos Lineales

¿Cuál es la estructura causal?

¿Qué implica para la regresión la de la derecha?



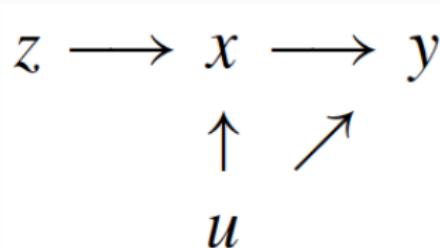
x afecta u (el ruido). El dato y cambia en función de ambos

$$y = \beta x + u(x)$$

$$\frac{dy}{dx} = \beta + \frac{du}{dx}$$

El estimador OLS de β está sesgado, incluye información proveniente del error. No es posible interpretarlo como fuerza causal, solo como correlación.

Solución: variable instrumental Z



Características

- z no está correlacionada con u
- z se correlaciona con x
- z se relaciona con y vía x

En términos generales, hay que buscar un instrumento si $\text{Cov}(x, u) \neq 0$ en

$$y = \beta_0 + \beta_1 x + u$$

Modelos Lineales

Ejemplos variables instrumentales (z)

¿Por qué hallazgo de pozos? Use las condiciones de VI

- $y = \text{precio}_{\text{petroleo}}$; $x = \text{demanda}_{\text{petroleo}}$; $z : \text{hallazgo pozos}$

QUIZ

¿Por qué cercanía a la escuela? Use las condiciones de VI

- $y = \text{Salario}$; $x = \text{Educacion}_{\text{nivel}}$; $z : \text{cercania escuela}$

Estimador β en $y = \beta x + u$ (min. squared error).

Asume que z se correlaciona con x y no con el error

$$\hat{\beta}_{IV} = (z'x)^{-1} z'y$$

Asuma que un incremento en z aumenta 0.2 años de escolaridad (i.e. $\frac{dx}{dz} = 0,2$) y \$500 el salario (i.e. $\frac{dy}{dz} = 500$)

Es decir, un año de escolaridad aumenta \$2500 el salario (i.e. $\frac{500}{0,2}$). Por lo tanto, β de educación es \$2500.

$$\hat{\beta} = \frac{dy/dz}{dx/dz}$$

Las derivadas son los β s de regresiones (e.g. $y = \beta z + u$), cuyos estimadores OLS son

$$\begin{aligned}\hat{\beta} &= \frac{(z'z)^{-1}z'y}{(z'z)^{-1}z'x} \\ &= (z'x)^{-1}z'y\end{aligned}$$

Modelos Lineales

Si el instrumento es binario (e.g tratado vs no tratado), se puede mostrar con la definición que el estimador es la siguiente pendiente,

$$\hat{\beta} = \frac{(\bar{y}_1 - \bar{y}_0)}{(\bar{x}_1 - \bar{x}_0)}$$

Por ejemplo, becar a un grupo y a otro no en primaria y ver su ingreso 5 años después de graduación. El efecto en ingreso sería,

$$\hat{\beta} = \frac{(\overline{Ingreso}_{beca} - \overline{Ingreso}_{no\ beca})}{(\overline{Ed.}_{beca} - \overline{Ed.}_{no\ beca})}$$

Ejemplo 2 binario: local average treatment effect (LATE).

ATE es el outcome promedio de ser tratado (e.g. asegurado) menos no serlo

$$ATE = E[y(1) - y(0)]$$

[Imbens and Angrist, 1994]

LATE es similar pero con mas restricciones. Hay un instrumento z (e.g. aleatorizacion), pero el individuo es libre de participar o no en el tratamiento o control.

Por ejemplo, z = loteria servicio militar. Si la loteria me salva, igual puedo decidir ir. En ese caso yo seria no asignado pero tratado.

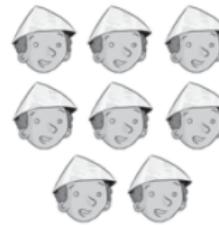
En LATE solo entran los que el instrumento asigna y son tratados i.e. compliers. En el ejemplo, lotería me salva y no voy ($w(1)=1$) + lotería no me salva y voy ($w(0)=0$)

$$LATE = E[y(1) - y(0)|compliers]$$

[Imbens and Angrist, 1994]

Modelos Lineales

Visualización. LATE es el efecto solo en los de la segunda fila i.e. \$50.

