

EXAMEN FINAL ORDINARIO

Econometría

Universidad Carlos III de Madrid

26/05/21

Escriba su nombre y grupo en cada hoja. Conteste todas las preguntas en 2:30 horas.

1. (40%) Sean $\{Y_i, X_{1i}, X_{2i}\}_{i=1}^n$ observaciones independientes e idénticamente distribuidas como las variables (Y, X_1, X_2) de una población, las cuales mantienen una relación causal de acuerdo con el modelo

$$Y = X_1\beta_1 + X_2\beta_2 + u, \quad (1)$$

donde u es un error con media cero, varianza σ^2 , e independiente de (X_1, X_2) , y β_1 y β_2 son parámetros desconocidos.

- a. (1/5) Demuestre que, si $\mathbb{E}(X_1X_2) = 0$ y $\mathbb{E}(X_1^2) > 0$,

$$\beta_1 = \frac{\mathbb{E}(X_1Y)}{\mathbb{E}(X_1^2)}.$$

- b. (2/5) Derive una expresión para el estimador de mínimos cuadrados ordinarios de β_1 en el modelo (1) como el coeficiente de *MCO* de una regresión simple donde la variable explicativa es el residuo de un ajuste *MCO*, sin constante, de X_1 sobre X_2 .
- c. (2/5) Suponga que X_2 no es observable y que está correlacionada con X_1 . Debemos estimar β_1 en el modelo

$$Y_i = \beta_0 + X_{1i}\beta_1 + v_i, \quad i = 1, \dots, n \quad (2)$$

donde $\beta_0 = \mu_{X_2}\beta_2$ y $v_i = (X_{2i} - \mu_{X_2})\beta_2 + u_i$ ($\mu_{X_2} = \mathbb{E}(X_2)$). Suponga que contamos con un instrumento Z que cumple las propiedades de exogeneidad y relevancia.

- i. (1/3 de 1.c) Demuestre que X_1 es una variable endógena en el modelo (2).
- ii. (2/3 de 1.c) Exprese los coeficientes y el error de la forma estructural (2), β_0 , β_1 y v , en términos de los coeficientes y errores en las ecuaciones de las formas reducidas de Y y X_1 .
2. (30%) El ratio de defectuosos ("scrap rate") es el número de unidades defectuosas (productos que deben ser desechados) de cada 100 unidades producidas. Estamos interesados en utilizar

el ratio de defectuosos para medir el efecto de la formación de los trabajadores sobre la productividad.

Para una muestra aleatoria de empresas se obtiene el siguiente ajuste por *MCO*,

$$\begin{aligned} \ln(\widehat{scrap}_i) &= \underset{(4.57)}{11.74} - \underset{(0.019)}{0.042} hrsemp_i - \underset{(0.370)}{0.951} \ln(sales_i) + \underset{(0.360)}{0.992} \ln(employ_i), & (3) \\ n &= 43, SCR = 65.91 \end{aligned}$$

donde *hrsemp* son las horas anuales de formación para cada trabajador, *sales* son las ventas anuales (en dólares) y *employ* es el número de empleados en la empresa. Además, se reporta que $\widehat{Cov}(\hat{\beta}_{\ln(sales)}, \hat{\beta}_{\ln(employ)}) = -0.11$. Los errores estándar y la covarianza estimada de los coeficientes son robustos a heterocedasticidad.

- a. (2/5) Se modifica la anterior especificación y se reporta el ajuste *MCO*,

$$\begin{aligned} \ln(\widehat{scrap}_i) &= \underset{(4.57)}{11.74} - \underset{(0.019)}{0.042} hrsemp_i - \underset{(0.370)}{0.951} \ln\left(\frac{sales_i}{employ_i}\right) + \underset{(?)}{0.041} \ln(employ_i), & (4) \\ n &= 43, SCR = 65.91 \end{aligned}$$

Demuestre que existe una relación uno-a-uno entre los coeficientes en (3) y (4) (1/2 de 2.a). Utilizando esta relación, obtenga el error estándar del coeficiente de $\ln(employ)$ no reportado en (4) (1/2 de 2.a).

- b. (1/5) ¿Cómo variarían los coeficientes estimados si las ventas (*sales*) se reportan en miles de dólares en vez de en dólares? Calcule el valor numérico de la variación para cada uno de los cuatro parámetros estimados.
- c. (2/5) Controlando por la formación de los trabajadores (*hrsemp*) y por el ratio de ventas por empleado (*sales/employ*), ¿tienen las empresas más grandes (con más trabajadores) un ratio de defectuosos mayor? Establezca las hipótesis nula y alternativa, la regla de decisión, y realice el contraste (3/4 de 2.c). ¿Cómo contrastaría que el efecto de las ventas (*sales*) y el tamaño de la empresa (*employ*) sobre $\ln(scrap)$ son idénticos, pero con signo contrario? ((1/4 de 2.c) Valores críticos de la normal estándar Z : $Z_{0.005} = 2.58$, $Z_{0.01} = 2.33$, $Z_{0.025} = 1.96$, $Z_{0.05} = 1.64$, $Z_{0.1} = 1.28$, donde $\mathbb{P}(Z > Z_\alpha) = \alpha$).

