PUBLICACIONES DE 4er CURSO

Curso: 4º

Grado: Economía

Asignatura: ECONOMETRÍA III

TRANSPARENCIAS PARTE 2: Regresores estocásticos

TEMA 5: Evaluación

Profesores: Antonio Aznar

Departamento de ANÁLISIS ECONÓMICO

Curso Académico 2015/16

Facultad de Economía y Empresa
Universidad Zaragoza

TEMA 5: EVALUACIÓN

1. Validez Interna y Externa

2. Amenazas a la Validez Interna

   a. Sesgo de Variable Omitida

   b. Forma functional errónea

   c. Sesgo por errores de observación.

   d. Sesgo por datos ausentes y por selección muestral.

   e. Sesgo de causalidad simultánea.

3. Contrastes de esfericidad

4. Criterios de información

1. Validez interna y externa

¿Qué hace que un estudio de regresión múltiple sea o no fiable?

¿Cuándo proporcionará la regresión múltiple un estimador útil- insesgado y consistente- del efecto causal? O lo que es lo mismo ¿Cuándo no lo hará?

C.C. 9.1 Validez Interna y Externa

- ***Validez Interna***: Las inferencias estadísticas sobre los efectos causales son válidas para la población que se está estudiando.

- ***Validez Externa***: Las inferencias estadísticas pueden ser generalizadas de la población y escenarios estudiados a otras poblaciones y escenarios, en donde escenarios se refiere a los entornos legal, histórico y político.

2. Amenazas a la Validez Externa

El estudio de las amenazas a la validez externa requiere un análisis detallado caso por caso.

Hasta donde podemos generalizar los resultados obtenidos para California?

- Diferencias en poblaciones
  - California en 2011?
  - Massachusetts en 2011?
  - Mexico en 2011?
- Diferencias de Escenarios
  - Diferente marco legal
  - Diferente tratamiento de la educación bilingüe
  - Diferentes características del profesorado

C.C. 9.7 Amenazas a la Validez Interna de un estudio de Regresión Múltiple

Existen cinco amenazas principales a la validez interna:

a. Sesgo de Variable Omitida
b. Forma functional errónea
c. Sesgo por errores de observación.
d. Sesgo por datos ausentes y por selección muestral.
e. Sesgo de causalidad simultánea.

Todos estos implican que $E(u_i|X_{1i},\ldots,X_{ki}) \neq 0$ (o que no se cumple la hipótesis de la independencia en la media condicional) –en cuyo caso los MCO son sesgados e inconsistentes.

ESFERICIDAD

El cálculo incorrecto de los errores estándar representa, asimismo, una amenaza a la validez interna. Los errores estándar válidos con homocedasticidad no son válidos en presencia de heterocedasticidad. Si las variables no son independientes entre distintas observaciones, lo cual puede ocurrir en datos de panel y en datos de series temporales, entonces se necesita un nuevo ajuste en la fórmula de los errores estándar a fin de obtener errores estándar válidos.

La aplicación de esta línea de amenazas a un estudio de regresión múltiple constituye un método sistemático de evaluar la validez interna del estudio

a) Sesgo de Variable Omitida

C.C. 9.2 Sesgo de Variable Omitida

El sesgo de variable omitida surge cuando una variable omitida es, al mismo tiempo,

(i) un determinante de Y y (ii) está correlacionada con al menos uno de los regresores incluidos.

Esto genera una correlación entre la variable independiente y el término de error que hace que el MCO sea sesgado e inconsistente

- Primero, discutiremos el caso con un solo regresor y, a continuación, lo extenderemos al caso con más regresores.
- Si la regresión múltiple incluye variables control necesitamos preguntarnos si estas variables realmente controlan la posible relación de la perturbación con el regresor de interés.

Sesgo de Variable Omitida .Soluciones

1. Si la variable causal omitida puede ser medida, hay que incluirla como un regresor adicional en la regresión múltiple;
2. Si se tienen datos sobre una o mas variables control y estas variables son efectivas(en el sentido de garantizar el cumplimiento de la hipótesis de la independencia en la media condicional) entonces incluye estas variables control;
3. Si es posible, utilizar datos panel en donde cada agente es observado en periodos sucesivos;
4. Utilizar Variable Instrumentales;
5. Realizar un experimento aleatorizado controlado.