	Group assigned to treatment	Group not assigned to treatment	Impact
	Percent enrolled = 90% Average Y for those assigned to treatment = 110	Percent enrolled = 10% Average Y for those not assigned to treatment = 70	$\Delta\% \text{ enrolled} = 80\%$ $\Delta Y = \text{ITT} = 40$ $\text{LATE} = 40/80\% = 50$
Never enroll			—
Only enroll if assigned to treatment			
Always enroll			—

y: Ingreso; sombra: los inscritos
[Gertler et al., 2016]

Critica de LATE: no es siempre fácil identificar compliers. La data solo tiene la respuesta (y), condición de tratado (w), el instrumento (z).

Por ejemplo, y: vacunado en 1850; w: ingreso alto o medio; z: distancia a hospital, lo compliers son los que se vacunan si hay hospital cerca y no si lejos

¿Cómo saber con y, w, z si se es complier? e.g. Pepe es no vacuna + ingreso medio + cerca a hospital. No complier?

Recordar que una VI debe,

- ▶ Correlacionarse fuertemente con el regresor x (\sim *paso1*)
- ▶ No correlacionarse con el ruido ($u + e$)
- ▶ No correlacionarse con la variable y (solo vía x)

Método usual para aplicar VI: Regresión OLS en 2 pasos (2SLS).

Pasos:

1. $x = \beta z + u + \epsilon_1$ para obtener \hat{x}
2. $y = \beta \hat{x} + u + \epsilon_2$

Sección 2SLS



Variables instrumentales en salud

- Colera en Londres 1853-1854: $y : cholera; x : waterpurity$

Por que $colera = \beta_{agua\,pureza} + \epsilon$ no probaria que colera se transmite por agua?

[Grootendorst, 2007]

Los que tomaban agua impura tambien eran pobres y sus condiciones de vivienda eran peores. La colera podria venir de otra fuente no relacionada con pureza de agua.

Que instrumento usar, que no se correlacione con el error/no observables (condiciones de vivienda) y si con pureza de agua?

[Grootendorst, 2007]

Variables instrumentales en salud

- ▶ Colera en Londres 1853-1854:

y : cholera; x : water purity; z : proveedor de agua

En Londres de 1800 había 2 proveedores que recogían agua en dos puntos diferentes, uno mas contaminado. Los 2 proveían aleatoriamente los diferentes barrios.

¿Por qué z es un instrumento valido? Se correlaciona con x (pureza de agua)? Se correlaciona con error (condiciones de vida)?

[Grootendorst, 2007]

Variables instrumentales en salud

- ▶ Colera en Londres 1853-1854
- ▶ Esperar por cirugía es peligroso?:
 $y : recuperacion; x : tiempo de espera$

¿Por qué $recuperacion = \beta tiempo_{espera} + \epsilon$ no probaría que recuperación es peor si no se opera rápido?

[Grootendorst, 2007]

La recuperación depende también de la severidad de la enfermedad
¿Qué instrumento usar, que no se correlacione con el error/no
observables (severidad de la enfermedad) y sí con tiempo de espera?

[Grootendorst, 2007]

Variables instrumentales en salud

- ▶ Colera en Londres 1853-1854
- ▶ ¿Esperar por cirugía es peligroso?:
 y : recuperacion; x : tiempo de espera; z : Tipodesangre

Hay tipos de sangre que reciben cualquier donación (AB) y otros es más difícil (A-).

¿Por qué z es un instrumento valido? Se correlaciona con x (tiempo de espera)? Se correlaciona con error/no observables (severidad enfermedad)?

[Grootendorst, 2007]

Variables instrumentales en salud

- ▶ Colera en Londres 1853-1854
- ▶ ¿Esperar por cirugía es peligroso?
- ▶ ¿Gastar más en salud es bueno?: $y : \text{salud}$; $x : \text{gasto en salud}$

¿Por qué $\text{salud} = \beta \text{gasto}_{\text{salud}} + \epsilon$ no probaría que salud es mejor si se gasta más?

