Econometría I
Temas Adicionales MCO

Ramiro de Elejalde

Facultad de Economía y Finanzas
Universidad Alberto Hurtado

Outline

Error de medida

Mínimos cuadrados generalizados (GLS)

Validación interna y validación externa

Referencias: Cap. 4.3.2 y 4.4 de Wooldridge, cap. 8 y 9 de Stock and Watson, y cap. 3.2.3 Angrist and Pischke.

Error de medida

Error de medida en la variable dependiente

- Modelo de regresión

$$y^* = \beta_0 + \beta_1 x_1 + ... + \beta_K x_K + v,$$

donde $\text{Cov}(v, x_j) = 0$ para todo j.

- No observamos y^* pero observamos un medida de y^* dada por y.
- Definimos el error de medida poblacional como,

$$e_0 = y - y^*.$$

El modelo que podemos estimar es,

$$y = \beta_0 + \beta_1 x_1 + ... + \beta_K x_K + v + e_0.$$

- ¿Es la estimación MCO consistente?

Error de medida en la variable dependiente

- Modelo de regresión

$$y^* = \beta_0 + \beta_1 x_1 + ... + \beta_K x_K + v,$$

donde $\text{Cov}(v, x_j) = 0$ para todo j.

- No observamos y^* pero observamos un medida de y^* dada por y.
- Definimos el error de medida poblacional como,

$$e_0 = y - y^*.$$

El modelo que podemos estimar es,

$$y = \beta_0 + \beta_1 x_1 + ... + \beta_K x_K + v + e_0.$$

- ¿Es la estimación MCO consistente? Si, si el error de medida no está correlado con las x's, $\text{Cov}(x_j, e_0) = 0$ para todo j.
- Mayor varianza del error: Si $\text{Cov}(v, e_0) = 0$, $\text{Var}(v + e_0) = \text{Var}(v) + \text{Var}(e_0) \geq \text{Var}(v)$.

Error de medida en la variable explicativa

- Modelo de regresión

$$y = \beta_0 + \beta_1 x_1^* + ... + \beta_K x_K + v,$$

donde $\mathrm{Cov}(v, x_j) = 0$ para todo $j \neq 1$ y $\mathrm{Cov}(v, x_1^*) = 0$.

- No observamos x_1^* pero observamos un medida de x_1^* dada por x_1. Suponemos que $\mathrm{Cov}(x_1, v) = 0$.
- Definimos el error de medida poblacional como,

$$e_1 = x_1 - x_1^*,$$

donde $\mathbb{E}(e_1) = 0$. Dados los supuestos anteriores, $\mathrm{Cov}(e_1, v) = 0$.

El modelo que podemos estimar es,

$$y = \beta_0 + \beta_1 x_1 + ... + \beta_K x_K + (v - \beta_1 e_1).$$

- ¿Es la estimación MCO consistente?

Error de medida en la variable explicativa

¿Es la estimación MCO consistente? Suponemos que $\text{Cov}(x_j, e_1) = 0$ para todo $j \neq 1$.
El supuesto clave es la relación entre e_1 con x_1 y x_1^*

- Error de medida no está correlado con la variable observada: $\text{Cov}(x_1, e_1) = 0$.

  Es consistente pero hay una pérdida de eficiencia.

- Error de medida clásico: Error de medida no está correlado con la variable NO observada, $\text{Cov}(x_1^*, e_1) = 0$.

  No es consistente $\implies$ Sesgo de atenuación

Error de medida no está correlado con la variable observada: $\mathrm{Cov}(x_1, e_1) = 0$

$$\begin{aligned}
\mathrm{Cov}(x_1, u) &= \mathrm{Cov}(x_1, (v - \beta_1 e_1)), \\
&= \mathrm{Cov}(x_1, v) - \beta_1 \mathrm{Cov}(x_1, e_1), \\
&= 0.
\end{aligned}$$

- MCO es consistente pero la varianza del error es mayor $\mathrm{Var}(v - \beta_1 e_1) = \mathrm{Var}(v) + \beta_1^2 \mathrm{Var}(e_1)$.

Error de medida clásico: $\text{Cov}(x_1^*, e_1) = 0$

$$\begin{aligned}
\text{Cov}(x_1, e_1) &= \text{Cov}((x_1^* + e_1), e_1), \\
&= \text{Cov}(x_1^*, e_1) + \text{Var}(e_1), \\
&= \text{Var}(e_1).
\end{aligned}$$

- Entonces $\text{Cov}(x_1, (v - \beta_1 e_1)) = -\beta_1 \text{Var}(e_1)$ y MCO es inconsistente para todos los β's.
- Sesgo asintótico? $\implies$ Sesgo de atenuación.