Porque funciona el experiment aleatorizado?Recordar-si X se asigna aleatoriamente, entonces X necesariamente se distribuye independientemente de u; esto es, $E(u|X = x) = 0$.

b) Sesgo de Forma Funcional Equivocada

C.C. 9.3 Error de Especificación de la Forma Funcional

Este error aparece cuando la forma funcional de la regresión estimada difiere de la forma funcional de la función de regresión poblacional. Si la especificación es incorrecta entonces el estimador del efecto parcial de un cambio en una de las variables será, en general, sesgado.

Soluciones

1. Variable dependiente continua: utilizar la especificación no lineal apropiada en X (logaritmos, interacciones, etc.). Ver gráfico de datos y de la regresión-residuos
2. Variable dependiente discreta (*ejemplo*: binaria): necesita una extensión de los modelos de regresión múltiple (análisis "probit" or "logit" para variables dependientes binarias).

c) Sesgo de Errores de medida

C.C. 9.4 Sesgo por errores en la variables

Este sesgo en los estimadores MCO se produce cuando una variable independiente se mide de forma imprecisa, lo que hace que la variable y el término de error estén correlacionados. Este sesgo depende de la naturaleza del error de medida y persiste incluso si el tamaño de la muestra es grande.

.Cusas de este error:

- Errores de manejo administrativo
- Errores en datos de encuestas.
- Cuestiones ambiguas (¿cuál fue la renta del último año?)
- Respuestas intencionalmente falsas en encuestas (¿cuál es el valor actual de tus activos financieros? ¿Con que frecuencia conduces bebido?)

En general, el error de medida en un regresor acaba en el sesgo de "errores en variables".

Un poco de matemáticas permite demostrar como el error de medida conduce a la correlación entre la perturbación y el regresor. Considerar el modelo simple:

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

Y suponer $E(u_i|X_i) = 0$). Sea

X_i = el verdadero valor desconocido de X

$\tilde{X}_i$ = el valor observado con error de X entonces

$$Y_i = \beta_0 + \beta_1 X_i + u_i$$
$$= \beta_0 + \beta_1 \tilde{X}_i + [\beta_1(X_i - \tilde{X}_i) + u_i]$$

La regression que se hace es,

$$Y_i = \beta_0 + \beta_1 \tilde{X}_i + \tilde{u}_i, \text{ donde } \tilde{u}_i = \beta_1(X_i - \tilde{X}_i) + u$$

Con error de medida, generalmente $\tilde{X}_i$ está correlacionado con $\tilde{u}_i$ por lo que $\hat{\beta}_1$ es sesgado:

$$\text{cov}(\tilde{X}_i, \tilde{u}_i) = \text{cov}(\tilde{X}_i, \beta_1(X_i - \tilde{X}_i) + u_i)$$
$$= \beta_1 \text{cov}(\tilde{X}_i, X_i - \tilde{X}_i) + \text{cov}(\tilde{X}_i, u_i)$$

Es possible que $\text{cov}(\tilde{X}_i,u_i) = 0$ (if $E(u_i|X_i) = 0$ entonces $\text{cov}(\tilde{X}_i,u_i) = 0$ si el error de medida en $\tilde{X}_i$ no está correlacionado con u_i). pero en general $\text{cov}(\tilde{X}_i, X_i - \tilde{X}_i) \neq 0$….

A. Error de Medida Clásico

Se supone que,

$$\tilde{X}_i = X_i + v_i,$$

En donde v_i es un ruido blanco con media cero y corr(X_i, v_i) = 0 y corr(u_i, v_i) = 0.