[Grootendorst, 2007]

Los que tienen enfermedades graves gastan más en salud
¿Qué instrumento usar, que no se correlacione con el error/no observables (severidad de la enfermedad) y sí con salud?

[Grootendorst, 2007]

Variables instrumentales en salud

- ▶ Colera en Londres 1853-1854
- ▶ ¿Esperar por cirugía es peligroso?
- ▶ ¿Gastar más en salud es bueno?:
 $y : \text{salud}; x : \text{gasto en salud}; z : \text{gastos salud en ultimos dias}$

Gastos de salud en ultimos días se refiere a los gastos relacionados antes de morir en toda la población

¿Por que z es un instrumento valido? ¿Se correlaciona con x (gasto en salud)? ¿Se correlaciona con error/no observables (severidad enfermedad)?

Variables instrumentales en salud

- ▶ Colera en Londres 1853-1854
- ▶ ¿Esperar por cirugía es peligroso?
- ▶ ¿Gastar más en salud es bueno?
- ▶ Efecto contexto físico en salud: y : *Expectativa de vida*; x : *educacion*

¿Por qué $\text{Expectativa}_{\text{vida}} = \beta_{\text{educacion}} + \epsilon$ no probaría que expectativa de vida mejora con más educación?

Los más educados pueden tener mejores ingresos
¿Qué instrumento usar, que no se correlacione con el error/no
observables (ingresos) y sí con expectativa de vida?

[Grootendorst, 2007]

Variables instrumentales en salud

- ▶ Colera en Londres 1853-1854
- ▶ ¿Esperar por cirugía es peligroso?
- ▶ ¿Gastar más en salud es bueno?
- ▶ Efecto contexto físico en salud: y : *Expectativa de vida*; x : *educacion*; z : *leyes de educación obligatoria*

Entre 1915 y 1939 muchos estados en EE.UU introdujeron en distintos momentos leyes que hacían la educación obligatoria

¿Por qué z es un instrumento válido? ¿Se correlaciona con x (educación)?

¿Se correlaciona con error/no observables (ingresos)?

Modelos Lineales

En tablas ... efecto de educación en mortalidad ... ¿Cómo interpretar?

		OLS	Instrumental Variable
Variables			
Data		Census	Census ^{(a)(b)(c)}
Method		WLS	2SLS
Level ^(d)		Aggregate ^(b)	Aggregate
Dependent variable		10-Year death rate	10-Year death rate
Individual characteristics	Education	-0.017** (0.004)	-0.051** (0.026)
	Female	-0.074** (0.003)	0.003 (0.005)
State-of-birth % Urban characteristics		1.0e-04 (9.4e-04)	-0.071** (0.004)
	0% Foreign	-5.6e-01	0.001

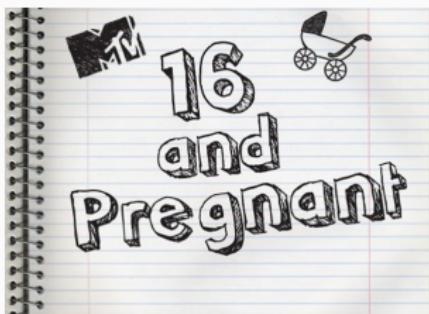
[Lleras-Muney, 2005]

Variables instrumentales en salud

- ▶ Colera en Londres 1853-1854
- ▶ ¿Esperar por cirugía es peligroso?
- ▶ ¿Gastar más en salud es bueno?
- ▶ Efecto contexto físico en salud

[Grootendorst, 2007]

