

**Duración del examen: 2 horas y 30 minutos.**

**Nota importante: Alguna información contenida en las salidas es redundante.**

**PREGUNTA 1** Utilizando los datos WAGE2.RAW se ha estimado el modelo

$$\log(\text{wage}) = \beta_0 + \beta_1 \text{educ} + \beta_2 \text{exper} + \beta_3 \text{tenure} + \beta_4 \text{married} + \beta_5 \text{black} + \beta_6 \text{south} + \beta_7 \text{urban} + U.$$

En las salidas de GRETL que se aportan a continuación se estima este y otros tres modelos utilizando mínimos cuadrados ordinarios. Utilice los listados para contestar las siguientes preguntas

- (0,5 puntos)** Manteniendo los demás factores fijos, en el contexto del modelo en el enunciado (Modelo 1 en las salidas adjuntas) ¿cuál es la diferencia aproximada entre el salario mensual de las personas de raza negra y las que no lo son? ¿Es esta diferencia estadísticamente significativa?
- (0,5 puntos)** Añadir las variables  $\text{exper}^2$  y  $\text{tenure}^2$  en la ecuación y demostrar que no son conjuntamente estadísticamente significativas incluso al 10 por ciento.
- (0,5 puntos)** Ampliar el modelo original para que el rendimiento de la educación dependa de la raza. Escriba el modelo a considerar y contraste la significatividad de esta dependencia.
- (0,5 puntos)** Partiendo del modelo inicial, permitir que el salario difiera entre cuatro grupos de individuos: negros casados, negros solteros, no negros casados y no negros solteros. Expresar el modelo de la forma más concisa posible y proporcionar el valor estimado de la diferencia salarial entre negros casados y no negros casados con los mismos valores para el resto de variables explicativas.

Modelo 1: MCO, usando las observaciones 1–935

	Variable dependiente: lwage			
	Coefficiente	Desv. Típica	Estadístico $t$	Valor p
const	5,39550	0,113225	47,6529	0,0000
educ	0,0654307	0,00625040	10,4683	0,0000
exper	0,0140430	0,00318519	4,4089	0,0000
tenure	0,0117473	0,00245297	4,7890	0,0000
married	0,199417	0,0390502	5,1067	0,0000
black	-0,188350	0,0376666	-5,0004	0,0000
south	-0,0909037	0,0262485	-3,4632	0,0006
urban	0,183912	0,0269583	6,8221	0,0000
Suma de cuad. residuos	123,8185	D.T. de la regresión	0,365471	
$R^2$	0,252558	$R^2$ corregido	0,246914	

Modelo 2: MCO, usando las observaciones 1–935

	Variable dependiente: lwage			
	Coefficiente	Desv. Típica	Estadístico $t$	Valor p
const	5,35868	0,125914	42,5581	0,0000
educ	0,0642761	0,00631148	10,1840	0,0000
exper	0,0172146	0,0126138	1,3647	0,1727
tenure	0,0249291	0,00812966	3,0664	0,0022
married	0,198547	0,0391103	5,0766	0,0000
black	-0,190664	0,0377011	-5,0572	0,0000
south	-0,0912153	0,0262356	-3,4768	0,0005
urban	0,185424	0,0269585	6,8781	0,0000
exper <sup>2</sup>	-0,000113801	0,000531871	-0,2140	0,8306
tenure <sup>2</sup>	-0,000796448	0,000471013	-1,6909	0,0912
Suma de cuad. residuos	123,4210	D.T. de la regresión	0,365278	
$R^2$	0,254958	$R^2$ corregido	0,247709	

Modelo 3: MCO, usando las observaciones 1–935

Variable dependiente: lwage				
	Coefficiente	Desv. Típica	Estadístico $t$	Valor p
const	5,37482	0,114703	46,8587	0,0000
educ	0,0671153	0,00642769	10,4416	0,0000
exper	0,0138259	0,00319063	4,3333	0,0000
tenure	0,0117870	0,00245289	4,8054	0,0000
married	0,198908	0,0390474	5,0940	0,0000
black	0,0948093	0,255399	0,3712	0,7106
south	-0,0894495	0,0262769	-3,4041	0,0007
urban	0,183852	0,0269547	6,8208	0,0000
educ*black	-0,0226237	0,0201827	-1,1209	0,2626
Media de la vble. dep.	6,779004	D.T. de la vble. dep.		0,421144
Suma de cuad. residuos	123,6507	D.T. de la regresión		0,365420
$R^2$	0,253571	$R^2$ corregido		0,247122
$F(8, 926)$	39,32158	Valor p (de $F$ )		$4,35 \cdot 10^{-54}$