Bajo este modelo, $\hat{\beta}_1$ es sesgado hacia cero. Esta es la idea: Suponer que tomas una variable aleatoria a la que añades una gran cantidad de ruido aleatorio-números aleatorios generados en el ordenador. En el límite de “todo ruido” $\tilde{X}_i$ no guardará ninguna relación con Y_i de forma que la esperanza del estimador del coeficiente será cero. Si $\tilde{X}_i$ tiene algún ruido pero no es “todo ruido” entonces habrá una relación entre $\tilde{X}_i$ e Y_i de forma que $\hat{\beta}_1$ es sesgado hacia cero.

$$\tilde{X}_i = X_i + v_i, \text{ donde } \mathrm{corr}(X_i, v_i) = 0 \text{ y } \mathrm{corr}(u_i, v_i) = 0.$$

entonces $\quad \mathrm{var}(\tilde{X}_i) = \sigma_X^2 + \sigma_v^2$

$$\mathrm{cov}(\tilde{X}_i, X_i - \tilde{X}_i) = \mathrm{cov}(X_i + v_i, -v_i) = -\sigma_v^2$$

$$\mathrm{cov}(\tilde{X}_i, \tilde{u}_i) = -\beta_1 \sigma_v^2$$

$$\hat{\beta}_1 \xrightarrow{p} \beta_1 - \beta_1 \frac{\sigma_v^2}{\sigma_{\tilde{X}}^2} = \left(1 - \frac{\sigma_v^2}{\sigma_{\tilde{X}}^2}\right)\beta_1$$

$$= \left(\frac{\sigma_{\tilde{X}}^2 - \sigma_v^2}{\sigma_{\tilde{X}}^2}\right)\beta_1 = \left(\frac{\sigma_X^2}{\sigma_X^2 + \sigma_v^2}\right)\beta_1$$

por ello $\hat{\beta}_1$ is sesgado hacia cero.

El modelo de error de medida clásico es especial porque asume que corr(X_i, v_i) = 0.

B. Error de Medida de “Mejor Conjetura”

Suponer que el que responde no recuerda X_i, pero hace una mejor conjetura de la forma $\tilde{X}_i = E(X_i|W_i)$, donde $E(u_i|W_i) = 0$. Entonces,

$$\begin{aligned}\text{cov}(\tilde{X}_i, \tilde{u}_i) &= \text{cov}(\tilde{X}_i, \beta_1(X_i - \tilde{X}_i) + u_i)\\ &= \beta_1\text{cov}(\tilde{X}_i, X_i - \tilde{X}_i) + \text{cov}(\tilde{X}_i, u_i)\end{aligned}$$

- $\text{cov}(\tilde{X}_i, X_i - \tilde{X}_i) = 0$ porque $\tilde{X}_i = E(X_i|W_i)$ (porque $\tilde{X}_i$ es la major conjetura, el error $X_i - \tilde{X}_i$ no está correlacionado con $\tilde{X}_i$).
- $\text{cov}(\tilde{X}_i, u_i) = 0$ porque $\text{E}(u_i|W_i) = 0$ ($\tilde{X}_i$ es una función de W_i y $\text{E}(u_i|W_i) = 0$).
- Así $\text{cov}(\tilde{X}_i, \tilde{u}_i) = 0$, de forma que $\hat{\beta}_1$ es insesgado.

- Bajo este modelo, sigues teniendo un error de medida –tu no observas el verdadero valor de X_i – pero este error de medida no introduce sesgo en $\hat{\beta}_1$!
- El modelo de “mejor conjetura es extremo ya que no basta con hacer una buena conjetura sino que se necesita la mejor $\tilde{X}_i = E(X_i|W_i)$, esto es, la esperanza condicional de *X dada W*, donde $E(u_i|W_i) = 0$.

Soluciones al sesgo de errores de medida

1. Obtener mejores datos (often easier said than done).
2. Desarrollar un modelo específico del proceso de error de medida. Esto solo es posible si se tiene un gran conocimiento del proceso del error de medida.
3. Regresión de Variables Instrumentales.

C. C. 9.5 Sesgo de datos ausentes y de selección muestral

Este sesgo se presenta cuando el proceso de selección de los elementos muestrales influye en le disponibilidad de los datos y el proceso está relacionado con la variable dependiente, además de depender de los regresores. Este proceso de selección induce correlación entre uno o más regresores y el término de error lo que da lugar a que los estimadores MCO sean sesgados e inconsistentes en muestras grandes.

A menudo los datos son incompletes. A veces, esta falta de datos produce sesgo y otras no. Es útil considerar tres casos:

1. La ausencia de datos es aleatoria.
2. La ausencia de datos depende del valor de una o mas *X*'s
3. La ausencia de datos depende de Y o de u.