2SLS variables instrumentales en salud

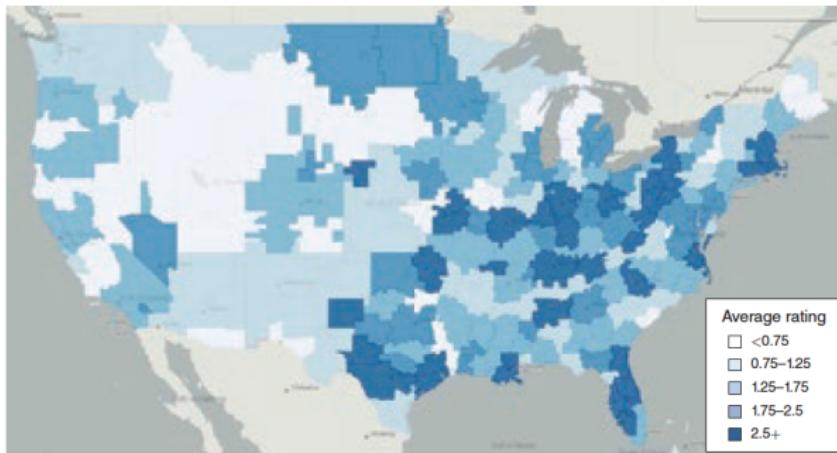


[Kearney and Levine, 2015]

- ▶ Reality de adolescentes embarazadas que empezó en el 2009 (<https://www.youtube.com/watch?v=LEjiM0sESus>)
- ▶ Que efecto tuvo en comportamiento sexual?

Modelos Lineales

El show fue popular (MTV produjo varias temporadas)



[Kearney and Levine, 2015]

Discusión:

¿Los medios causan comportamiento?

¿La relación comportamiento-medios es causal o sufre de endogenidad?

[Kearney and Levine, 2015]

Modelos Lineales

Usaron 2SLS. Primera etapa

$$\text{Rating}16P_j \times \text{post}_t = \beta_0 + \beta_1 \text{MTV}0809_j \times \text{post}_t + \beta_2 U_{jy} + X_{jy}\gamma + \theta_t + \lambda_{js} + \epsilon_{jt}$$

t: quarters; j: media markets (regiones); s: temporada del agno; y: agno

$\text{Rating}16P_j$: rating edades 12-24; post_t : 0 fechas antes de empezar show, despues es el rating

$\text{MTV}0809 \times \text{post}_t$: instrumento, rating shows MTV

U_{jy} : desempleo; X_{jy} : raza; θ_t : efectos fijos quarter; λ_{js} : efecto fijo temporada region (nacimientos son ciclicos)

Segunda etapa

$$\ln(B_j t) = \beta_0 + \beta_1 \widehat{\text{Rating}16P_j} \times \text{post}_t + \beta_2 U_{jy} + X_{jy}\gamma + \theta_t + \lambda_{js} + \epsilon_{jt}$$

$B_j t$: tasa de nacimientos

[Kearney and Levine, 2015]

Discusión: ¿Ratings de MTV en el 2008 y 2009 es una variable instrumental?

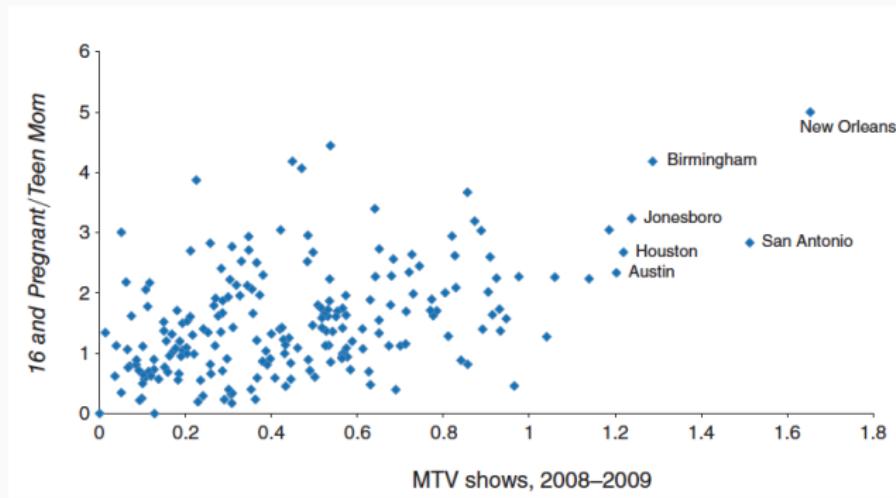
¿se correlaciona con ratings de 16P?

¿se correlaciona con el ruido o variables no observadas que afecten tasas de nacimiento?