Modelo 4: MCO, usando las observaciones 1–935

Variable dependiente: lwage				
	Coefficiente	Desv. Típica	Estadístico $t$	Valor p
const	5,40379	0,114122	47,3509	0,0000
educ	0,0654751	0,00625302	10,4710	0,0000
exper	0,0141462	0,00319103	4,4331	0,0000
tenure	0,0116628	0,00245795	4,7449	0,0000
married	0,188915	0,0428777	4,4059	0,0000
black	-0,240820	0,0960229	-2,5079	0,0123
south	-0,0919894	0,0263212	-3,4949	0,0005
urban	0,184350	0,0269778	6,8334	0,0000
married*black	0,0613538	0,103275	0,5941	0,5526
Media de la vble. dep.	6,779004	D.T. de la vble. dep.		0,421144
Suma de cuad. residuos	123,7714	D.T. de la regresión		0,365599
$R^2$	0,252842	$R^2$ corregido		0,246388
$F(8, 926)$	39,17047	Valor p (de $F$ )		$6,78e-54$

**PREGUNTA 2.** Considere un modelo de ecuaciones simultáneas en la "forma de oferta y demanda", donde la misma variable dependiente  $y_1$  (típicamente, la "cantidad") en el lado izquierdo de cada ecuación:

$$\begin{aligned} Y_1 &= \alpha_1 Y_2 + \beta_1 Z_1 + U_1 \\ Y_1 &= \alpha_2 Y_2 + \beta_2 Z_2 + U_2. \end{aligned}$$

Como es habitual las  $Z$ 's son exógenas.

- (0,5 puntos)** Si  $\alpha_1 \neq 0$ ,  $\alpha_2 \neq 0$  y  $\alpha_1 \neq \alpha_2$ , encontrar la forma reducida de  $Y_1$ . En este caso ¿tiene  $Y_2$  forma reducida?.
- (0,5 puntos)** ¿Es razonable suponer que  $\alpha_1 \neq \alpha_2$  en un sistema de oferta y demanda?
- (0,5 puntos)** ¿Qué ecuación nunca puede estar identificada cuando  $\beta_2 = 0$ ?

**PREGUNTA 3:** Se considera un modelo lineal de probabilidad para explicar la probabilidad de concesión de una hipoteca:

$$E(Y|X_1, X_2) = \beta_0 + \beta_1 X_1 + \beta_2 X_2$$

donde:

$Y$  toma el valor 1 si NO le conceden la hipoteca;

$X_1 = DEUDA$  = tasa de endeudamiento en tanto por uno (sin incluir el préstamo hipotecario considerado) del solicitante;

$X_2 = NEG$  = toma el valor 1 si el solicitante es negro y cero si es blanco.

Con una muestra de 2380 solicitudes de hipotecas en Boston (EE.UU.), se ha estimado el modelo por mínimos cuadrados ordinarios. Entre corchetes  $[\cdot]$  se encuentran los errores estándar robustos en presencia de heterocedasticidad.

$$\hat{Y} = -0,091 + 0,559X_1 + 0,177X_2.$$

(0,029)	(0,089)	(0,025)
[0,040]	[0,129]	[0,125]

- a. **(0,5 puntos)** Interprete el coeficiente  $\beta_1$ . ¿Cuál es la probabilidad de que le concedan la hipoteca a un solicitante blanco con una tasa de endeudamiento del 30%?
- b. **(0,5 puntos)** ¿Existe heterocedasticidad en este modelo? Proporcionar el valor del estadístico para contrastar si el ser blanco es significativo en la concesión de una hipoteca y realice el contraste al 1% de significación.

**PREGUNTA 4:** La siguiente ecuación explica las horas semanales de televisión vistas por un niño en función de su edad, la educación de la madre, la educación del padre y el número de hermanos:

$$tvhours^* = \beta_0 + \beta_1 age + \beta_2 age^2 + \beta_3 mothereduc + \beta_4 fathereduc + \beta_5 sibs + U.$$

Nos preocupa que  $tvhours^*$  se mida con error en nuestra encuesta. Supongamos que  $tvhours$  son las horas de televisión declaradas en la encuesta.

- a. **(0,5 puntos)** ¿Qué debe cumplirse en esta aplicación para que el estimador de mínimos cuadrados ordinarios sea insesgado?
- b. **(0,5 puntos)** ¿Qué ocurriría si la variable  $mothereduc$  estuviera también medida con error bajo los supuestos del apartado anterior?