Error de medida clásico: $\text{Cov}(x_1^*, e_1) = 0$

- Sesgo asintótico? $\implies$ Sesgo de atenuación.

  Demostración (sketch para el modelo de regresión simple $y = \beta_0 + \beta_1 x_1^* + v$):

$$
\begin{aligned}
\text{plim}\, \hat{\beta}_1 &= \frac{\text{Cov}(x_1, y)}{\text{Var}(x_1)}, \\
&= \frac{\text{Cov}(x_1, \beta_0 + \beta_1 x_1 + (v - \beta_1 e_1))}{\text{Var}(x_1)}, \\
&= \beta_1 + \frac{\text{Cov}(x_1, (v - \beta_1 e_1))}{\text{Var}(x_1)}, \\
&= \beta_1 - \beta_1 \frac{\text{Var}(e_1)}{\text{Var}(x_1)}, \\
&= \beta_1 \left(1 - \frac{\text{Var}(e_1)}{\text{Var}(x_1)}\right), \\
&= \beta_1 \left(\frac{\text{Var}(x_1^*)}{\text{Var}(x_1^*) + \text{Var}(e_1)}\right).
\end{aligned}
$$

Mínimos cuadrados generalizados (GLS)

Mínimos cuadrados generalizados (GLS)

$$y_i = \beta_0 + \beta_1 x_{1i} + ... + \beta_K x_{Ki} + u_i = x_i' \beta + u_i,$$

donde y_i es una variable aleatoria, x_i es un vector aleatorio de $K \times 1$ y u_i es una variable aleatoria.

Mínimos cuadrados generalizados (GLS)

$$y_i = \beta_0 + \beta_1 x_{1i} + \ldots + \beta_K x_{Ki} + u_i = x_i'\beta + u_i,$$

donde y_i es una variable aleatoria, x_i es un vector aleatorio de $K \times 1$ y u_i es una variable aleatoria.

- Supuestos:

GLS1 Muestra aleatoria de tamaño $N \implies \{y_i, x_i\}_{i=1}^N$ i.i.d.,

GLS2 Independencia en media: $\mathbb{E}(u_i|x_i) = 0$,

GLS3 $\text{Var}(u_i|x_i) = \sigma^2(x_i)$,

GLS4 $\mathbb{E}\left(\frac{x_i x_i'}{\sigma^2(x_i)}\right)$ no singular.

Mínimos cuadrados generalizados (GLS)

- Idea: Transformar el modelo en homoscedástico y luego aplicar MCO que sabemos es eficiente bajo homoscedasticidad e independencia en media.

Mínimos cuadrados generalizados (GLS)

- Idea: Transformar el modelo en homoscedástico y luego aplicar MCO que sabemos es eficiente bajo homoscedasticidad e independencia en media.

$$\frac{y_i}{\sigma(x_i)} = \frac{x_i}{\sigma(x_i)}'\beta + \frac{u_i}{\sigma(x_i)},$$

donde

$$\begin{aligned} \text{Var}\left(\frac{u_i}{\sigma(x_i)}|x_i\right) &= \mathbb{E}\left(\frac{u_i^2}{\sigma^2(x_i)}|x_i\right), \\ &= \frac{\mathbb{E}(u_i^2|x_i)}{\sigma^2(x_i)}, \\ &= 1. \end{aligned}$$

Mínimos cuadrados generalizados (GLS)

- Estimador por GLS

$$\hat{\beta}_{GLS} = \left(\frac{1}{N} \sum_{i=1}^{N} \frac{x_i x_i'}{\sigma^2(x_i)} \right)^{-1} \frac{1}{N} \sum_{i=1}^{N} \frac{x_i y_i}{\sigma^2(x_i)}.$$

Propiedades: Consistencia

$$
\begin{aligned}
\operatorname{plim} \hat{\beta}_{GLS} &= \operatorname{plim} \left[\left(\frac{1}{N} \sum_{i=1}^{N} \frac{x_i x_i'}{\sigma^2(x_i)} \right)^{-1} \left(\frac{1}{N} \sum_{i=1}^{N} \frac{x_i y_i}{\sigma^2(x_i)} \right) \right], \\
&= \beta + \left(\operatorname{plim} \frac{1}{N} \sum_{i=1}^{N} \frac{x_i x_i'}{\sigma^2(x_i)} \right)^{-1} \operatorname{plim} \left(\frac{1}{N} \sum_{i=1}^{N} \frac{x_i u_i}{\sigma^2(x_i)} \right), \\
&= \beta,
\end{aligned}
$$

usando LGN, TFC y $\mathbb{E}(u_i|x_i) = 0$.