[Kearney and Levine, 2015]

Modelos Lineales

Se correlaciona con ratings de 16P



ejes x & y son ratings
[Kearney and Levine, 2015]

Modelos Lineales

Aumento en ratings disminuía tasa de nacimientos (aprox. -2.37%; ver instrumental variable IV)

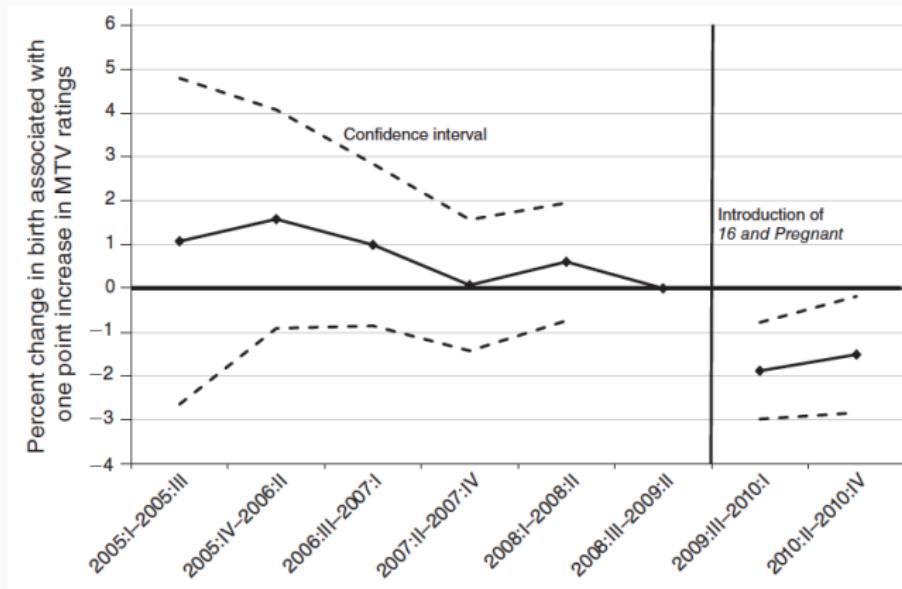
TABLE 1—ESTIMATES OF THE IMPACT OF *16 AND PREGNANT* RATINGS ON TEEN BIRTH RATES

	OLS (1)	First stage (2)	IV (3)	Reduced form (4)
Dependent variable:	ln(birth rate)	<i>16 and Pregnant</i> ratings	ln(birth rate)	ln(birth rate)
<i>16 and Pregnant</i> ratings	-1.020* (0.552)		-2.368** (0.942)	
MTV ratings 2008–2009		1.511*** (0.204)		-3.581** (1.512)
Unemployment rate	-1.440*** (0.401)	-0.001 (0.026)	-1.487*** (0.375)	-1.485*** (0.432)
<i>F</i> -statistic on omitted instrument		48.1		

[Kearney and Levine, 2015]

Modelos Lineales

Visualización: antes del show el cambio en la tasa de nacimientos no es diferente a cero (ver confidence intervals). Después del show es negativa.



[Kearney and Levine, 2015]

Modelos Lineales

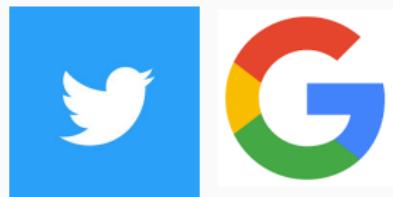
El efecto ocurrió en todas las edades

TABLE 2—ESTIMATES OF THE IMPACT OF *16 AND PREGNANT* RATINGS ON BIRTH RATES, BY DEMOGRAPHIC GROUP

	Age 15–19	Age 20–24	Age 25–29	Age 30–34	White, non-Hispanic Age 15–19	Black, non-Hispanic Age 15–19	Hispanic Age 15–19
<i>Panel A. IV</i>							
Rating	-2.368** (0.942)	-2.422** (0.993)	-1.606** (0.817)	-0.328 (0.689)	-2.413** (1.075)	-0.140 (1.380)	-3.782 (2.063)
Unemployment rate	-1.487*** (0.375)	-1.810*** (0.308)	-0.800*** (0.229)	-0.910*** (0.146)	-1.399*** (0.329)	-0.243 (1.049)	-2.512*** (0.784)

