

Examen Parcial 1
Econometría 1
Dr. Francisco Cabrera
OT2024

Centro de Investigación y Docencia Económicas (CIDE)

Fecha: 14 de octubre 2024

Instrucciones: coloque su nombre en todas las hojas de respuesta; responda a todas las preguntas; puede utilizar calculadora; no se permiten celulares o laptops; no responda sobre las hojas de preguntas. Tiene 1 hora y 50 minutos para responder.

Pregunta 1 (20%)

En esta pregunta analizaremos la relación entre el logaritmo de los ingresos semanales y los años de educación. Usando una muestra de 470 individuos del Estudio Longitudinal Nacional de la Juventud de EE.UU., obtenemos los siguientes resultados de una regresión del logaritmo del salario semanal sobre los años de educación, experiencia, experiencia al cuadrado y una constante.

$$\log(\text{ingresos}_i) = 4.016 + 0.092 \cdot \text{educacion}_i + 0.079 \cdot \text{experiencia}_i - 0.002 \cdot \text{experiencia}_i^2$$

(0.222) (0.008) (0.025) (0.001)

- a) La desviación estándar de la educación de los individuos en nuestra base de datos es de 3 con media de 12 ¿Cuál es el efecto marginal de aumentar dos desviaciones estándar la educación en los log-ingresos de i ?
- b) Construya un intervalo de confianza al 95% del efecto marginal de la educación en los log-ingresos.
- c) ¿Qué supuestos se necesitan cumplir para asegurar que el efecto de la educación en log-ingresos es *causal* y que el intervalo de confianza estimado es válido?
- d) Interprete el efecto de pasar de 10 años de experiencia a 11, y de pasar de 2 a 3 años de experiencia, en los ingresos del individuo i .
- e) En la ecuación estimada, una variable omitida es la habilidad ¿esto sesga β_1 ? De ser así, demuéstrela e indique la dirección del sesgo.
- f) Suponga que quiere controlar el efecto de la habilidad (no observable) y que en su base de datos puede identificar a gemelos idénticos –que comparten el 100% de sus genes y viven la misma casa ¿qué regresión estimaría para obtener el efecto de la educación en ingresos libre del sesgo de habilidad?

Pregunta 2 (20%)

Para estimar el efecto del salario mínimo en el empleo, David Card y Alan Krueger recopilaron datos sobre empleo en restaurantes i de comida rápida en dos estados vecinos, Nueva Jersey y Pensilvania y en dos momentos del tiempo: antes y después de que entrara en vigor un aumento en el salario mínimo legal en Nueva Jersey. Antes de eso, el salario mínimo era el mismo en Nueva Jersey y Pensilvania. Las variables con que cuenta son: y_i el número de empleados en el restaurante i , nj_i un indicador de que el restaurante está en Nueva Jersey, y $post_i$ una variable dummy que identifica el empleo en el periodo después del cambio en el salario mínimo.

- a) Suponga que sólo tiene datos de Nueva Jersey ¿Cuál sería la regresión para estimar el efecto del cambio en el salario mínimo en el empleo? ¿cómo interpretaría las variables incluidas?
- b) Suponga que sólo tenemos datos de empleo después del cambio en salario mínimo en ambos estados y no antes ¿Qué regresión estimaría y cómo interpretaría las variables incluidas?
- c) ¿cuál es el problema con ambas estimaciones? sea breve y específico/a.
- d) interprete δ_2 la siguiente regresión:

$$y_i = \beta_0 + \delta_0 nj_i + \beta_1 post_i + \delta_2 nj_i \times post_i$$

- e) Dibuje las rectas de regresión que surgen de la estimación en (d) y discuta si el efecto del salario mínimo en el empleo es el esperado.

Pregunta 3 (30%)

Contexto:

¿Cuál es la importancia relativa del “nature” (genes) vs. “nurture” (contexto familiar y social) a la hora de determinar resultados económicos? Esta parte del examen es sobre examinar los resultados de regresiones utilizando una base de datos sobre niños/as nacidos en Corea del Sur, adoptados por padres en Estados Unidos entre 1964 y 1985. La idea es que los niños adoptados no heredan genes de sus padres adoptivos y sólo están expuestos al contexto socioeconómico de éstos.

A los padres adoptivos en Estados Unidos “se les asignó”, a través de una agencia de adopción, al menos a un niño/a coreano sin que pudieran “escogerlo” (e.g. “first come first served”). La base de datos contiene información sobre padres e hijos/as adoptados y no adoptados (“naturales”) en el momento de la adopción y al final del estudio, cuando los hijos ya son adultos. Algunos hogares tienen a más de un niño/a coreano adoptado que no están relacionados entre sí (no son hermanos/as). Al final del estudio, los hijos naturales y adoptados (ya adultos) en esta base de datos tienen al menos 25 años de edad.

Las variables de interés en la base de datos se presentan a continuación:

Variables en la base datos sobre hijos naturales y adoptados y sus padres.

