

UNIVERSIDAD CARLOS III DE MADRID
ECONOMETRÍA I
Curso 2008/09
EXAMEN FINAL (Convocatoria ordinaria)

19 de Enero de 2009

Muy importante: Tenga en cuenta que:

1. Cada pregunta del cuestionario, salvo que se indique expresamente lo contrario, requiere un análisis completo de todas las salidas del problema al que se refiere.
Por ejemplo, para responder aquellas preguntas que se refieren a “estimaciones apropiadas”, o “dadas las estimaciones” o “dadas las condiciones del problema”, deben usarse los resultados basados en los estimadores consistentes y más eficientes de entre las distintas salidas.
2. Cada salida incluye todas las variables explicativas utilizadas en la estimación correspondiente.
3. Algunos resultados correspondientes a las salidas presentadas han podido ser omitidos.
4. La variable dependiente puede variar en cada salida presentada dentro del mismo problema.
5. Para simplificar, diremos que un modelo está “bien especificado” cuando el modelo sea lineal en las variables en que se condiciona (tal y como aparecen en el modelo) y el error sea independiente en media de dichas variables.
6. MCO y MC2E son las abreviaturas de mínimos cuadrados ordinarios y mínimos cuadrados en 2 etapas, respectivamente.
7. Se adjuntan tablas estadísticas al final de este documento.

ENUNCIADOS DE PROBLEMAS

PROBLEMA 1: Rendimiento de la educación

Estamos interesados en estimar una ecuación salarial con datos de individuos varones, utilizando el logaritmo neperiano del salario mensual, $lwage$, como variable dependiente. Las variables explicativas para cada individuo son los años de educación, $educ$, su edad, age , una variable binaria que indica su estado civil, $married$ (que toma el valor uno si el individuo está casado y cero en caso contrario), y su habilidad, $abil$.

$$lwage = \beta_0 + \beta_1 educ + \beta_2 age + \beta_3 married + \beta_4 abil + \varepsilon \quad (1.1)$$

donde $E(\varepsilon | educ, age, married, abil) = 0$. Esperaríamos que $\beta_4 > 0$.

Dado que $abil$ es inobservable, la ecuación que podemos estimar omite dicha variable, de manera que estimamos el siguiente modelo de determinación salarial:

$$lwage = \delta_0 + \delta_1 educ + \delta_2 age + \delta_3 married + u \quad (1.2)$$

Tenemos razones sólidas para creer que los años de educación, $educ$, están correlacionados con la habilidad omitida, $abil$. Por el contrario, podemos asumir que dicha variable omitida no está correlacionada con las otras dos variables explicativas observables, age y $married$.

Como posibles instrumentos para $educ$, disponemos de dos variables observables no incluidas en el modelo, que sabemos que no están correlacionadas con la habilidad:

- $urban$ (una variable binaria que toma el valor uno si el individuo reside en una ciudad de más de 50000 habitantes y cero en caso contrario);
- $feduc$ (años de educación del padre).

Nuestro objetivo es obtener estimadores consistentes de los parámetros de (1.1), $\beta_1, \beta_2, \beta_3$, a partir de la información observable.

Salida 1: estimaciones MCO utilizando las 663 observaciones 1–663

Variable dependiente: $lwage$

| Variable | Coefficiente | Desv. típica | Estadístico t | valor p |
|-----------|--------------|--------------|-----------------|---------|
| const | 5.0578 | 0.1831 | 27.62 | 0.0000 |
| $educ$ | 0.0597 | 0.0066 | 9.03 | 0.0000 |
| age | 0.0228 | 0.0048 | 4.71 | 0.0000 |
| $married$ | 0.2101 | 0.0494 | 4.25 | 0.0000 |

| | |
|-----------------------------------|---------|
| Suma de cuadrados de los residuos | 94.4723 |
| R^2 | 0.1601 |
| \bar{R}^2 corregido | 0.1563 |
| $F(3, 659)$ | 41.88 |