Propiedades: Distribución asintótica normal

$$\sqrt{N}(\hat{\beta}_{GLS} - \beta) = \left(\frac{1}{N} \sum_{i=1}^{N} \frac{x_i x_i'}{\sigma^2(x_i)} \right)^{-1} N^{1/2} \frac{1}{N} \sum_{i=1}^{N} \frac{x_i u_i}{\sigma^2(x_i)}.$$

Usando LGN y TCL tenemos,

$$\begin{aligned} N^{1/2} \frac{1}{N} \sum_{i=1}^{N} \frac{x_i u_i}{\sigma^2(x_i)} &\xrightarrow{d} \text{Normal} \left(0, \mathbb{E} \left(\frac{u_i^2}{\sigma^2(x_i)} \frac{x_i x_i'}{\sigma^2(x_i)} \right) \right), \\ &\xrightarrow{d} \text{Normal} \left(0, \mathbb{E} \left(\frac{x_i x_i'}{\sigma^2(x_i)} \right) \right), \\ \left(\frac{1}{N} \sum_{i=1}^{N} \frac{x_i x_i'}{\sigma^2(x_i)} \right)^{-1} &\xrightarrow{p} \mathbb{E} \left(\frac{x_i x_i'}{\sigma^2(x_i)} \right)^{-1}. \end{aligned}$$

Usando el teorema de Slutsky

$$\sqrt{N}(\hat{\beta}_{GLS} - \beta) \xrightarrow{d} \text{Normal} \left(0, \mathbb{E} \left(\frac{x_i x_i'}{\sigma^2(x_i)} \right)^{-1} \right).$$

Mínimos cuadrados generalizados factibles (FGLS)

- En la práctica no conocemos $\sigma^2(x_i)$.

  Si suponemos una forma funcional $\sigma^2(x_i, \gamma)$, podemos estimar $\sigma^2(x_i, \gamma)$ consistentemente usando los residuos de la estimación por MCO,

  $$\hat{u}_i^2 = \sigma^2(x_i, \gamma) + error.$$

  Luego podemos estimar por *GLS* usando $\sigma^2(x_i, \hat{\gamma})$: Mínimos cuadrados generalizados factibles.

  Se puede demostrar que *GLS* y *FGLS* son asintóticamente equivalentes.

  En la práctica es preferible utilizar errores robustos.

Test de Heteroscedasticidad

- Dado el modelo,

$$y_i = \beta' x_i + u_i,$$

donde $\mathbb{E}(u_i|x_i) = 0$.

- Contraste de hipótesis: $H_0 : \mathbb{E}(u_i^2|x_i) = \sigma^2$.
Dado una función de x_i, $h(x_i)$ bajo la hipótesis nula $\text{Cov}(u_i^2, h(x_i)) = 0$.
Podemos utilizar el modelo:

$$u_i^2 = \delta_0 + h(x_i)'\delta + v_i,$$

bajo los supuestos usuales.
Bajo H_0, tenemos que $\delta_0 = \sigma^2$ y $\delta = 0$, y podemos hacer el contraste $H_0 : \delta = 0$ utilizando un test multiplicadores de Lagrange ($LM_N = N\hat{R}^2 \xrightarrow{p} \chi^2_q$).

- En realidad no observamos u_i^2 pero lo podemos reemplazar por $\hat{u}_i^2$ y asintóticamente nada cambia.

Test de Heteroscedasticidad

- Dado el modelo,

$$y_i = \beta' x_i + u_i,$$

  donde $\mathbb{E}(u_i|x_i) = 0$.
- Contraste de hipótesis: $H_0 : \mathbb{E}(u_i^2|x_i) = \sigma^2$.
- Cuando $h(x_i) = x_i$ tenemos el test de Breusch-Pagan, y para $h(x_i)$ incluye x_i y $x_i x_i'$ tenemos el test de White.
- En Stata despues de reg usamos `estat hettest, iid`.