Casos 1 y 2 no generan sesgo: los errores estándar son mayores que serían si se dispusiera de los datos pero $\hat{\beta}_1$ es insesgado.

El Caso 3 introduce el sesgo de "Selección Muestral".

Caso 1 Ausencia Aleatoria

Suponer que tomas una muestra aleatoria de 100 trabajadores y anotas las respuestas en hojas de papel. – pero tu perro se come 20 de las hojas (seleccionadas al azar) antes de que las puedas introducir en el ordenador. Esto equivale a tomar una muestra aleatoria de 80 trabajadores de forma que tu perro no introdujo ningún sesgo.

Caso 2. Ausencia que depende de valores de X.

En el ejemplo de las calificaciones suponer que restringes tu análisis al subconjunto formado por los distritos con REM<20. Los resultados obtenidos solo se aplicaran al caso cuando el ratio REM sea bajo pero eso no introduce sesgo. Esto es equivalente a una falta de los datos para cuando REM>20. Por lo tanto cuando los datos faltan dependiendo de los datos de X eso no produce que el estimador MCO sea sesgado.

Caso 3. Selección de muestra

La ausencia de datos depende de los valores de Y o de u.
En general, este tipo de ausencias de datos introduce sesgo en los estimadores MCO. Ese tipo de sesgo se llama también sesgo de selección muestral.

El sesgo de selección muestral se produce cuando un proceso de selección:

(i) influye en la disponibilidad de datos y
(ii) está relacionado con la variable dependiente.

Ejemplo #1: Altura de estudiantes universitarios

Se trata de realizar un estudio para estimar la altura media de los estudiantes masculinos. Recoges tus datos (la muestra) en la puerta del equipo de baloncesto, apuntando la altura de los que entran.

- Es este un buen diseño? – proporcionará una estimación insesgada de la altura media de los estudiantes?
- Se recoge la muestra de forma que depende del resultado de Y (la altura), lo que produce el sesgo.

Ejemplo #2: Fondos Activos

- Logran los fondos activos mejores resultados que los fondos seguidores de Mercado?
- Estrategia Empírica:
  - Muestreo: eleatorio simple de los fondos activos disponibles al público en una determinada fecha.
  - Datos: rendimientos en los últimos diez años.
  - Estimador: rendimiento medio en los diez años de los fondos activos en la muestra menos el rendimiento en los diez últimos años del S&P500
  - ¿Hay un sesgo de selección de muestra? (o lo que es lo mismo, ¿los datos perdidos dependen del valor de Y o de u?)
  - ¿En qué se parece este ejemplo al de los jugadores de baloncesto?

Ejemplo #2: Fondos Activos

Los fondos activos que sobreviven son los “jugadores de baloncesto” de los fondos activos. Pero habría que tener en cuenta también a los que han fracasado y han desaparecido

Ejemplo #3: rendimientos de educación

- ¿Cuál es el rendimiento de un año adicional de educación?
- Estrategia Empírica:
  - Muestreo: aleatorio simple de graduados empleados. (empleados, de los que se tienen datos salariales)
  - Datos: ganancias y años de educación
  - Estimador: hacer la regresión de ln(*earnings*) sobre *years_education*
  - Ignorar cuestiones de sesgo de variable omitiday de error de medida – Hay sesgo de selección de muestra?
  - ¿Cómo se relaciona esto con el ejemplo del jugador de baloncesto?

Soluciones al Sesgo de Selección Muestral

- Tomar la muestra evitando el muestreo selectivo.
  - *Basketball player example*:Obtener una verdadera muestra aleatoria por ejemplo a partia de la lista de matricula de todos los estudiantes.
  - *Mutual funds example*: Cambiar la población para obtener la muestra y en lugar de coger los fondos existentes en el ultimo año coger los del primer año. (incluidos los que fracasan a lo largo de los diez años y no están en la lista del último año.
  - *Returns to education example*: muestreo sobre los graduados y no solo los que trabajan.
- Experimento aleatorizado controlado.

