

Examen Final

Profesor: Juan Palomino

December 6, 2022

Instrucciones

- (1) El examen se encuentra habilitado en el sistema PAIDEIA a partir de las 10:00am y tiene hasta las 12:00pm para entregarme los cuadernillos. No se recibirá ningún cuadernillo pasado esa hora.
- (2) Pueden subir la solución de su programación a PAIDEIA hasta la hora final del examen.
- (3) Redacción y ortografía **SERÁN TOMADAS EN CUENTA** para la calificación de la prueba. **SI EL ENVÍO DEL MANUSCRITO NO ES LEGIBLE RECIBIRÁN LA PENALIDAD DE MENOS UN PUNTO.**
- (4) Si se evidencia plagio en las respuestas, automáticamente la prueba será anulada y remitida a las autoridades correspondientes de la Facultad.
- (5) La pregunta sobre Heterocedasticidad lo pueden hacer en excel, Stata, RStudio, Colab u otro software. Pero solo quiero sus respuestas numéricas y su interpretación. En esta pregunta no se evaluará el procedimiento.

Teoría

Responder brevemente los siguientes conceptos:

- (1) Explicar las causas de la endogeneidad. **(1 punto)**
- (2) Defina lo que es un instrumento y las propiedades que debe cumplir. **(1 punto)**

Modelo de Ecuaciones Simultáneas

Considere el siguiente modelo de demanda y oferta de dinero:

$$\begin{aligned}\text{Demanda de dinero } M_t^d &= \beta_0 + \beta_1 Y_t + \beta_2 R_t + \beta_3 P_t + \epsilon_{1t} \\ \text{Oferta de dinero } M_t^s &= \alpha_0 + \alpha_1 Y_t + \epsilon_{2t}\end{aligned}$$

en donde M =dinero, Y =ingreso, R =tasa de interés, P =precio. Suponga que R y P están predeterminados.

- (a) Halle la forma reducida del modelo. **(1 punto)**
- (b) ¿Está identificada la función de demanda? **(1 punto)**
- (c) ¿Está identificada la función de oferta? **(1 punto)**
- (d) ¿Cuál método se utilizaría para estimar los parámetros de la(s) ecuación(es) identificada(s)? Por qué? Explicar los pasos de la estimación. **(1.5 puntos)**
- (e) Suponga que se modifica la función de oferta agregando las variables explicativas Y_{t-1} y M_{t-1} ¿Qué sucede con el problema de la identificación? ¿Se utilizará aún el método que utilizó en (c)? ¿Por qué sí o por qué no? Explicar los pasos de la estimación. **(1.5 puntos)**

Heterocedasticidad

Considere la base de datos `base_anemia.dta` que se encuentra en PAIDEIA. Esta base de datos cuenta con información sobre la anemia en el Perú, construido a partir de la Encuesta Nacional Demográfica de Salud (ENDES). A partir de esta base de datos, se pide estimar el siguiente modelo:

$$\text{hemoglobina} = \beta_0 + \beta_1 \text{sexo} + \beta_2 \text{edadmeses} + \beta_3 \text{pesoalnacer} + \beta_4 \text{educmadre} + \beta_5 \text{area} + \epsilon \quad (1)$$

donde:

Variables	Valor
sexo	sexo del niño 1: Mujer 0: Hombre
edadmeses	Edad del niño en meses
pesoalnacer	Peso del niño al nacer (kilos)
educmadre	Nivel Educativo de la madre 0: Primaria o menos 1: Secundaria 2: Superior
area	1: Urbano 0: Rural

Utilizando cualquier software (STATA, RStudio o Colab), se le pide:

- 1. Realizar el test de White y el test de Breusch-Pagan para detectar heterocedasticidad. ¿Existe heterocedasticidad en el modelo? **(4 puntos)**

Pista: Recuerden que si una variable tiene n categorías, debe haber al menos $n - 1$ parámetros estimados para esa variable.

Variables Instrumentales

Card (1994) utilizó datos de salarios y educación para una muestra de hombres en 1976 para estimar el retorno a la educación utilizando la siguiente ecuación de Mincer:

$$\ln w_i = \beta_1 + \beta_2 S_i + \beta_3 E_i + \beta_4 E_i^2 + h_i' \delta + \epsilon_i$$

donde $\ln w_i$ denota el logaritmo de salarios individuales, S_i denota años de escolaridad y E_i denota años de experiencia. En ausencia de información sobre la experiencia real, E_i a veces se reemplaza por “experiencia potencial”, medida como $\text{age}_i - S_i - 6$, suponiendo que las personas empiezan la escuela a la edad de 6 años. Esta especificación generalmente se aumenta con variables explicativas adicionales, h_i , que uno quiere controlar, como dummies regionales y raciales.

