Econometría
Universidad Carlos III de Madrid
Examen Extraordinario
25 de Junio de 2014

Instrucciones para la realización del examen:

Dispone de 2 horas y media para responder al examen
La evaluación consta de 10 preguntas referidas a 2 ejercicios cuya puntuación total son 10 puntos.
Utilice hojas de respuesta separadas para los dos ejercicios. Escriba su nombre claramente en todas ellas.
La puntuación total de cada una de los ejercicios se indica al comienzo de cada pregunta.
En caso que no se mencione el nivel de significación, utilice un 5%.
En caso de que se le pida contrastar una hipótesis, se evaluarán los siguientes puntos:
*Definir claramente las hipótesis nula y alternativa.
*Definir claramente el estadístico.
*Definir claramente el nivel de significación y el valor crítico. Se adjuntan tablas a continuación.
*Concluir e interpretar la regla de rechazo.

VALORES CRÍTICOS:

$\Pr(N(0,1) > 2,576) = 0,005$
$\Pr(N(0,1) > 2,326) = 0,01$
$\Pr(N(0,1) > 1,960) = 0,025$
$\Pr(N(0,1) > 1,645) = 0,05$
$\Pr(N(0,1) > 1,282) = 0,10$

$c : \Pr(\chi^2_q > c) = \alpha$	$q = 1$	$q = 2$	$q = 3$	$q = 4$	$q = 5$	$q = 6$
$\alpha = 0,01$	$6,63$	$9,21$	$11,34$	$13,28$	15.09	16.81
$\alpha = 0,05$	$3,84$	$5,99$	$7,81$	$9,49$	$11,07$	$12,59$
$\alpha = 0,10$	$2,71$	$4,61$	$6,25$	$7,78$	$9,24$	$10,65$

Recordamos que una t de Student con n grados de libertad se comporta como un $N(0,1)$ para n razonablemente grande ($n > 30$). Por otro lado, una F de $Fisher$ con q grados de libertad en el numerador y n grados de libertad en el denominador se comporta como una $\chi^2_{(q)}/q$ para n grande.

1. [7 puntos/sobre 10] En el estudio de los factores que determinan el Índice de Masa Corporal (Body Mass Index, BMI), definido como el peso (en kilogram0s) dividido pore el cuadrado de la altura (en metros). Para una muestra aleatoria se ha estimado el siguiente modelo de regresión (Modelo 1, Salida 1),

$$\widehat{BMI} = \widehat{\beta}_0 + \widehat{\beta}_1 * drinks + \widehat{\beta}_2 * drinks^2 + \widehat{\beta}_3 * female \qquad \text{(Modelo 1)}$$

donde

$$\begin{aligned} drinks &= \text{número de días durante el último año} \\ & \text{en los que el individual ha bebido 5 o más vasos de alcohol.} \\ drinks^2 &= \text{cuadrado de } drinks. \\ female &= \text{es una variable binaria que toma el valor uno} \\ & \text{para mujeres y cero para el resto.} \end{aligned}$$

a) Usando el Modelo 1, ¿Cuál es el impacto marginal de *drinks* sobre el BMI esperado? ¿Es ese efecto constante? ¿Para qué valores de *drinks* es ese efecto positivo (negativo)? **[1 punto/sobre 10]**

Respuesta:

$$\text{Efecto marginal} = \frac{\partial B\hat{M}I}{\partial drinks} = \hat{\beta}_1 + 2 * \hat{\beta}_2 drinks.$$

Como el efecto marginal depende de *drinks* podemos decir que el impacto no es constante (ya que $\hat{\beta}_2$ es significativo al 5% con el correspondiente test de la t).

El valor a partir del cual el impacto marginal de *drinks* cambia de signo es $drinks^* = -\frac{\hat{\beta}_1}{2*\hat{\beta}_2} =$ 47.5.