4. Sesgo de Simultaneidad

Aparece en una regresión de Y sobre X, cuando, además del vínculo causal de interés que va desde X a Y, existe un vínculo causal desde Y hacia X. Esta causalidad inversa provoca que X esté correlacionado con el termino de error en la regresión poblacional de interés lo que da lugar a que los estimadores MCO sean sesgados e inconsistentes.

Hasta el momento hemos supuesto que X causa a Y.
¿Qué pasa si Y causa a X, también?

Ejemplo: Efecto del tamaño de la clase sobre las calificaciones

- Un valor bajo de REM lleva a mejores calificaciones.
- Pero suponer que se pone en marcha una política que concede recursos extra para los distritos con bajas calificaciones.
- Que significa esto en una regresión de las calificaciones sobre el tamaño?

e) **Sego de Causalidad simultánea**

(a) Efecto causal de X sobre Y: $Y_i = \beta_0 + \beta_1 X_i + u_i$

(b) Efecto Causal de Y sobre X: $X_i = \gamma_0 + \gamma_1 Y_i + v_i$

- Resulta corr$(X_i,u_i) \neq 0$
- Como consecuencia $\hat{\beta}_1$ es sesgado e inconsistente.
- *Ejemplo: Un distrito con muy malas calificaciones, dado el nivel de REM (negativo u_i*) recibe recursos extra por lo que baja su REM, por lo tanto, REM y u_i están correlacionados.

Soluciones al sesgo de simultaneidad

1. Realizar un experimento aleatorio controlado. Como X_i se elige aleatoriamente no hay feedback.
2. Especificar y estimar un modelo completo que incorpore las dos direcciones de causalidad. Esta es la idea que está detrás de los modelos macro-econométricos construidos en los últimos años. Esto es extremadamente difícil en la práctica.
3. Utilizar variables instrumentales para estimar el efecto causal de interés. (efecto de X sobre Y, ignorando el efecto de Y sobre X).

Validez Interna y Externa para predicción

- La predicción y la estimación de efectos causales son objetivos totalmente diferentes.
- Para predecir,
  - El $\bar{R}^2$ importa (mucho!)
  - El sesgo de variable omitida no es un problema!
  - La interpretación de los coeficientes en predicción no es importante – lo que importa es un buen ajuste.
  - La validez externa es lo más importante: el modelo estimado con datos históricos debe ser válido en el futuro cercano.
  - Se verán más detalles cuando estudiemos los datos de series temporales.

Aplicación de la validez interna y externa al ejemplo de California.

Calificaciones y tamaño muestral

Objetivo: Estudiar las amenazas a la validez interna y externa del estudio econométrico de las calificaciones de California.

- Validez Externa
  - Comparar los resultados para California y Massachusetts
  - Think hard…
- Validez Interna
  - Recorrer la lista de las cinco amenazas potenciales a la validez interna y think hard…

Chequeo de la Validez Externa

Compararemos el estudio de California con otro que utiliza los datos de Massachusetts.