[Kearney and Levine, 2015]

Modelos Lineales



[Kearney and Levine, 2015]

$$searchBC_t = \beta_0 + \beta_1 NewEpisode_t + X_t\gamma + \epsilon_t$$

$searchBC_t$: búsqueda de birth control

Modelos Lineales

Cuando se estrenaba un capítulo las busquedas y tendencias de 16P y birth control se correlacionaban, especialmente a nivel estatal (Panel B)

TABLE 4—IMPACT OF SEARCHES/TWEETS ABOUT *16 AND PREGNANT* ON SEARCHES/TWEETS ABOUT BIRTH CONTROL AND ABORTION

	Google Trends: search index			Twitter: ln(tweet rate)	
	“How get birth control”	“How get birth control pill”	“How get abortion”	“Birth control”	“Abortion”
<i>Panel A. National, high frequency data</i>					
Search index/ln(tweet rate) <i>16 and Pregnant</i>	0.012 (0.026)	0.069 (0.042)	-0.074** (0.036)	0.077** (0.034)	0.064*** (0.025)
Number of weeks (searches)/days (tweets)	209	209	209	336	336
<i>Panel B. State-level, lower frequency data</i>					
Search index/ln(tweet rate) <i>16 and Pregnant</i>	0.751*** (0.127)	—	0.505** (0.239)	0.137** (0.054)	-0.087 (0.075)
Number of states/periods	30	—	24	537	537

[Kearney and Levine, 2015]

En suma, ver las dificultades que una madre adolescente experimenta tuvo un efecto en las tasas de nacimiento.

Uso de ratings como variable instrumental. ¿Valido o no?

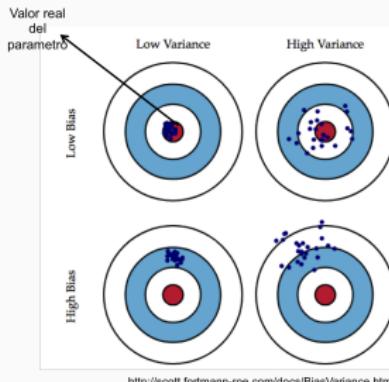
[Kearney and Levine, 2015]

Modelos Lineales

¿Qué es un instrumento bueno?

1. Es exogeno a la ecuación i.e $\text{Cov}(z, u) = 0$
2. Es relevante i.e. $\text{Cov}(z, x) > 0$ y preferiblemente alta. Si es baja el estimador del efecto de x en y va a estar sesgado, lejos del valor real, porque el instrumento no añade información nueva dada la condición 1.

Visualizacion sesgo (y varianza)



¿Qué es un instrumento débil?

- ▶ La correlación r_{xz}^2 entre x (predictor endógeno) & z (instrumento) es pequeña.
- ▶ En la primera etapa del 2SLS el R^2 o F son pequeños i.e. de $x = \beta z + u + \epsilon$
- ▶ La mejora del ajuste, e.g. R^2 o F , que trae z es baja e.g. el parcial R^2 es pequeño
- ▶ El instrumento z se correlaciona con el error i.e. es endógeno.

¿Qué costos traen los instrumentos? Potencial mayor varianza en el estimador

Varianza OLS

$$Var(\hat{\beta}_{OLS}) = \frac{\sigma_u^2}{N\sigma_w^2}$$

Varianza variable instrumental (VI)

$$Var(\hat{\beta}_{VI}) = \frac{\sigma_u^2}{N\sigma_w^2} \times \frac{1}{\rho_{z,x}^2}$$

Varianza relativa OLS vs VI depende de la correlación entre z & x. ¿Qué pasa si es baja?

$$\frac{se(\hat{\beta}_{VI})}{se(\hat{\beta}_{OLS})} = \frac{1}{|\rho_{z,x}|}$$

Modelos Lineales

Educación católica y puntajes. IV via 2SLS mayores s.e.