Como $\hat{\beta}_1 < 0$ y $\left|\hat{\beta}_1\right| > \left|\hat{\beta}_2\right|$ para valores de *drinks* por debajo de $drinks^*$ un aumento del consumo de alcohol se asocia a una reducción de BMI. Sin embargo, un incremento del consumo para valores de *drinks* por encima de $drinks^*$ se asocia a un incremento de BMI.

b) Usando el Modelo 1, ¿cuál es la diferencia predicha en BMI entre un hombre con $drinks = 2$ y una mujer con $drinks = 6$? **[1 punto/sobre 10]**

Respuesta:

$$\begin{aligned} E[BMI|female &= 0, drinks = 2] - E[BMI|female = 1, drinks = 6] \\ &= \widehat{\beta_1} * (2-6) + \widehat{\beta}_2 * (4-36) - \widehat{\beta}_3 = -\widehat{\beta_1} * 4 - \widehat{\beta}_2 * 32 - \widehat{\beta}_3 \\ &= 0.0095 * 4 - 32 * 0.0001 + 1.1418 = 1.1766 \end{aligned}$$

Alternativamente, se ha estimado el siguiente modelo (Modelo 2, Salida 2)

$$\begin{aligned} \widehat{BMI} &= \widehat{\beta}_0 + \widehat{\beta}_1 * drinks + \widehat{\beta}_2 * drinks^2 \qquad \text{(Modelo 2)} \\ &\quad + \widehat{\beta}_3 * female + \widehat{\beta}_4 * drinks * female + \widehat{\beta}_5 * drinks^2 * female \end{aligned}$$

c) Usando el Modelo 2, ¿cuál es la diferencia predicha en el BMI medio entre un hombre con $drinks = 2$ y una mujer con $drinks = 6$? **[1 punto/sobre 10]**

$$\begin{aligned} & E[BMI|female = 0, drinks = 2] - E[BMI|female = 1, drinks = 6] \\ = \ & \widehat{\beta_1} * (2-6) + \widehat{\beta}_2 * (4-36) - \widehat{\beta}_3 - \hat{\beta}_4 * 6 - \hat{\beta}_5 * 36 = -\widehat{\beta_1} * 4 - \widehat{\beta}_2 * 32 - \widehat{\beta}_3 - \hat{\beta}_4 * 6 - \hat{\beta}_5 * 36 \\ = \ & -0.0424 * 4 + 32 * 0.0004 + 0.8753 + 6 * 0.1639 - 36 * 0.0016 = 1.6443 \end{aligned}$$

d) Usando el Modelo 2, ¿cuál es el impacto sobre BMI de un cambio marginal en *drinks* para un hombre? ¿Cuál es el impacto sobre BMI de un cambio marginal en *drinks* para una mujer? [**1 punto/sobre 10**]

$$\begin{aligned}
\text{Efecto marginal} &= \frac{\partial B\hat{M}I}{\partial drinks} = (\hat{\beta}_1 + \hat{\beta}_4 * female) + 2 * drinks * (\hat{\beta}_2 + \hat{\beta}_5 * female) \\
\text{Efecto m. p. hombres} &= \hat{\beta}_1 + 2 * \hat{\beta}_2 drinks = 0.0424 - 2 * 0.0004 * drinks = 0.0424 - 0.0008 * dri \\
\text{Efecto m. p. mujeres} &= (\hat{\beta}_1 + \hat{\beta}_4) + 2 * drinks * (\hat{\beta}_2 + \hat{\beta}_5) = -0.1215 + 0.0024 * drinks
\end{aligned}$$

e) Usando el Modelo 2 como referencia (modelo no restringido), explique y contraste que el efecto sobre BMI de un cambio marginal en *drinks* es lineal para los hombres. [**1 punto/sobre 10**]

Hipótesis a contrastar:

$$\begin{aligned}
H_0 &: \beta_2 = 0 \\
H_1 &: \beta_2 \neq 0
\end{aligned}$$

Construimos el estadístico t

$$t = \frac{\hat{\beta}_2 - 0}{se(\hat{\beta}_2)} = \frac{-0.0004}{0.00006} = -6.6.$$

como $|t| > 1.96$, podemos rechazar H_0 al nivel de significación del 5%. Esto quiere decir que se rechaza que el efecto marginal de *drinks* sea constante y por tanto confirmamos que es lineal para los hombres.

f) Usando el Modelo 2 como referencia (modelo no restringido), explique y contraste que el efecto sobre BMI de un cambio marginal en *drinks* es lineal para las mujeres. [**1 punto/sobre 10**]

Hipótesis a contrastar:

$$\begin{aligned}
H_0 &: \beta_2 + \beta_5 = 0 \\
H_1 &: \beta_2 + \beta_5 \neq 0
\end{aligned}$$