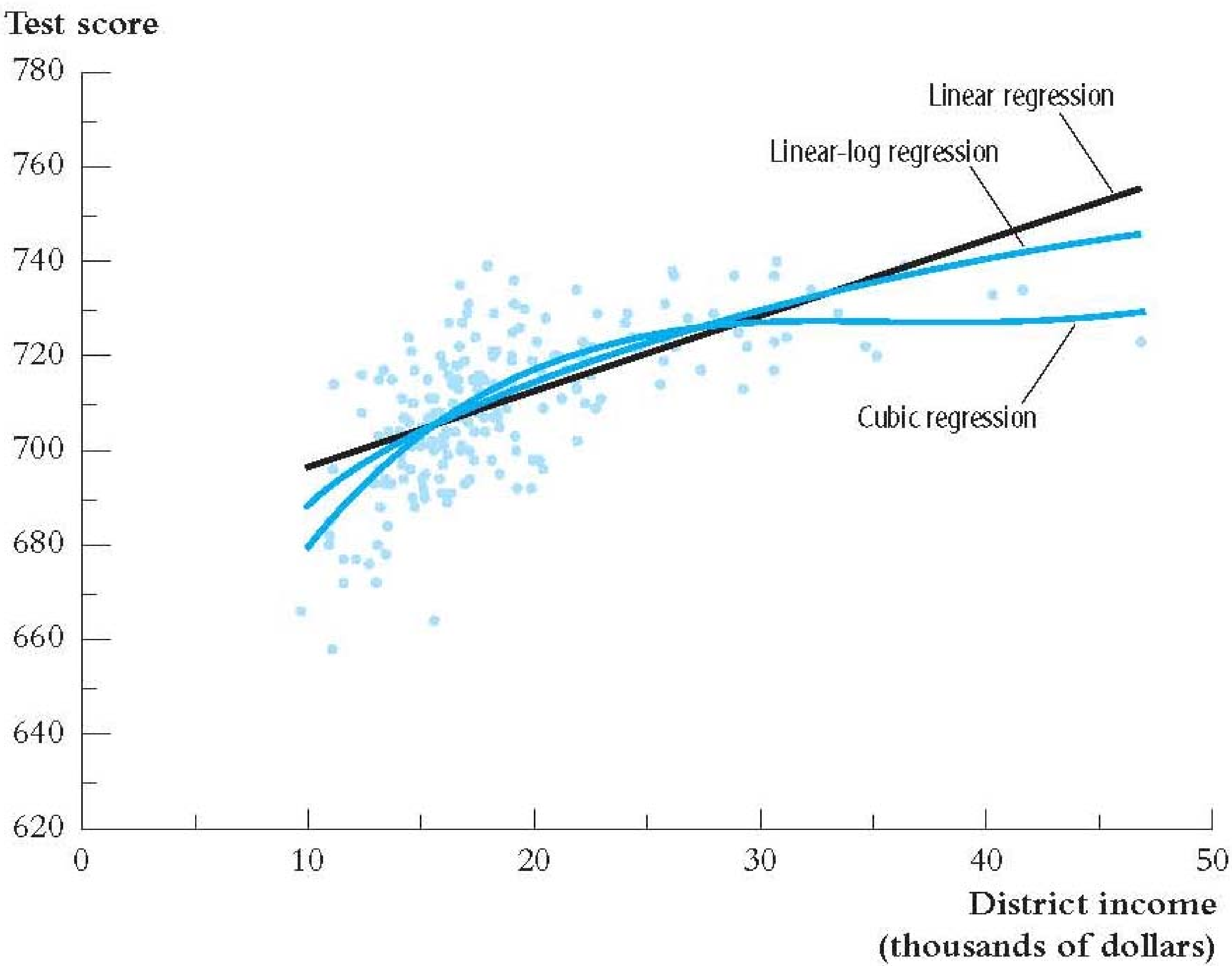
La base de datos de Massachusetts

- 220 distritos escolares
- Test: 1998 MCAS test – fourth grade total (Math + English + Science)
- Variables: REM(*STR*), Calificaciones(*TestScore)*, *PctEL*, *LunchPct*, *Income*

Datos de Massachusetts : Estadísticos Resumen

TABLE 9.1 Summary Statistics for California and Massachusetts Test Score Data Sets

	California		Massachusetts	
	Average	Standard Deviation	Average	Standard Deviation
Test scores	654.1	19.1	709.8	15.1
Student–teacher ratio	19.6	1.9	17.3	2.3
% English learners	15.8%	18.3%	1.1%	2.9%
% Receiving lunch subsidy	44.7%	27.1%	15.3%	15.1%
Average district income ($)	$15,317	$7226	$18,747	$5808
Number of observations	420		220	
Year	1999		1998	



Test scores vs. Income & regression lines: Massachusetts data

TABLE 9.2 Multiple Regression Estimates of the Student–Teacher Ratio and Test Scores: Data from Massachusetts

Dependent variable: average combined English, math, and science test score in the school district, fourth grade; 220 observations.