Validación interna y validación externa

Validación interna y validación externa

- Validez interna: la inferencia estadística de los efectos causales de interés son válidos para la población estudiada.
- Validez externa: la inferencia estadística sobre los efectos causales se puede generalizar a poblaciones con diferentes condiciones institucionales, legales, sociales y económicas.

Amenazas a la validación externa

- Diferencias entre la población y contexto estudiados, y la población y contexto de interés.
  - Diferencias en las poblaciones: Efectos sobre alumnos en escuela media.
  - Diferencias en las condiciones institucionales, legales, sociales y económicas: diferentes características de los profesores, escuelas privadas y públicas, etc.
- ¿Cómo evaluar la validez externa?
  - Evaluar si hay diferencias importantes entre la población bajo estudio y la población de interés.
  - Comparar resultados de estudios similares en poblaciones diferentes.

Amenazas a la validación interna

- Sesgo por omisión de variable
- Mala especificación de la forma funcional
- Sesgo por error de medida
- Sesgo por selección de muestra
- Sesgo por simultaneidad

  Nota: Todas estos problemas generan correlación entre variables explicativas y el error, y se suelen llamar problemas de endogeneidad.

Sesgo por omisión de variable

- Surge si una variable omitida
  - Está correlada con la variable dependiente y
  - Está correlada con al menos una variable explicativa x_j.
- Soluciones
  - Si la variable puede ser medida, incluirla en la regresión como variable explicativa adicional.
  - Si tenemos una buena proxy, utilizar la proxy en la regresión como variable explicativa adicional.
  - Utilizar variables intrumentales (el tema de la siguiente unidad del curso).
  - Si observamos la misma unidad en distintos momentos del tiempo utilizar datos de panel que nos permite controlar por factores inobservados que permanecen fijos a lo largo del tiempo (el tema de Econometría II).
  - Realizar un experimento aleatorio controlado (randomized controlled experiment).

Mala especificación de la forma funcional

- Surge si la forma funcional de la esperanza condicional no es lineal.
- Soluciones
  - Variable dependiente continua: utilizar una forma no lineal apropiada: logaritmos, polinomios, interacciones,etc.
  - Variable dependiente discreta: modelos econométricos específicos para este tipo de variables. Por ejemplo, para variables binarias tenemos el modelo Logit y Probit (el tema de la última unidad del curso).

Sesgo por error de medida

- Surge por error de medida en las variables por problemas en el registro de datos administrativos, en la recolección de datos de las encuestas, en el "wording" de las preguntas, respuestas intencionalmente falsas, etc.
- Soluciones
  - Obtener mejores datos.
  - Utilizar variables instrumentales.

Sesgo por selección de muestra

- El sesgo por selección de muestra surge si la disponibilidad datos depende de un proceso de selección que está relacionado con la variable dependiente.
- Soluciones
  - Recolectar los datos de forma tal de aliviar el problema de selección
  - Construir un modelo del proceso de selección y estimarlo (Heckman).
  - Realizar un experimento aleatorio controlado.

Sesgo por simultaneidad

- Hasta ahora hemos supuesto que x causa a y .¿Qué pasa si y causa a x? (feedback effect)
- Soluciones
  - Realizar un experimento aleatorio controlado.
  - Estimar un modelo completo con las ambas causalidades: Sistema de ecuaciones.
  - Utilizar variables instrumentales.

Inconsistencia en la estimación de los errores estándar

- Heterocedasticidad
- Correlación del término de error entre observaciones: Datos de panel y series de tiempo.

Ejemplo: Tamaño de clase y rendimiento académico para California

- Objetivo: Analizar las amenazas a la validez interna y externa en los resultados empíricos con los datos de California.
- Validez externa
  - Comparar los resultados de California con los resultados de Massachusetts.
- Validez interna
  - Evaluar la lista de 5 amenazas a la validez interna.