**PREGUNTA 5:** Existe evidencia de que el comportamiento laboral de las mujeres está determinado en particular por sus decisiones de fertilidad. Para evaluar el impacto de la fertilidad (medido por el número de niños) en las horas trabajadas, nos concentramos en la siguiente especificación:

$$\begin{aligned} HRS = & \beta_0 + \beta_1 WHITE + \beta_2 BLACK + \beta_3 HISPAN \\ & + \beta_4 HIGHSC + \beta_5 UNIV + \beta_6 AGE + \beta_7 AGE2 \\ & + \beta_8 SPOUSE + \beta_9 (SPOUSE \times AGE) + \beta_{10} (SPOUSE \times AGE2) \\ & + \beta_{11} NCHILD + U \end{aligned} \tag{E.1}$$

donde, para cada mujer:

- HRS = número de horas trabajadas a la semana;
- WHITE = variable binaria que vale 1 si la mujer es blanca y cero en caso contrario;
- BLACK = variable binaria que vale 1 si la mujer es negra y cero en caso contrario;
- HISPAN = variable binaria que vale 1 si la mujer es hispana y cero en caso contrario;
- HIGHSC = variable binaria que vale 1 si la mujer ha completado solamente la educación secundaria y cero en caso contrario;
- UNIV = variable binaria que vale 1 si la mujer tiene un título universitario y cero en caso contrario;
- AGE = edad en años;
- AGE2 = edad al cuadrado;
- SPOUSE = variable binaria que vale 1 si la mujer tiene un cónyuge que convive en su domicilio y cero en caso contrario;
- NCHILD = número de hijos menores de 18 años que conviven en su domicilio.

**Nota:** Hay cuatro grupos étnicos mutuamente excluyentes: blanca, negra, hispana y asiática. Sólo hay tres niveles de educación (mutuamente excluyentes): educación primaria o menos, educación secundaria, educación universitaria. Para estimar esta ecuación, utilizamos datos de 69852 mujeres con hijos del censo de los EE.UU. de 1980.

Además, sabemos que las decisiones de fertilidad están correlacionadas con características inobservables que afectan a su vez a las decisiones laborales. Por ejemplo, aquellas mujeres mejor preparadas para el mercado laboral pueden tener en promedio no sólo mayores salarios (y por tanto un coste de oportunidad de asumir tareas del hogar más alto), sino también un mayor coste personal asociado al cuidado de los hijos. Por tanto, esperaríamos que

$$Cov(U, NCHILD) \neq 0.$$

mientras que el resto de las variables del lado derecho de (E.1) **no están correlacionadas** con las variables inobservables ( $U$ ).

Además de las variables arriba mencionadas, tenemos información sobre si la mujer ha experimentado un parto múltiple ( $MB$ ), es decir, si ha tenido mellizos, trillizos, cuatrillizos o quintillizos en el mismo parto. Podemos por tanto definir la variable Parto Múltiple (*Multiple Births*)  $MB$ :

$MB$  = variable binaria que vale 1 si la mujer ha tenido un parto múltiple y cero en caso contrario.

Además, sabemos que  $Cov(MB, U) = 0$ .

Al final del enunciado se presentan diferentes estimaciones:

- a. **(0,5 puntos)** Suponiendo que  $Cov(MB, U) = 0$ , contraste si  $MB$  es un instrumento válido. Contraste al 1% de significación si  $NCHILD$  es exógena.
- b. **(0,5 puntos)** Suponiendo que  $Cov(NCHILD, U) \neq 0$  y que  $MB$  es un instrumento válido, proporcione el valor estimado de la disminución promedio de las horas trabajadas que conlleva un niño adicional, para una raza, nivel de educación, edad y estatus marital fijos.

<b>SALIDA 1</b>					
Variable Dependiente: HRS					
Método: Mínimos Cuadrados Ordinarios					
Muestra: 69852					
Variables incluidas: 69852					
Variable	Coefficiente	Std. Error	t-Statistic	Prob.	
WHITE	1.3224	0.7901	1.67	0.094	
BLACK	3.3607	0.8005	4.20	0.000	
HISPAN	-0.5399	1.0007	-0.54	0.590	
HIGHSC	4.4627	0.1770	25.22	0.000	
UNIV	4.3868	0.2005	21.88	0.000	
AGE	1.4505	0.0288	50.37	0.000	
AGE2	-0.0207	0.0006	-36.04	0.000	
SPOUSE	0.4258	0.4343	0.98	0.327	
SPOUSE * AGE	-0.5668	0.0335	-16.90	0.000	
SPOUSE * AGE2	0.0101	0.0006	15.42	0.000	
NCHILD	-1.8506	0.0610	-30.32	0.000	
C	0.2952	0.8568	0.34	0.730	
R-cuadrado		0.1344			
R-cuadrado Ajustado		0.1343			
S.E. of regression		335.4738			