Regressor	(1)	(2)	(3)	(4)	(5)	(6)
Student–teacher ratio (*STR*)	−1.72** (0.50)	−0.69* (0.27)	−0.64* (0.27)	12.4 (14.0)	−1.02** (0.37)	−0.67* (0.27)
STR^2				−0.680 (0.737)		
STR^3				0.011 (0.013)		
% English learners		−0.411 (0.306)	−0.437 (0.303)	−0.434 (0.300)		
% English learners > median? (Binary, *HiEL*)					−12.6 (9.8)	
$HiEL \times STR$					0.80 (0.56)	
% Eligible for free lunch		−0.521** (0.077)	−0.582** (0.097)	−0.587** (0.104)	−0.709** (0.091)	−0.653** (0.72)
District income (logarithm)		16.53** (3.15)				
District income			−3.07 (2.35)	−3.38 (2.49)	−3.87* (2.49)	−3.22 (2.31)
District income^2			0.164 (0.085)	0.174 (0.089)	0.184* (0.090)	0.165 (0.085)
District income^3			−0.0022* (0.0010)	−0.0023* (0.0010)	−0.0023* (0.0010)	−0.0022* (0.0010)
Intercept	739.6** (8.6)	682.4** (11.5)	744.0** (21.3)	665.5** (81.3)	759.9** (23.2)	747.4** (20.3)

(*Table 9.2 continued*)

F-Statistics and p-Values Testing Exclusion of Groups of Variables

Regressor	(1)	(2)	(3)	(4)	(5)	(6)
All STR variables and interactions = 0				2.86 (0.038)	4.01 (0.020)	
$STR^2, STR^3 = 0$				0.45 (0.641)		
$Income^2, Income^3$			7.74 ($<$ 0.001)	7.75 ($<$ 0.001)	5.85 (0.003)	6.55 (0.002)
$HiEL, HiEL \times STR$					1.58 (0.208)	
SER	14.64	8.69	8.61	8.63	8.62	8.64
$\overline{R}^2$	0.063	0.670	0.676	0.675	0.675	0.674

These regressions were estimated using the data on Massachusetts elementary school districts described in Appendix 9.1. Standard errors are given in parentheses under the coefficients, and p-values are given in parentheses under the F-statistics. Individual coefficients are statistically significant at the *5% level or **1% level.

How do the Mass and California results compare?

- Logarithmic v. cubic function for *STR*?
- Evidence of nonlinearity in *TestScore*-*STR* relation?
- Is there a significant $HiEL \times STR$ interaction?

Efectos previstos de una reducción del tamaño de la clase de 2

Especificación para Mass:

$$\widehat{NotaExamen} = \underset{(21.3)}{744.0} - \underset{(0.27)}{0.64}REM - \underset{(0.303)}{0.437}PctEL - \underset{(0.097)}{0.582}LchPct$$

$$- \underset{(2.35)}{3.07}Renta + \underset{(0.085)}{0.164}Renta^2 - \underset{(0.0010)}{0.0022}Renta^3$$

- Estimated effect = $-0.64\times(-2) = 1.28$
- Estándar error = $2\times 0.27 = 0.54$

NOTE: $\text{var}(aY) = a^2\text{var}(Y)$; $SE(a\hat{\beta}_1) = |a|SE(\hat{\beta}_1)$

- 95% CI = $1.28 \pm 1.96\times 0.54 = (0.22, 2.34)$

Computo de los efectos estimados en modelos no lineales

Utilizar el método de "antes" y "después":

$$\widehat{NotaExamen} = 655.5 + 12.4REM - 0.680REM^2 + 0.0115REM^3$$
$$- 0.434PctEL - 0.587LchPct$$
$$- 3.48Renta + 0.174\,Renta^2 - 0.0023\,Renta^3$$

Reducción estimada al pasar de 20 a 18 estudiantes en el tamaño:

$$\Delta\widehat{NotaExamen} = [12.4\times 20 - 0.680\times 20^2 + 0.0115\times 20^3]$$
$$- [12.4\times 18 - 0.680\times 18^2 + 0.0115\times 18^3] = 1.98$$

- Comparar con la estimación del modelo lineal de 1.28
- *Para calcular el error estándar de esta estimación utilizar el método de transformación.*

Resumen de los resultados encontrados para Massachusetts

- El coeficiente de REM cae de –1.72 a –0.69 cuando se incluyen las variables control – que es una indicación de que la estimación original estaba afectada por el sesgo de variable omitida.
- El efecto del tamaño de la clase es estadísticamente significativo al 1%, después de controlar con variables de las características de los estudiantes y del distrito.
- Ninguna evidencia estadística de no linealidades en la relación NotasExamen-REM.
- Ninguna evidencia de la interacción *REM* – *PctEL*.