Table 2 OLS and 2SLS estimates of Catholic Schooling Effects NELS:88 and NLS-72 Weighted, (Huber-White Standard Errors in Parentheses)			
	Excluded Instruments		
	(1) <i>Catholic</i> (C_i)	(2) <i>Distance</i> (D_i)	(3) <i>Catholic</i> \times <i>Distance</i> ($C_i \times D_i$)
12th Grade Reading Score (NELS:88)			
OLS	1.16 (0.37)	1.03 (0.37)	1.14 (0.38)
2SLS	1.40 (1.54)	-1.09 (1.84)	1.24 (1.82)
12th Grade Math Score (NELS:88)			
OLS	1.03 (0.31)	1.00 (0.31)	0.92 (0.32)
2SLS	2.64 (1.21)	2.43 (1.45)	-2.63 (1.57)
12th Grade Reading Score (NLS-72)			
OLS	2.06 (0.34)	2.54 (0.37)	2.50 (0.36)
2SLS	-1.34 (0.99)	8.69 (4.53)	0.50 (2.32)
12th Grade Math Score (NLS-72)			
OLS	1.52 (0.33)	1.77 (0.35)	1.71 (0.36)
2SLS	-0.07 (0.96)	11.05 (4.47)	-3.94 (2.27)

¿Distancia a escuela católica es un buen instrumento para puntajes?
Altonji et al argumentan que no en parte por los s.e. relativos a OLS
($\rho_{z,x}$ ¿pequeño o grande?)
[Altonji et al., 2005]

Modelos Lineales

Mortalidad y educación. Misma critica que Altonji et al; IV es más ruidoso? No es tan claro.

Lo que es claro es que IV aumenta varianza del estimador. ¿Bueno o malo?

		OLS	Instrumental Variable
Variables			
Data	Census		
Method	WLS		
Level ^(d)	Aggregate ^(b)		
Dependent variable	10-Year death rate		
Individual characteristics	-0.017** (0.004)	-0.017** (0.004)	-0.051** (0.026)
Female	-0.074** (0.003)		0.003 (0.005)
State-of-birth % Urban characteristics	1.0e-04 (9.4e-04)		-0.071** (0.004)
R.C. Environment	-5.6e-04		0.001

[Lleras-Muney, 2005]

Además del ruido, hay que tener en cuenta el sesgo de OLS no desaparece con instrumentos débiles.

Asuma el siguiente sistema (z: variable instrumental; x: endógena; y: outcome),

$$x = \lambda u + \epsilon$$

$$z = \gamma \epsilon + v$$

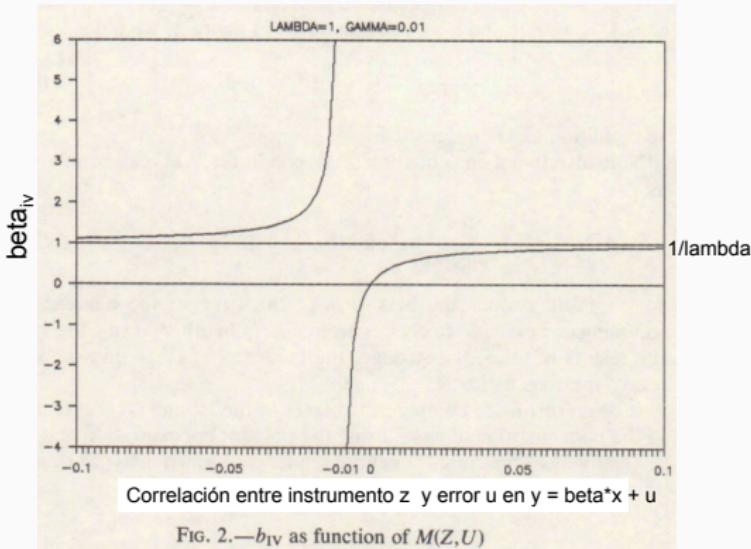
$$y = \beta x + u$$

γ determina la correlación del instrumento z con x (queremos que sea alta).

λ determina la correlación de x con el error u (queremos que sea baja).
[Nelson and Startz, 1990]

Modelos Lineales

Si la correlación entre z, x ($\sim \gamma$) es baja, el estimativo de β vía instrumentos depende de la correlación entre z, u , & tiende a $(\frac{1}{\lambda})$.