Se usa los datos de 3010 hombres de la Encuesta Longitudinal Nacional de Hombres Jóvenes de EE. UU. (NLS-Y), empleados en Card (1994). En esta encuesta, se sigue a un grupo de individuos desde 1966 cuando tenían entre 14 y 24 años, y se los entrevistaba en varios años consecutivos. La información del mercado laboral que se utiliza abarca 1976. Varios resultados se presentan en la Tabla 2.1. Los controles utilizados son los siguientes:

- *experience*: experiencia potencial en años
- *experience*²: experiencia potencial al cuadrado
- *black*: dummy para raza negra (1 si el individuo es de raza negra)
- *south*: dummy si individuo vive en el sur
- *smsa*: dummy para residencia en area metropolitana
- *smsa66*: dummy para residencia en area metropolitana en 1966
- *reg662* – *reg668*: un set de variables dummies regionales

1. La columna (1) de la Tabla 2.1 muestra el resultado de una regresión MCO. Interprete el coeficiente de escolaridad (Educ), que mide los años de escolaridad. ¿Es este coeficiente potencialmente endógeno? ¿Por qué? **(1 punto)**
2. En la columna (2) controlamos por la habilidad (IQ). ¿Cómo interpretas el resultado para el coeficiente de escolaridad (Educ)? **(1 punto)**
3. La columna (3) presenta una estimación MC2E donde el instrumento para Educ es Nearc4, que es una variable dummy para determinar si alguien creció cerca de una universidad durante cuatro años como una variable instrumental para la educación. ¿Este modelo está subidentificado, identificado o sobreidentificado? Comente el coeficiente para Educ. **(2 puntos)**
4. La columna (4) muestra la estimación de la ecuación de forma reducida. ¿Nearc4 es un instrumento válido? **(1 punto)**
5. Explica por qué Nearc4 podría correlacionarse con la habilidad (que está en el término de error). **(1 punto)**
6. Explique la idea de las columnas 5 y 6 de la tabla de acuerdo con su respuesta en (5). ¿Qué concluye sobre estos resultados? **(2 puntos)**

Tabla 2.1. Retornos a la Educación

Variables	Dependent variable:					
	log(wage) (1)	log(wage) (2)	log(wage) (3)	educ (4)	IQ (5)	IQ (6)
Educ	0.075*** (0.003)	0.070*** (0.005)	0.132** (0.055)			
Nearc4				0.320*** (0.088)	2.596*** (0.745)	0.348 (0.814)
Experience	0.085*** (0.007)	0.095*** (0.009)	0.108*** (0.024)	-0.413*** (0.034)		
<i>Experience</i> ²	-0.002*** (0.0003)	-0.003*** (0.0005)	0.002*** (0.0003)	0.001 (0.002)		
Black	-0.199*** (0.018)	-0.148*** (0.027)	-0.147*** (0.054)	-0.936*** (0.094)		
South	-0.148*** (0.026)	-0.100*** (0.032)	-0.145*** (0.027)	-0.052 (0.135)		
SMSA	0.136*** (0.020)	0.123*** (0.024)	0.112*** (0.032)	0.402*** (0.105)		
reg661	-0.119*** (0.039)	-0.122*** (0.044)	-0.108*** (0.042)	-0.210 (0.202)		2.892 (1.797)
reg662	-0.022 (0.028)	-0.019 (0.032)	-0.007 (0.033)	-0.289** (0.147)		3.991*** (1.294)
reg663	0.026 (0.027)	0.004 (0.031)	0.040 (0.032)	-0.238* (0.143)		1.333 (1.259)
reg664	-0.063* (0.036)	-0.073* (0.040)	-0.058 (0.038)	-0.093 (0.186)		2.349 (1.635)
reg665	0.009 (0.036)	0.002 (0.043)	0.038 (0.047)	0.483** (0.188)		-5.584*** (1.322)
reg666	0.022 (0.040)	0.037 (0.051)	0.055 (0.053)	-0.513** (0.210)		-4.529*** (1.698)
reg667	-0.001 (0.039)	-0.031 (0.047)	0.027 (0.049)	-0.427** (0.206)		-5.502*** (1.520)
reg668	-0.175*** (0.046)	-0.170*** (0.052)	-0.191*** (0.051)	0.314 (0.242)		-0.033 (2.111)
smsa66	0.026 (0.019)	0.036 (0.023)	0.019 (0.022)	0.025 (0.106)		1.089 (0.809)
IQ		0.002*** (0.001)				
Constant	4.739*** (0.072)	4.501*** (0.106)	3.774*** (0.935)	16.849*** (0.211)	100.611*** (0.627)	101.882*** (1.292)
Observations	3010	2061	3010	3010	2061	2061
R^2	0.300	0.237	0.238	0.477	0.006	0.063
Adjusted R^2	0.296	0.231	0.234	0.474	0.005	0.058

Note: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$