Variables al momento de adopción	Definición
Adopted	=1 si adoptado, 0 si no adoptado (natural)
Weight at adoption	Peso del hijo/a al ser adoptado
Height at adoption	Estatura del hijo/a al ser adoptado
Variables al final del estudio (cuando niños adoptados y naturales son adultos)	
Child's education	Años de educación alcanzados por el niño cuando es adulto.
College grad	=1 si el niño, cuando es adulto, se graduó de la universidad, =0 si no.
Child's BMI	Índice de Masa Corporal del niño cuando es adulto (en unidades). Es igual al peso del niño (cuando es adulto) en kilos, dividido entre su estatura en metros al cuadrado (km/m ²)
Child drinks	=1 si el niño consume alcohol cuando es adulto, 0 si no.
Child's income	Ingreso del niño cuando es adulto.
Variables sobre los padres (en el año pre-adopción)	
Fathers's education	Años de educación del padre
Mothers's education	Años de educación de la madre
Log Parent's Income	Logaritmo natural del ingreso de los padres
Mother's BMI	Índice de Masa Corporal de la madre (unidades)
Father's BMI	Índice de Masa Corporal del padre (unidades)
Mother Drinks	=1 si la madre consume alcohol, =0 si no.
Father Drinks	=1 si el padre consume alcohol, =0 si no
Year dummy variables	Set de variables dummy indicando el año de adopción del niño. El primer año en que se adoptó a cualquier niño, es el año base.

- a) Usted decide estimar la relación abajo mostrada, utilizando toda la muestra en la base de datos:

$$childseducation_i = \beta_0 + \beta_1 fatherseducation_i + \beta_2 \log(parentsincome_i) + \mu$$

interprete el efecto de $\log(parentsincome)$ en $childseducation$.

- b) en la relación:

$$childseducation_i = \beta_0 + \beta_1 fatherseducation_i + v$$

¿Existe un sesgo en β_1 por omitir $\log(parentsincome_i)$? Demuéstrelo teóricamente.

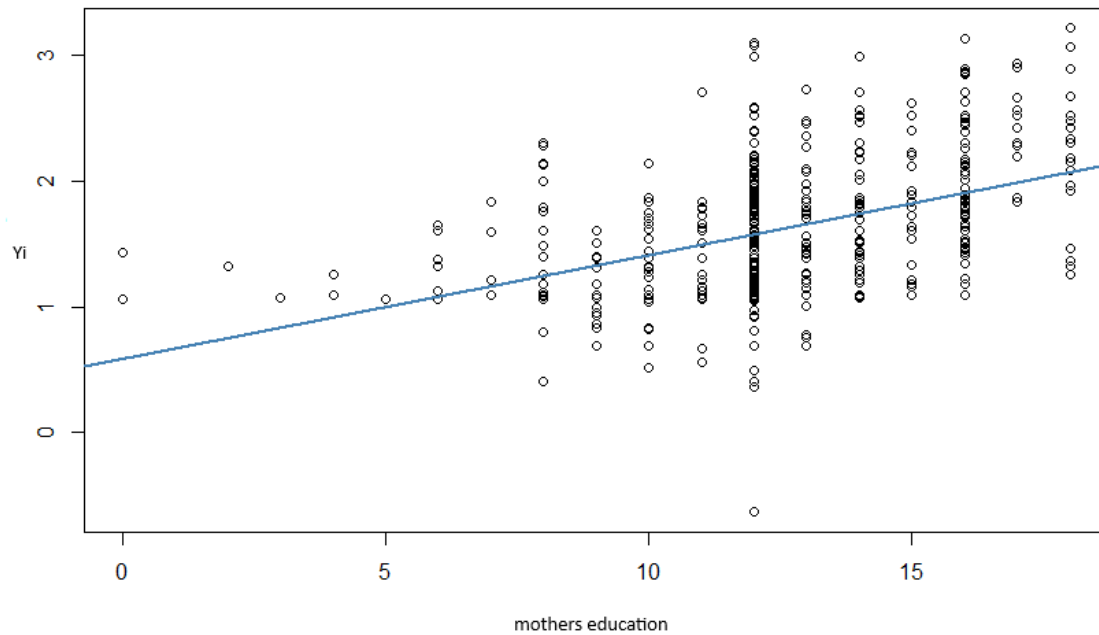
- c) Si, empíricamente: $\beta_2 = 0.5$ y $cov[\log(parentsincome), fatherseducation] = 0.3$ ¿De qué tamaño sería el sesgo por omisión si se estima la relación en b)?