[Nelson and Startz, 1990]

Modelos Lineales

Otra forma de verlo:

Con correlaciones bajas entre z , x ($\sim \gamma$) el sesgo del estimativo de β hacia $\frac{1}{\lambda}$, se vuelve más fuerte

TABLE 1 Estimated Fractiles for b_{IV} with $\lambda = 1$, $n = 100$
(Based on 500 Replications)

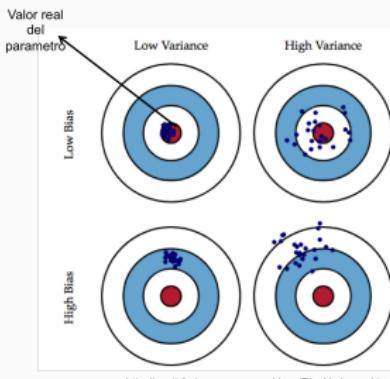
Fractiles	$\lambda = 1:$ b_{OLS}	$\gamma = .05$		$\gamma = .01$		$\gamma = .001$	
		b_{IV}	ASY	b_{IV}	ASY	b_{IV}	ASY
1	.43	-27.45	-5.14	-2.45	-25.7	-1.03	-257
10	.47	-1.63	-2.56	.59	-12.8	.77	-128
50	.52	.63	.00	.95	.00	.97	0
90	.56	2.49	2.56	1.29	12.8	1.20	128
99	.59	47.76	5.14	3.07	25.7	5.98	257

NOTE.—ASY column gives fractiles implied by asymptotic distribution.

[Nelson and Startz, 1990]

Modelos Lineales

En suma, los instrumentos pueden aumentar ruido en el estimativo y producir sesgos si las correlaciones no son fuertes



Sección Specification & overidentification test



-  Altonji, J. G., Elder, T. E., and Taber, C. R. (2005).
An evaluation of instrumental variable strategies for estimating the effects of catholic schooling.
Journal of Human resources, 40(4):791–821.
-  Christelis, D., Georgarakos, D., and Sanz-de Galdeano, A. (2020).
The impact of health insurance on stockholding: A regression discontinuity approach.
Journal of Health Economics, 69:102246.
-  Clark, D. and Royer, H. (2013).
The effect of education on adult mortality and health: Evidence from britain.
American Economic Review, 103(6):2087–2120.

References ii

-  Galiani, S., Gertler, P., and Schargrodsky, E. (2005).
Water for life: The impact of the privatization of water services on child mortality.
Journal of political economy, 113(1):83–120.
-  Gertler, P. J., Martinez, S., Premand, P., Rawlings, L. B., and Vermeersch, C. M. (2016).
Impact evaluation in practice.
The World Bank.
-  Grootendorst, P. (2007).
A review of instrumental variables estimation of treatment effects in the applied health sciences.
Health Services and Outcomes Research Methodology, 7(3-4):159–179.

-  Hansen, B. (2015).
Punishment and deterrence: Evidence from drunk driving.
American Economic Review, 105(4):1581–1617.
-  Imbens, G. W. and Angrist, J. D. (1994).
Identification and estimation of local average treatment effects.
Econometrica, 62(2):467–475.
-  Kearney, M. S. and Levine, P. B. (2015).
Media influences on social outcomes: The impact of mtv's 16 and pregnant on teen childbearing.
American Economic Review, 105(12):3597–3632.

-  Levy, D. and Ohls, J. (2010).
Evaluation of jamaica's path conditional cash transfer programme.
Journal of Development Effectiveness, 2(4):421–441.
-  Lleras-Muney, A. (2005).
The relationship between education and adult mortality in the united states.
The Review of Economic Studies, 72(1):189–221.
-  Nelson, C. and Startz, R. (1990).
The distribution of the instrumental variables estimator and its t-ratio when the instrument is a poor one.
The Journal of Business, 63(1):S125–40.

-  Nilsson, J. P. (2017).
Alcohol availability, prenatal conditions, and long-term economic outcomes.
Journal of Political Economy, 125(4):1149–1207.