Forma funcional

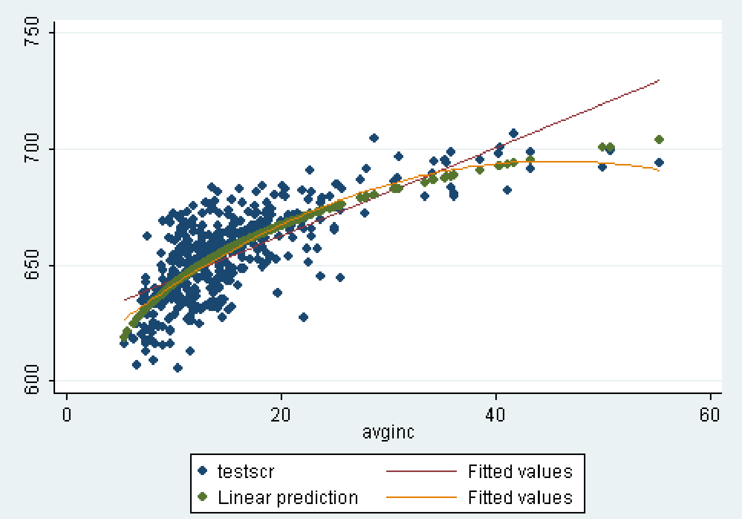
- ¿Efecto del ingreso en el rendimiento académico es lineal?
- ¿Efecto marginal de aumentar el tamaño del clase depende del porcentaje de niños que deben aprenden inglés?
- ¿Efecto marginal de aumentar el tamaño del clase depende del tamaño de clase?

Forma funcional

- Teoría económica puede sugerir la forma funcional.
- Análisis gráfico puede ayudar en determinar la forma funcional.
- Si la variable dependientes es la misma y los modelos alternativos son anidados, podemos utilizar un test de Wald.
- Si la variable dependientes es la misma pero los modelos no son anidados, podemos utilizar el R^2 ajustado.
- Si la variable dependiente es distinta, podemos utilizar un test del modelo de Box-Cox

Efecto del ingreso

Figure 1: Scatter plot: Lineal, nivel-log, cuadrático



Efecto del ingreso

Figure 2: Regresión particionada: avplot avginc

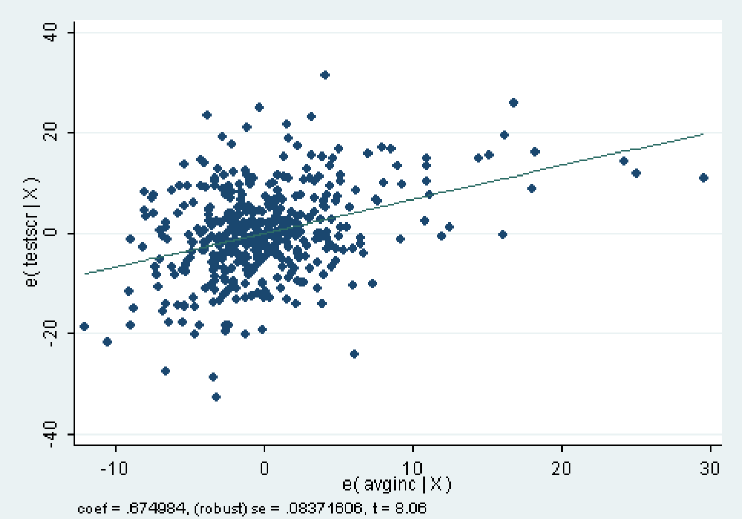
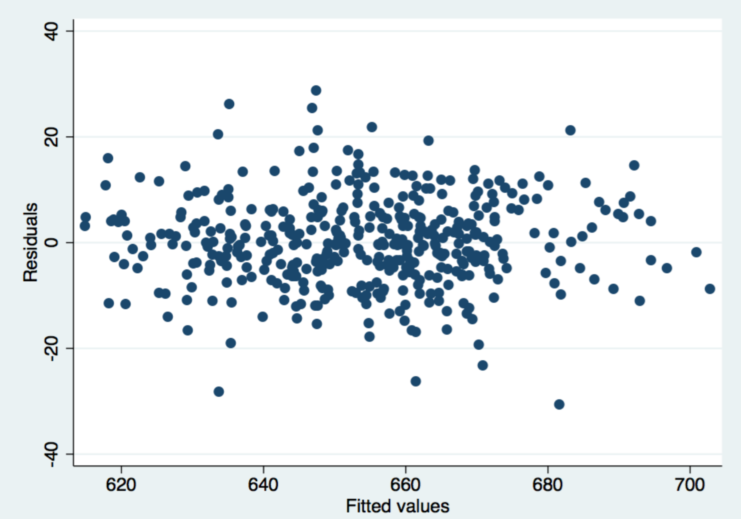