Comparación de los efectos del tamaño: CA vs. MA

TABLE 9.3 Student–Teacher Ratios and Test Scores: Comparing the Estimates from California and Massachusetts

	OLS Estimate $\hat{\beta}_{STR}$	Standard Deviation of Test Scores Across Districts	Estimated Effect of Two Fewer Students per Teacher, In Units of: Points on the Test	Standard Deviations
California				
Linear: Table 8.3(2)	−0.73 (0.26)	19.1	1.46 (0.52)	0.076 (0.027)
Cubic: Table 8.3(7) *Reduce STR from 20 to 18*	—	19.1	2.93 (0.70)	0.153 (0.037)
Cubic: Table 8.3(7) *Reduce STR from 22 to 20*	—	19.1	1.90 (0.69)	0.099 (0.036)
Massachusetts				
Linear: Table 9.2(3)	−0.64 (0.27)	15.1	1.28 (0.54)	0.085 (0.036)

Standard errors are given in parentheses.

Resumen; Comparación de California y Massachusetts

- Los efectos de la clase caen en los dos CA, MA cuando se añaden variables control referidas a estudiantes y distrito.
- El efecto del tamaño es estadísticamente significativo para los dos CA, MA data.
- El efecto estimado de una reducción de 2-estudiantes en REM es cuantitativamente similar para CA, MA.
- En ninguno de los dos casos se encuentra evidencia de la interacción $STR – PctEL$.
- Se encuentra evidencia de no linealidad en los datos de CA pero no en los de MA.

Recapitulando: Qué queda de amenazas de la validez interna en el contraste de Calificaciones/tamaño de clase.

¿Qué factores causales faltan?

- Las características de los estudiantes tales como el origen.
- Acceso a oportunidades de aprendizaje externas
- Calidad del profesorado.

Las regresiones intentan controlar estos factores omitidos utilizando variables control que no necesariamente incorporan causalidad pero que están correlacionados con las variables causales omitidas:

- distrito (renta, % concesión desayuno libre)
- Porcentaje de estudiantes de ingles.

1. Sesgo de variable Omitida

¿Son las variables de control efectivas? Esto es después de ser incluidas ¿se relaciona el error de la regresión con la variable REM?

- La respuesta requiere juicio no respuesta mecánica.
- En este ejemplo hay evidencia de que las variables control pudieran estar hacienda su labor:
  - El coeficiente de REM no cambia mucho cuando se cambian las variables control en las diferentes especificaciones.
  - Los resultados para California y Massachusetts son similares – por lo que si queda algo de sesgo de variable omitida tendrá que ser el mismo para los dos.
- *Que otras variables control podrían utilizarse?*

2. ¿Forma functional equivocada?

- Se han utilizado formas funcionales diferentes a la lineal en los dos casos, California and Mass.
- Poca evidencia de efectos no lineales
- No parece que sea una gran amenaza en este caso.

3. ¿Sesgo de Error en Variables?

- Los datos son administrativos por lo que es difícil que se hayan producido riesgos relevantes.
- REM es una medida global de distrito por lo que los estudiantes que hacen el examen pueden no entrar en el cálculo de esa variable – es un tipo de error complicado.
- Idealmente nos gustaría n datos sobre estudiantes individuales según el grado.

4. ¿Sesgo de selección muestral?

- La muestra se toma a partir de todos los colegios públicos de los distritos (en California y en Mass.) – No puede haber datos perdidos por lo que no hay razón para pensar que pueda haber un sesgo de selección de muestra.

5. ¿Sesgo de Simultaneidad?

- La política de subvenciones para igualar basada en las calificaciones puede generar causalidad simultanea.
- Esta política no estaba vigente ni en California ni en Mass. cuando se recogieron las muestras por lo que no cabe pensar en la existencia del sesgo de simultaneidad.

3. Contrastes de esfericidad

(Suplemento docente 3: Contrastes de esfericidad)

4. Criterios de información

(Suplemento docente 4: Criterios de Selección)