Salida 2: estimaciones MCO utilizando las 663 observaciones 1–663
Variable dependiente: *educ*

| Variable | Coeficiente | Desv. típica | Estadístico t | valor p |
|-----------------------------------|-------------|--------------|-----------------|---------|
| const | 12.4497 | 0.9632 | 12.92 | 0.0000 |
| <i>age</i> | 0.0383 | 0.0282 | 1.35 | 0.1761 |
| <i>married</i> | -0.4256 | 0.2890 | -1.47 | 0.1413 |
| <i>urban</i> | 0.4888 | 0.4161 | 1.17 | |
| Suma de cuadrados de los residuos | 3244.91 | | | |
| R^2 | | 0.0156 | | |
| $F(3, 659)$ | | 3.47 | | |

Salida 3: estimaciones MCO utilizando las 663 observaciones 1–663
Variable dependiente: *educ*

| Variable | Coeficiente | Desv. típica | Estadístico t | valor p |
|-----------------------------------|-------------|--------------|-----------------|---------|
| const | 9.1774 | 0.9140 | 10.04 | 0.0000 |
| <i>age</i> | 0.0565 | 0.0257 | 2.20 | 0.0281 |
| <i>married</i> | -0.3868 | 0.2623 | -1.47 | 0.1407 |
| <i>feduc</i> | 0.2907 | 0.0239 | 12.18 | 0.0000 |
| Suma de cuadrados de los residuos | 2674.57 | | | |
| R^2 | | 0.1886 | | |
| $F(3, 659)$ | | 51.06 | | |

Salida 4: estimaciones MC2E utilizando las 663 observaciones 1–663
Variable dependiente: *lwage*
Instrumentos: *urban*

| Variable | Coefficiente | Desv. típica | Estadístico t | valor p |
|----------------|--------------|--------------|-----------------|---------|
| <i>const</i> | -0.2134 | 2.2910 | -0.09 | 0.9258 |
| <i>educ</i> | 0.4700 | 0.1753 | 2.68 | 0.0073 |
| <i>age</i> | 0.0073 | 0.0142 | 0.51 | 0.6087 |
| <i>married</i> | 0.3964 | 0.1515 | 2.62 | 0.0089 |

Suma de cuadrados de los residuos 646.232
 $F(3, 659)$ 29.27

Contraste de Hausman –

Hipótesis nula: Los estimadores de MCO son consistentes

Estadístico de contraste asintótico: $\chi_1^2 = 40.42$

con valor p = 2.04972e-010

Salida 5: estimaciones MC2E utilizando las 663 observaciones 1–663

Variable dependiente: *lwage*

Instrumentos: *feduc*

| Variable | Coefficiente | Desv. típica | Estadístico t | valor p |
|----------------|--------------|--------------|-----------------|---------|
| <i>const</i> | 4.4176 | 0.2669 | 16.55 | 0.0000 |
| <i>educ</i> | 0.1095 | 0.0161 | 6.81 | 0.0000 |
| <i>age</i> | 0.0209 | 0.0051 | 4.13 | 0.0000 |
| <i>married</i> | 0.2327 | 0.0519 | 4.49 | 0.0000 |

Suma de cuadrados de los residuos 102.610
 $F(3, 659)$ 38.79

Contraste de Hausman –

Hipótesis nula: Los estimadores de MCO son consistentes

Estadístico de contraste asintótico: $\chi_1^2 = 13.11$

con valor p = 0.000293096

PROBLEMA 2: Determinantes del test SAT

La variable sat es la puntuación en el test SAT de aptitud escolar, $hsize$ es el tamaño de la promoción (medido en cientos de alumnos) a la que pertenece el alumno, $female$ es una variable binaria de sexo (que toma el valor 1 si el estudiante es una mujer y 0 en caso contrario), y $black$ es una variable binaria racial (que toma el valor 1 si el estudiante es de raza negra y 0 en caso contrario). Se propone el modelo siguiente para estimar los efectos de varios factores sobre los resultados del test SAT de aptitud escolar,

$$sat = \beta_0 + \beta_1 hsize + \beta_2 hsize^2 + \beta_3 female + \beta_4 black + \beta_5 femaleblack + u \quad (2.1)$$

donde $hsize^2$ es el cuadrado de la variable $hsize$ y la variable $femaleblack$ es el término de interacción $female \times black$.