Figure 3: Valores predichos y residuos: rvfplot



Estimación MCO

	Model 1	Model 2	Model 3	Model 4
Stud/teacher (str)	-0.998***	-0.734***	-0.947***	65.286**
	(0.270)	(0.257)	(0.347)	(25.259)
English learn % (el)	-0.122***	-0.176***	-0.482*	-0.166***
	(0.033)	(0.034)	(0.246)	(0.034)
Lunch subsidy	-0.547***	-0.398***	-0.401***	-0.402***
	(0.024)	(0.033)	(0.033)	(0.033)
Log(Average income)		11.569***	11.433***	11.509***
		(1.819)	(1.818)	(1.806)
el X str			0.015	
			(0.012)	
str^2				-3.466***
				(1.271)
str^3				0.060***
				(0.021)
Constant	700.150***	658.552***	663.158***	244.803
	(5.568)	(8.642)	(10.054)	(165.722)
R-squared	0.775	0.796	0.797	0.801
N	420.000	420.000	420.000	420.000

* p<0.10, ** p<0.05, *** p<0.01

Efectos marginales de Student/teacher

	Model 1	Model 2	Model 3	Model 4
Student/teacher	-0.998***	-0.734***	-0.703***	-0.874***

* p<0.05, ** p<0.01, *** p<0.001

Efectos marginales de Student/teacher

$$testscr_i = \beta_0 + \beta_1 str_i + \beta_2 el\_pct_i + \beta_3 meal\_pct_i$$
$$+ \beta_4 lavginc_i + \beta_5 (el\_pct_i - \overline{el\_pct}) \times str_i + u_i$$

```
.     reg testscr str el_pct meal_pct lavginc el_str_demean, robust

Linear regression                                      Number of obs =     420
                                                       F(  5,   414) =  343.25
                                                       Prob > F      =  0.0000
                                                       R-squared     =  0.7969
                                                       Root MSE      =  8.6385

------------------------------------------------------------------------------
             |               Robust
     testscr |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         str |  -.7025698   .2550764    -2.75   0.006    -1.203976   -.2011634
      el_pct |  -.4818238   .2461291    -1.96   0.051    -.9656424    .0019949
    meal_pct |  -.4013796   .0332698   -12.06   0.000    -.4667784   -.3359809
     lavginc |   11.43307   1.817734     6.29   0.000     7.859935    15.00621
el_str_dem~n |   .0154803    .012334     1.26   0.210    -.0087647    .0397254
       _cons |   663.1576   10.05435    65.96   0.000     643.3937    682.9216
------------------------------------------------------------------------------
```

Resumen de forma funcional

- Es importante controlar por el ingreso (en forma logarítmica o cuadrática)
- No hay evidencia que el efecto marginal del tamaño de clase dependa del porcentaje de habitantes que deben aprender inglés.
- Existe evidencia que el efecto marginal del tamaño de clase depende del tamaño de clase.

Validez interna

- Omisión de variable: Pueden existir otras variable omitidas como la calidad de los profesores.
- Forma funcional: Se evaluaron distintas formas funcionales y el efecto marginal es robusto a las distintas formas funcionales.
- Error de medida: Puede existir error de medida en el tamaño de clase y el ingreso.
- Selección de muestra: La muestra cubre todas la escuelas públicas primarias de California.
- Simultaneidad: No existen programas de fondos o becas basados en el rendimiento de los estudiantes en California.
- Heteroscedasticidad: Se utilizaron errores robustos.

Validez externa

- Para evaluar la validez externa necesitamos un población de interés.
- Población de interés: Escuelas primarias públicas en otros estados de USA.
- Utilizar datos de Massachusetts para estimar el efecto del tamaño de clase.

Estadísticas descriptivas

```
.     use caschool.dta, clear

.
. * Descriptive statistics for test score and student-teacher ratio
.     tabstat testscr str  el_pct meal_pct avginc, statistics(mean sd)

   stats |   testscr       str    el_pct  meal_pct    avginc
---------+--------------------------------------------------
    mean |  654.1565  19.64043  15.76816  44.70524  15.31659
      sd |  19.05335  1.891812  18.28593  27.12338   7.22589
------------------------------------------------------------
.     use mcas.dta, clear

.
. * Descriptive statistics for test score and student-teacher ratio
.     tabstat  totsc4 tchratio  pctel lnch_pct percap, statistics(mean sd)

   stats |    totsc4  tchratio     pctel  lnch_pct    percap
---------+--------------------------------------------------
    mean |  709.8273  17.34409  1.117676  15.31591  18.74676
      sd |  15.12647  2.276666   2.90094  15.06007  5.807637
------------------------------------------------------------
```

Estadísticas descriptivas: CA versus MA

- Resultado promedio es mayor pero los tests son diferentes asi que no son directamente comparables.
- Las clases son mas pequeñas en MA, el ingreso medio es 20% mayor en MA, % de habitantes que deben aprender inglés es mayor en CA, y el % de pobres es mayor en CA.