- d) Escriba la ecuación teórica, si quisiera saber la diferencia del efecto de la educación del padre (*fatherseducation*) en los ingresos futuros de su hijo ($\log(\text{childsincome})$) para hijos con madres que beben y que no beben (*motherdrinks*), dejando fijo o controlando por $\log(\text{parentsincome})$.
- e) Dibuje la(s) recta(s) de regresión que se obtendría(n) en (d) sin considerar $\log(\text{parentsincome})$. Considere que el efecto de las madres que consumen alcohol en los hijos es negativo, pero que se compensa con mayores niveles de educación del padre.
- f) Escriba la ecuación teórica que necesitaría estimar si quisiera saber el efecto de tener una madre que consume alcohol (x_{i1}) en los ingresos futuros de su hijo ($\log(\text{childsincome})$) para niños adoptados y no adoptados (*adopted*) dejando fijo o controlando por $\log(\text{parentsincome})$.
- g) En teoría, los hábitos de consumo de alcohol de la madre podrían influir en el ingreso de sus hijos a través de dos vías: porque les genera un ambiente adverso mientras crecen y porque les heredan genéticamente dichos hábitos que pueden afectar su desempeño profesional. Si usted quiere probar que la genética no importa, dada la ecuación escrita en (f) ¿Cuál sería la hipótesis nula que desea probar y el test estadístico correspondiente?
- h) Dibuje en dos gráficos las rectas de regresión que se obtendrían de (g) cuando la H_0 planteada se rechaza y cuando H_0 no se rechaza (hint: no considere $\log(\text{parentsincome})$ y asuma *motherdrinks* “continua”).

Pregunta 4 (30%)

En la Tabla 1 (abajo) se presentan los resultados de distintas regresiones utilizando los datos de adopciones en EE.UU descritos en la pregunta 3. Estos utilizan la submuestra de niños adoptados (se excluyen hijos naturales).

- a) ¿Cuál es el efecto promedio en los años de educación de los hijos de madres que beben alcohol (*motherdrinks*) ¿es el efecto estadísticamente significativo a niveles convencionales de confianza?
- b) ¿Cuál es el efecto promedio en la probabilidad de que los hijos de madres que beban alcohol también lo consuman? ¿qué puede concluir respecto a la discusión “nature” vs. “nurture”? dados los resultados en (a) y (b).
- c) Los errores estándar reportados en la Tabla 1 (columna 1) para el estimador de $\log(\text{parentsincome})$ provienen de la siguiente relación (abajo) ¿cuál es la fórmula correcta para su estimación y con qué supuesto de Gauss-Markov se relaciona este caso?



- d) Usted sospecha que el BMI de la madre puede causar problemas de multicolinealidad ¿cuál sería la razón intuitiva (en el contexto específico de esta regresión) y ¿cómo podría probar empíricamente que (no) existe dicho problema?
- e) Si su énfasis está en estudiar el efecto de *motherdrinks* considere el test F reportado en la Tabla 1 y responda ¿qué modelo escogería entre la columna 1 y 2? justifique su respuesta.

Table 1. Regression of adoptee outcome variables on pre-adoption parental characteristics

Dependent variable	(1) Child's Years of Education	(2) Child's Years of Education	(3) College Grad	(4) Log Child's Income	(5) Child's BMI	(6) Child Drinks
Regressors:						
Mother's Education	0.097** (0.027)	0.084** (0.031)	0.021* (0.008)	0.016 (0.013)	-0.081 (0.061)	0.010 (0.009)
Father's Education	-0.001 (0.032)	-0.041 (0.055)	-0.004 (0.007)	-0.004 (0.011)	-0.037 (0.052)	0.010 (0.007)
Log parent's income	-0.018 (0.113)	-0.005 (0.032)	0.011 (0.027)	0.024 (0.040)	-0.412 (0.219)	0.015 (0.028)
Mother's BMI	-0.088** (0.024)	0.180 (0.183)	-0.017** (0.004)	-0.004 (0.006)	0.006 (0.028)	-0.001 (0.004)
Father's BMI	0.007 (0.020)	-0.008 (0.112)	-0.000 (0.004)	-0.000 (0.007)	-0.004 (0.038)	0.004 (0.004)
(Mother's BMI) ²		-0.091 (0.118)				
(Father's BMI) ²		-0.081 (0.202)				
(Mother's BMI) x (Father's BMI)		0.274 (0.206)				
Mother Drinks	-0.039 (0.205)	-0.715** (0.175)	-0.043 (0.046)	-0.007 (0.066)	-0.345 (0.392)	0.135** (0.045)
Father Drinks	0.263 (0.212)	0.000 (0.002)	0.050 (0.048)	0.030 (0.070)	0.580 (0.396)	0.061 (0.046)
Child is Male	-0.723** (0.177)	-0.004 (0.003)	-0.159** (0.041)	-0.259** (0.059)	1.927** (0.301)	0.068 (0.040)
Constant	16.902** (1.063)	0.002 (0.004)	0.766** (0.264)	3.758** (0.466)	31.183** (2.350)	0.121 (0.315)
Year binary variables?	Yes	Yes	Yes	Yes	Yes	Yes
F-statistic testing: (Mother's BMI) ² , (Father's BMI) ² , (Mother's BMI) x (Father's BMI) = 0 (p-value)		0.57 (0.634)				
Observations	897	897	897	874	878	893
Adjusted R-squared	0.03	0.03	0.01	0.06	0.04	0.04

Notes: All regressions are estimated by OLS. Clustered standard errors are given in parentheses, where the clustering occurs at the level of the family. * significant at 5%; ** significant at 1%.