Se considera también un modelo más general que incluye el efecto adicional de que el alumno sea deportista, mediante la variable $athlete$ (que toma el valor 1 si la observación corresponde a un estudiante deportista y 0 en caso contrario), así como la variable interacción $athleteblack = athlete \times black$.

Salida 1: estimaciones MCO utilizando las 4137 observaciones 1–4137

Variable dependiente: sat

| | Coefficiente | Desv. típica | estadístico t | valor p |
|---------------|--------------|--------------|-----------------|-----------|
| $const$ | 1028.1000 | 6.2902 | 163.44 | 0.0000 |
| $hsize$ | 19.2971 | 3.8323 | 5.03 | 0.0000 |
| $hsize^2$ | -2.1948 | 0.5272 | -4.16 | 0.0000 |
| $female$ | -45.0910 | 4.2911 | -10.51 | 0.0000 |
| $black$ | -169.8100 | 12.7131 | -13.36 | 0.0000 |
| $femaleblack$ | 62.3064 | 18.1542 | 3.43 | 0.0006 |

| | |
|--|-------------|
| Media de la var. dependiente | 1030.33 |
| D.T. de la variable dependiente | 139.401 |
| Suma de cuadrados de los residuos | 7.34791e+07 |
| Desviación típica de la regresión ($\hat{\sigma}$) | 133.369 |
| R^2 | 0.0858 |
| \bar{R}^2 corregido | 0.0847 |
| $F(5, 4131)$ | 77.52 |

Salida 2: estimaciones MCO utilizando las 4137 observaciones 1–4137

Variable dependiente: $usq1$ (residuos de la *Salida 1* al cuadrado)

| | Coefficiente | Desv. típica | estadístico t | valor p |
|--------------------|--------------|--------------|-----------------|---------|
| <i>const</i> | 19456.50 | 1195.80 | 16.27 | 0.0000 |
| <i>hsize</i> | 25.94 | 728.54 | 0.04 | 0.9716 |
| <i>hsize2</i> | -43.98 | 100.22 | -0.44 | 0.6608 |
| <i>female</i> | -3226.80 | 815.76 | -3.96 | 0.0001 |
| <i>black</i> | 7445.69 | 2416.85 | 3.08 | 0.0021 |
| <i>femaleblack</i> | -9217.30 | 3451.23 | -2.67 | 0.0076 |

Suma de cuadrados de los residuos 3.2907e+12
 R^2 0.0357
 \bar{R}^2 corregido 0.0346
 $F(5, 4131)$ 30.61

Salida 3: estimaciones MCO utilizando las 4137 observaciones 1-4137
Variable dependiente: *usq1* (residuos de la *Salida 1* al cuadrado)

| | Coefficiente | Desv. típica | estadístico t | valor p |
|--------------------|--------------|--------------|-----------------|---------|
| <i>const</i> | 2024.31 | 607.05 | | |
| <i>female</i> | -3892.49 | 907.94 | | |
| <i>black</i> | 26726.56 | 2689.55 | | |
| <i>femaleblack</i> | -11921.12 | 3838.87 | | |

R^2 0.0355
 $F(3, 4131)$ 50.78

Salida 4: estimaciones MCO utilizando las 4137 observaciones 1-4137
Variable dependiente: *sat*

| | Coefficiente | Desv. típica | estadístico t | valor p |
|---------------------|--------------|--------------|-----------------|---------|
| <i>const</i> | 1033.6200 | 6.2360 | 165.75 | 0.0000 |
| <i>hsize</i> | 18.0214 | 3.7787 | 4.77 | 0.0000 |
| <i>hsize2</i> | -1.8931 | 0.5202 | -3.64 | 0.0003 |
| <i>female</i> | -48.4500 | 4.2450 | -11.41 | 0.0000 |
| <i>black</i> | -130.2500 | 14.1122 | -9.23 | 0.0000 |
| <i>femaleblack</i> | 36.7119 | 18.3839 | 2.00 | 0.0459 |
| <i>athlete</i> | -97.5820 | 11.0091 | -8.86 | 0.0000 |
| <i>athleteblack</i> | -59.1710 | 25.2181 | -2.35 | 0.0190 |