Estimación MCO

	Model 1	Model 2	Model 3	Model 4	Model 5
Stud/teacher (str)	-1.718***	-0.801***	-0.641**	-0.713**	12.426
	(0.499)	(0.279)	(0.268)	(0.301)	(14.010)
English learn % (el)		-0.032	-0.437	-1.103	-0.434
		(0.287)	(0.303)	(1.096)	(0.300)
Lunch subsidy		-0.762***	-0.582***	-0.593***	-0.587***
		(0.064)	(0.097)	(0.101)	(0.104)
Average income			-3.067	-3.142	-3.382
			(2.353)	(2.391)	(2.491)
Average income^2			0.164*	0.166*	0.174*
			(0.085)	(0.087)	(0.089)
Average income^3			-0.002**	-0.002**	-0.002**
			(0.001)	(0.001)	(0.001)
el X str				0.038	
				(0.055)	
str^2					-0.680
					(0.737)
str^3					0.011
					(0.013)
Constant	739.621***	735.413***	744.025***	746.236***	665.497***
	(8.607)	(4.896)	(21.318)	(22.238)	(81.332)
R-squared	0.067	0.629	0.685	0.686	0.687
N	220.000	220.000	220.000	220.000	220.000

Efectos marginales de Student/teacher

	Model 1	Model 2	Model 3	Model 4	Model 5
Student/teacher	-1.718***	-0.801***	-0.641**	-0.671**	-0.640**

* p<0.05, ** p<0.01, *** p<0.001

Resumen de forma funcional

- Es importante controlar por el ingreso (en forma logarítmica o cúbica)
- No hay evidencia que el efecto marginal del tamaño de clase dependa del porcentaje de habitantes que deben aprender inglés.
- No hay evidencia que el efecto marginal del tamaño de clase depende del tamaño de clase.

Efecto de disminuir el tamaño de clase en 2 estudiantes: CA versus MA

Table 1: Ratios de estudianes/profesor y rendimiento de los estudiantes: CA versus MA

			Efecto estimado de 2 estudiantes menos por profesor	
	Estimador MCO	sd(test scores)	Puntos	sd
California				
Modelo lineal	-0.73	19.1	1.46	0.076
Modelo cúbico	-0.87	19.1	1.74	0.091
Massachusetts				
Modelo lineal	-0.64	15.1	1.28	0.085

Proyecto STAR (Student-teacher Achievement Ratio): Tennessee

- Experimento aleatorio que duró 4 años y comenzo 1985-86 en 80 escuelas.
- 3 clases distintas: regulares con 22-25 alumnos, pequeñas con 13-17 alumnos, y regulares con un ayudante para el profesor.
- Alumnos desde kinder a 3er grado.

Proyecto STAR (Student-teacher Achievement Ratio): Tennessee

	Kinder	1er grado	2do grado	3er grado
Small class	13.899***	29.781***	19.394***	15.587***
	(2.409)	(2.807)	(2.710)	(2.395)
Regular size with aide	0.314	11.959***	3.479	-0.291
	(2.310)	(2.686)	(2.566)	(2.302)
_cons	918.043***	1039.393***	1157.807***	1228.506***
	(1.641)	(1.836)	(1.849)	(1.715)
R-squared	0.007	0.017	0.009	0.010
N	5786.000	6379.000	6049.000	5967.000

* p<0.10, ** p<0.05, *** p<0.01

Resultados

- Tener un ayudante no tiene efecto, excepto en en 1er grado.
- Disminuir el tamaño de clase en 7 alumnos tiene un efecto estadísticamente significativo.

Efecto de disminuir el tamaño de clase

Table 2: Ratios de estudianes/profesor y rendimiento de los estudiantes: STAR versus CA-MA

	Estimador MCO	Cambio en estudiantes/profesor	sd(test scores entre estudiantes)	Efecto estimado
STAR (Kinder)	-13.90	Small vs Regular	73.8	0.19
California	-0.73	-7.5	38.0	0.14
Massachusetts	-0.64	-7.5	39.0	0.12