Como no se proporciona la matriz de varianzas-covarianzas de los estimadores, se necesita estimar los modelos restringidos y no restringidos y realizar un test de la F bajo el supuesto de homocedasticidad. El modelo restringido bajo H_0 se puede expresar como

$$\begin{aligned}
\widehat{BMI} &= \widehat{\beta}_0 + \widehat{\beta}_1 * drinks + \widehat{\beta}_2 * drinks^2 + \widehat{\beta}_3 * female + \widehat{\beta}_4 * drinks * female \\
&\quad - \hat{\beta}_2 * drinks^2 * female \\
&= \widehat{\beta}_0 + \widehat{\beta}_1 * drinks + \widehat{\beta}_2 * drinks^2 * (1 - female) + \widehat{\beta}_3 * female + \widehat{\beta}_4 * drinks * female \\
&= \widehat{\beta}_0 + \widehat{\beta}_1 * drinks + \widehat{\beta}_2 * drinks^2 * male + \widehat{\beta}_3 * female + \widehat{\beta}_4 * drinks * female
\end{aligned}$$

Del listado de salidas proporcionado, se puede observar que el modelo restringido se corresponde con el modelo 5. Como los modelos restringido y no restringido (modelo 2, salida 2) tienen la misma variable dependiente, se puede construir nuestro estadístico F usando los R^2:

$$F = \frac{R^2_{unrestricted} - R^2_{restricted}}{(1 - R^2_{unrestricted})} * \frac{n}{q} = \frac{(0.0151 - 0.0143)}{(1 - 0.0151)} * \frac{233239}{1} = 189.45.$$

Como el estadístico F sigue una distribución χ^2_1 asintóticamente bajo la nula, con valor crítico 3.8 al nivel de significación del 5%, se rechaza H_0, lo que implica que se rechaza que el efecto marginal de *drinks* sea constante y por tanto confirmamos que es lineal para las mujeres.

g) Usando el Modelo 2 como referencia (modelo no restringido), explique y contraste que el el efecto sobre BMI de un cambio marginal en *drinks* es lineal para un individuo, cualquiera que sea su género. [**1 punto/sobre 10**]

Hipótesis a contrastar:

$$\begin{aligned} H_0 &: \beta_2 = \beta_5 = 0 \\ H_1 &: H_0 \text{ es falsa} \end{aligned}$$

Modelo restringido

$$\widehat{BMI} = \widehat{\beta}_0 + \widehat{\beta}_1 * drinks + \widehat{\beta}_3 * female + \widehat{\beta}_4 * drinks * female,$$

que se corresponde con la salida 4. Como el modelo restringido y el no restringido (Modelo 2, Salida 2) tienen la misma variable dependiente, se puede construir un contraste de la F usando los R^2:

$$F = \frac{R^2_{unrestricted} - R^2_{restricted}}{(1 - R^2_{unrestricted})} * \frac{n}{q} = \frac{(0.0151 - 0.0017)}{(1 - 0.0151)} * \frac{233239}{2} = 1586.7.$$

Como el estadístico F sigue una distribución asintótica $\chi^2_2/2$ bajo la nula, con valor crítico 5.99/2≈ 3 al nivel de significación del 5%. Por tanto se rechaza H_0 lo que significa que el impacto de *drinks* no es constante en general y es lineal para al menos un tipo de individuos.

SALIDA 1: OLS, using observations 1–233239
Dependent variable: bmi

	Coeff	S.E	t	p-value
const	26.8065	0.0182106	1472.0235	0.0000
drinks	–0.00953349	0.00484486	-1.9678	0.0491
drinks2	0.000102875	4.98842e–005	2.0623	0.0392
female	–1.14183	0.020171		

Mean Dependent Variable.	26.17626
Sum of the Squared Residuals	5196636
R^2	0.014139

SALIDA 2: OLS, using observations 1–233239
Dependent variable: bmi

	Coeff	S.E	t	p-value
const	26.6876	0.0197415	1351.8542	0.0000
drink	0.0424549	0.00586037	7.2444	0.0000
drink2	−0.000419308	6.03490e–005	−6.9481	0.0000
female	−0.875380	0.0264042	–33.1531	0.0000
drinks*female	–0.163936	0.0104052	–15.7552	0.0000
drinks2*female	0.00165051	0.000107120		

Mean Dependent Variable.	26.17626
Sum of the Squared Residuals	5191109
R^2	0.015188