---

---

**SALIDA 2**

---

---

Variable Dependiente: *NCHILD*

Método: Mínimos Cuadrados Ordinarios

Muestra: 69852

Variables incluidas: 69852

---

---

Variable	Coefficiente	Std. Error	t-Statistic	Prob.
WHITE	-0.3527	0.0482	-7.32	0.000
BLACK	0.2822	0.0488	5.78	0.000
HISPAN	0.0638	0.0610	1.05	0.296
HIGHSC	-0.4758	0.0106	-44.68	0.000
UNIV	-0.5453	0.0121	-45.22	0.000
AGE	0.0801	0.0017	46.26	0.000
AGE2	-0.0012	0.00003	-35.08	0.000
SPOUSE	0.2622	0.0265	9.90	0.000
SPOUSE * AGE	-0.0095	0.0020	-4.65	0.000
SPOUSE * AGE2	0.0001	0.00004	3.42	0.001
MB	1.2192	0.0255	47.79	0.000
C	1.5957	0.0519	30.73	0.000

---

---

R-cuadrado	0.1568
R-cuadrado Ajustado	0.1566

---

---

---

---

**SALIDA 3**

---

---

Variable Dependiente: *HRS*

Método: Mínimos Cuadrados Bietápicos

Muestra: 69852

Variables incluidas: 69852

Lista de Instrumentos: *MB*

---

---

Variable	Variable	Coefficiente	Std. Error	t-Statistic	Prob.
WHITE		1.7045	0.8019	2.13	0.034
BLACK		3.0750	0.8077	3.81	0.000
HISPAN		-0.5925	1.0029	-0.59	0.555
HIGHSC		4.9467	0.2398	20.63	0.000
UNIV		4.9483	0.2746	18.02	0.000
AGE		1.3697	0.0395	34.69	0.000
AGE2		-0.0195	0.0007	-27.66	0.000
SPOUSE		0.1529	0.4446	0.34	0.731
SPOUSE * AGE		-0.5575	0.0337	-16.52	0.000
SPOUSE * AGE2		0.0100	0.0007	15.15	0.000
NCHILD		-0.8362	0.3437	-2.43	0.015
C		-1.3616	1.0209	-1.33	0.182

---

---

R-cuadrado	0.1310
R-cuadrado Ajustado	0.1308

---

---

---

---

**SALIDA 4**

---

---

Variable Dependiente: HRS

Método: Mínimos Cuadrados Ordinarios

Muestra: 69852

Variables incluidas: 69852

---

---

Variable	Coefficiente	Std. Error	t-Statistic	Prob.
WHITE	1.7045	0.8002	2.13	0.033
BLACK	3.07502	0.8061	3.81	0.000
HISPAN	-0.5925	1.0008	-0.59	0.554
HIGHSC	4.9467	0.2393	20.67	0.000
UNIV	4.9483	0.2741	18.05	0.000
AGE	1.3697	0.0394	34.76	0.000
AGE2	-0.0195	0.0007	-27.71	0.000
SPOUSE	0.1529	0.4437	0.34	0.730
SPOUSE * AGE	-0.5575	0.0337	-16.55	0.000
SPOUSE * AGE2	0.0100	0.0007	15.19	0.000
NCHILD	-0.8362	0.3430	-2.44	0.015
RES	-1.0475	0.3486	-3.01	0.003
C	-1.3616	1.0188	-1.34	0.181

---

---

R-cuadrado	0.1345
Adj. R-cuadrado	0.1344

---

---

(NOTA: RES son los residuos de la SALIDA 2)

**VALORES CRÍTICOS**

$Z_{0,025} = 1,96$	$Z_{0,05} = 1,645$	$Z_{0,01} = 2,326$	$Z_{0,005} = 2,576$
$Z_{0,1} = 1,282$	$\chi_{3;0,01}^2 = 11,34$	$\chi_{3;0,05}^2 = 7,82$	$\chi_{5;0,05}^2 = 11,07$
$\chi_{2;0,05}^2 = 5,99$	$\chi_{2;0,01}^2 = 9,21$	$\chi_{6;0,05}^2 = 12,59$	$\chi_{2;0,1}^2 = 4,61$
$\chi_{6;0,01}^2 = 16,81$	$\chi_{4;0,05}^2 = 9,49$	$\chi_{3;0,1}^2 = 6,25$	$\chi_{4;0,01}^2 = 13,28$

$Z$  es la normal de media cero y varianza uno y  $\chi_q^2$  es la chi cuadrado con  $q$  grados de libertad,  $\Pr(Z > Z_\alpha) = \alpha$ ;  
 $\Pr(\chi_q^2 > \chi_{q;\alpha}^2) = \alpha$ .

Nótese que la distribución  $F$  se puede aproximar por la de la  $\chi^2$ . Esto es,  $\chi_q^2 \sim q \cdot F_{q,n}$  cuando  $n$  es grande,  
 $\Pr(\chi_q^2 > \chi_{q;\alpha}^2) \simeq \Pr(q \cdot F_{q,n} > \chi_{q;\alpha}^2)$ .