3. (30%) Considere la estimación de la oferta de trabajo de mujeres casadas. La demanda de trabajo expresa el salario ofrecido en términos de las horas demandadas. Una vez que imponemos la

condición de equilibrio, las dos ecuaciones estructurales a estimar son

$$\begin{aligned} \text{hours} = & \beta_{10} + \beta_{11} \ln(\text{wage}) + \beta_{12} \text{educ} + \beta_{13} \text{age} + \beta_{14} \text{kidslt6} \\ & + \beta_{15} \text{nwifeinc} + \beta_{16} (\text{kidslt6} \times \text{nwifeinc}) + u_1 \end{aligned} \quad (5)$$

y

$$\ln(\text{wage}) = \beta_{20} + \beta_{21} \text{hours} + \beta_{22} \text{exper} + \beta_{23} \text{exper}^2 + u_2, \quad (6)$$

donde *age* es la edad de la mujer, *educ* sus años de educación, *kidslt6* el número de hijos menores de 6 años, *nwifeinc* son los ingresos del hogar excluyendo el salario de la mujer en miles de dólares (que incluye los ingresos del marido), y *exper* son los años de experiencia laboral. Sabemos que todas las variables de control (*educ*, *age*, *kidslt6*, *nwifeinc* y *exper*) son independientes de los dos términos de error (u_1 y u_2), los cuales tienen media cero y son independientes entre sí. Utilice las salidas de GRETL que se encuentran al final del examen para contestar a las preguntas. Valores críticos de la χ^2/q al 5% $q = 1, \dots, 5$: $\chi_{1,0.05}^2 = 3.84$, $\chi_{2,0.05}^2/2 = 3.00$, $\chi_{3,0.05}^2/3 = 2.60$, $\chi_{4,0.05}^2/4 = 2.37$, $\chi_{5,0.05}^2/5 = 2.21$.

- a. (2/5) ¿Qué variables instrumentales están disponibles para estimar los parámetros de la ecuación (5) (oferta de horas) por *MC2E*? Explique con todo detalle cómo contrastaría que estas variables instrumentales son relevantes: i) ¿cuál es el estadístico del contraste? ii) ¿cómo se computaría el mismo? y iii) ¿cuál es la regla de decisión? (1/2 de 3.a). Realice el contraste. (1/2 de 3.a).
- b. (2/5) ¿Cuál es la diferencia en media de horas ofertadas entre dos mujeres con idénticas características excepto que una tiene 2 hijos menores de 6 años y su salario es la única fuente de ingresos del hogar, mientras que la otra mujer no tiene ningún hijo menor de seis años pero dispone de unos ingresos en su hogar de 20 mil dólares adicionales a su salario? (1/2 de 3.b) . Proporcione un intervalo de confianza para esta diferencia y contraste que la diferencia es significativamente diferente de cero al 5% de significación (1/2 de 3.b) Valores críticos de la normal estándar Z : $Z_{0.005} = 2.58$, $Z_{0.01} = 2.33$, $Z_{0.025} = 1.96$, $Z_{0.05} = 1.64$, $Z_{0.1} = 1.28$, donde $\mathbb{P}(Z > Z_\alpha) = \alpha$.
- c. (1/5) Explique detalladamente cómo contrastaría que los instrumentos son exógenos: i) Establezca la hipótesis nula y alternativa, ii) explique cómo computar el estadístico del contraste y iii) establezca la regla de decisión (2/3 de 3.c). Realice el contraste al 5% de significación. (1/3 de 3.c).

Modelo 1: MCO, usando las observaciones 1–428

Variable dependiente: lwage

| | Coefficiente | Desv. Típica | Estadístico t | Valor p |
|------------------------|--------------|-----------------------|-----------------|---------|
| const | −0.449268 | 0.285534 | −1.5734 | 0.1164 |
| age | −0.00269880 | 0.00520903 | −0.5181 | 0.6047 |
| educ | 0.101004 | 0.0149790 | 6.7430 | 0.0000 |
| kidslt6 | 0.00268457 | 0.163924 | 0.0164 | 0.9869 |
| nwifeinc | 0.00615072 | 0.00362019 | 1.6990 | 0.0901 |
| nwifeincXkidslt6 | −0.00322245 | 0.00795267 | −0.4052 | 0.6855 |
| exper | 0.0414884 | 0.0132833 | 3.1233 | 0.0019 |
| expersq | −0.000747477 | 0.000402880 | −1.8553 | 0.0642 |
| Media de la vble. dep. | 1.190173 | D.T. de la vble. dep. | 0.723198 | |
| Suma de cuad. residuos | 186.7847 | D.T. de la regresión | 0.666877 | |
| R^2 | 0.163629 | R^2 corregido | 0.149689 | |
| $F(7, 420)$ | 11.73846 | Valor p (de F) | 1.14e−13 | |