SALIDA 3: OLS, using observations 1–233239

Dependent variable: bmi

	Coeff	S.E	t	p-value
const	26.0776	0.0129662	2011.2017	0.0000
drinks	0.0543016	0.00474404	11.4463	0.0000
drinks2	–0.000511294	4.90232e–005		

Mean Dependent Variable.	26.17626
Sum of the Squared Residuals	5268033
R^2	0.000594

SALIDA 4: OLS, using observations 1–233239

Dependent variable: bmi

	Coeff	S.E.	t	p-value
const	26.1650	0.0102729	2546.9867	0.0000
drinks	0.0214939	0.00139848	15.3695	0.0000
drinks*female	–0.0445572	0.00232373	–19.1749	0.0000
female	–1.113876	0.0205		

Mean Dependent Variable.	26.17626
Sum of the Squared Residuals	5262194
R^2	0.001702

SALIDA 5: OLS, using observations 1–233239

Dependent variable: bmi

	Coeff	S.E.	t	p-value
const	26.6876	0.0197496	1351.2966	0.0000
drink	0.0424549	0.00586279	7.2414	0.0000
female	−1.02614	0.0240875	–42.6007	0.0000
drinks*female	–0.0473893	0.00617347	–7.6763	0.0000
drinks2*male	–0.000419308	6.03739e–005		

NOTA: *male* es una variable binaria que toma el valor uno cuando una persona es un hombre y cero para el resto.

2. **[3 puntos/sobre 10]** Un modelo de ecuaciones simultáneas para estudiar si la apertura de la economía (*open*) lleva a menores tasas de inflación (*inf*) es

$$\begin{aligned} inf &= \delta_{10} + \gamma_{12} open + \delta_{11} \log(pcinc) + u_1 \\ open &= \delta_{20} + \gamma_{21} inf + \delta_{21} \log(pcinc) + \delta_{22} \log(land) + u_2. \end{aligned}$$

Se asume que *pcinc* (renta per cápita) y *land* (tierra de cultivo) son exógenos en todo el ejercicio.

Se han obtenido las siguientes estimaciones por MCO y MC2E.

Salida 1: estimaciones MCO utilizando las 114 observaciones 1–114
Variable dependiente: inf

Variable	Coeficiente	Desv. típica	Estadístico t	valor p
const	25,1040	15,2052	1,6510	0,1016
open	-0,215070	0,0946289	-2,2728	0,0250
lpcinc	0,0175673	1,97527	0,0089	0,9929

Media de la var. dependiente	17,2640
D.T. de la variable dependiente	23,9973
Suma de cuadrados de los residuos	62127,5
Desviación típica de los residuos ($\hat{\sigma}$)	23,6581
R^2	0,0452708
$\bar{R}^2$ corregido	0,0280685
$F(2, 111)$	2,63167
valor p para $F()$	0,0764453

Salida 2: estimaciones MCO utilizando las 114 observaciones 1–114
Variable dependiente: open

Variable	Coeficiente	Desv. típica	Estadístico t	valor p
const	116,226	15,8808	7,3187	0,0000
inf	-0,0680353	0,0715556	-0,9508	0,3438
lpcinc	0,559501	1,49395	0,3745	0,7087
lland	-7,3933	0,834814	-8,8563	0,0000

Media de la var. dependiente	37,0789
D.T. de la variable dependiente	23,7535
Suma de cuadrados de los residuos	34865,3
Desviación típica de los residuos ($\hat{\sigma}$)	17,8033
R^2	0,453162
$\bar{R}^2$ corregido	0,438249
$F(3, 110)$	30,3855
valor p para $F()$	¡ 0,00001

Salida 3: estimaciones MCO utilizando las 114 observaciones 1–114
Variable dependiente: inf

Variable	Coeficiente	Desv. típica	Estadístico t	valor p
const	-12,615	21,0313	-0,5998	0,5498
lpcinc	0,191394	1,98158	0,0966	0,9232
lland	2,55380	1,08049	2,3635	0,0198

Media de la var. dependiente	17,2640
D.T. de la variable dependiente	23,9973
Suma de cuadrados de los residuos	61903,2
Desviación típica de los residuos ($\hat{\sigma}$)	23,6154
R^2	0,0487174
$\bar{R}^2$ corregido	0,0315772
$F(2, 111)$	2,84229
valor p para $F()$	0,0625432

Salida 4: estimaciones MCO utilizando las 114 observaciones 1–114
Variable dependiente: open