Modelo 2: MCO, usando las observaciones 1–428

Variable dependiente: lwage

| | Coefficiente | Desv. Típica | Estadístico t | Valor p |
|------------------------|--------------|-----------------------|-----------------|---------|
| const | −0.434017 | 0.270691 | −1.6034 | 0.1096 |
| age | 0.00503829 | 0.00455412 | 1.1063 | 0.2692 |
| educ | 0.107789 | 0.0151864 | 7.0977 | 0.0000 |
| kidslt6 | −0.0241129 | 0.166747 | −0.1446 | 0.8851 |
| nwifeinc | 0.00315312 | 0.00356191 | 0.8852 | 0.3765 |
| nwifeincXkidslt6 | −0.00318766 | 0.00806527 | −0.3952 | 0.6929 |
| Media de la vble. dep. | 1.190173 | D.T. de la vble. dep. | 0.723198 | |
| Suma de cuad. residuos | 195.0670 | D.T. de la regresión | 0.679885 | |
| R^2 | 0.126543 | R^2 corregido | 0.116193 | |
| $F(5, 422)$ | 12.22748 | Valor p (de F) | 4.40e−11 | |

Modelo 3: MC2E, usando las observaciones 1–428

Variable dependiente: hours

Mediante Instrumentos: lwage

Instrumentos: const age educ kidslt6 nwifeinc nwifeincXkidslt6 exper expersq

| | Coefficiente | Desv. Típica | z | Valor p |
|------------------------|--------------|-----------------------|---------|-----------|
| const | 2232.35 | 578.202 | 3.8608 | 0.0001 |
| lwage | 1643.67 | 471.901 | 3.4831 | 0.0005 |
| age | -7.78173 | 9.40253 | -0.8276 | 0.4079 |
| educ | -184.111 | 59.2247 | -3.1087 | 0.0019 |
| kidslt6 | -240.214 | 333.275 | -0.7208 | 0.4711 |
| nwifeinc | -10.6106 | 7.26892 | -1.4597 | 0.1444 |
| nwifeincXkidlt6 | 2.43423 | 16.1806 | 0.1504 | 0.8804 |
| Media de la vble. dep. | 1302.930 | D.T. de la vble. dep. | | 776.2744 |
| Suma de cuad. residuos | 7.76e+08 | D.T. de la regresión | | 1358.086 |
| R^2 | 0.000534 | R^2 corregido | | -0.013710 |
| $F(6, 421)$ | 2.866484 | Valor p (de F) | | 0.009519 |

Matriz de covarianzas de los coeficientes en el modelo 3

| const | lwage | age | educ | kidslt6 | nwifeinc | nwifXkids | |
|------------|------------|---------|---------|------------|----------|-----------|-----------|
| 3.3432e+05 | 96652. | -4042.1 | -21294. | -26188. | -225.98 | 833.62 | const |
| | 2.2269e+05 | -1122.0 | -24004. | 5369.7 | -702.17 | 709.86 | lwage |
| | | 88.408 | 131.55 | 427.63 | -3.2554 | 1.2589 | age |
| | | | 3507.6 | -1403.5 | 16.223 | -67.471 | educ |
| | | | | 1.1107e+05 | 817.65 | -4502.4 | kidslt6 |
| | | | | | 52.837 | -48.120 | nwifeinc |
| | | | | | | 261.81 | nwifXkids |

Modelo 4: MCO, usando las observaciones 1–428

Variable dependiente: uhat

IMPORTANTE: uhat son los residuos correspondientes al modelo 3.

| | Coefficiente | Desv. Típica | Estadístico t | Valor p |
|------------------------|--------------|-----------------------|-----------------|-----------|
| const | 194.274 | 581.606 | 0.3340 | 0.7385 |
| age | -3.68530 | 10.6103 | -0.3473 | 0.7285 |
| educ | 0.190471 | 30.5108 | 0.0062 | 0.9950 |
| kidslt6 | 16.7347 | 333.899 | 0.0501 | 0.9601 |
| nwifeinc | 1.19491 | 7.37399 | 0.1620 | 0.8713 |
| nwifeincXkidslt6 | -1.50531 | 16.1989 | -0.0929 | 0.9260 |
| exper | -16.9634 | 27.0570 | -0.6270 | 0.5310 |
| expersq | 0.673158 | 0.820630 | 0.8203 | 0.4125 |
| Media de la vble. dep. | 6.44e-13 | D.T. de la vble. dep. | | 1348.511 |
| Suma de cuad. residuos | 7.75e+08 | D.T. de la regresión | | 1358.368 |
| R^2 | 0.001962 | R^2 corregido | | -0.014672 |
| $F(7, 420)$ | 0.117922 | Valor p (de F) | | 0.997131 |

Contraste sobre el modelo 4:

Hipótesis nula: los parámetros de regresión son cero para las variables exper y expersq

Estadístico del contraste: $F(2,420)=0.412728$, valor p 0.66211