Variable	Coeficiente	Desv. típica	Estadístico t	valor p
const	117,085	15,8483	7,3878	0,0000
lpcinc	0,546479	1,49324	0,3660	0,7151
lland	-7,5671	0,814216	-9,2937	0,0000

Media de la var. dependiente	37,0789
D.T. de la variable dependiente	23,7535
Suma de cuadrados de los residuos	35151,8
Desviación típica de los residuos ($\hat{\sigma}$)	17,7956
R^2	0,448668
$\bar{R}^2$ corregido	0,438734
$F(2, 111)$	45,1654
valor p para $F()$	¡ 0,00001

Salida 5: estimaciones MC2E utilizando las 114 observaciones 1–114
Variable dependiente: inf
Instrumentos: lland

Variable	Coeficiente	Desv. típica	Estadístico t	valor p
const	26,8993	15,4012	1,7466	0,0807
open	-0,337487	0,144121	-2,3417	0,0192
lpcinc	0,375823	2,01508	0,1865	0,8520

Media de la var. dependiente	17,2640
D.T. de la variable dependiente	23,9973
Suma de cuadrados de los residuos	63064,2
Desviación típica de los residuos ($\hat{\sigma}$)	23,8358
$F(2, 111)$	2,62498
valor p para $F()$	0,0769352

Contraste de Hausman –
Hipótesis nula: Los estimadores de MCO son consistentes
Estadístico de contraste asintótico: $\chi^2_1 = 1{,}35333$
con valor p = 0,244697

a) Discuta la posible identificación de cada ecuación del sistema, la debilidad de los instrumentos disponibles y lleve a cabo los correspondientes contrastes de hipótesis cuando sea posible. [**1 punto/sobre 10**]

Respuesta: La segunda ecuación no está identificada por que no hay instrumentos para el regresor endógeno *inf*. La primera ecuación puede estar (exactamente) identificada usando $\log(land)$ como instrumento para *open*. Mirando en la Salida 4 (forma reducida para *open*) se comprueba que el instrumento $\log(land)$ es significativo con $t = -9,2937$ y el correspondiente estadístico $F = t^2 = -9,2937 \approx 86.4$ indica que el instrumento no es débil.

b) Explique cómo realizaría un contraste de la exogeneidad de los instrumentos usados en la estimación en dos etapas de una ecuación y si es posible llevarlo a cabo para las ecuaciones del sistema planteado. [**1 punto/sobre 10**]

Respuesta: El contraste de exogeneidad se realizaría usando los residuos de cada ecuación estructural después de la estimación por MC2E, para ajustar una regresión con ellos con variable dependiente sobre todas las variables exógenas disponibles, y comprobando la significación conjunta de (sólo) los instrumentos omitidos en esa ecuación estructural. Este procedimiento sólo se puede llevar a cabo en casos de sobreidentificación: más instrumentos disponibles que regresores endógenos en la ecuación.

En el sistema planteado este no es el caso y no se puede comprobar la exogeneidad de $\log(land)$ en la primera ecuación porque está exactamente identificada y tampoco en la segunda porque no hay instrumentos disponibles.

c) Contraste si el efecto de *open* sobre inf es menor que $-0,2$. Si *open* no fuese un determinante de inf, (pero inf sí lo fuese de *open*), explique las propiedades de los estimadores de la Salida 1. [**1 punto/sobre 10**]

Respuesta: El test requerido es

$$\begin{aligned} H_0 &: \beta_{open} = -0.2 \\ H_1 &: \beta_{open} < -0.2 \end{aligned}$$

y se realiza con un test de la t

$$t = \frac{\hat{\beta}_{open} - (-0.2)}{se\left(\hat{\beta}_{open}\right)} = \frac{-0.3375 + 0.2}{0.1441} = -0.9542$$

usando la Salida 5 (MC2E), que no es significativo comparado con el valor crítico de un contraste bilateral al 5%, -1.65, por lo que no podemos confirmar que el efecto es menor que -0.2.

En ese caso, $\gamma_{12} = 0$ y $\gamma_{21} \neq 0$, *open* seguiría siguiendo endógena en la primera ecuación, ya que sustituyendo la primera ecuación en la segunda se obtiene que

$$\begin{aligned} open &= \delta_{20} + \gamma_{21}\left\{\delta_{10} + \gamma_{12} open + \delta_{11}\log(pcinc) + u_1\right\} + \delta_{21}\log(pcinc) + \delta_{22}\log(land) + u_2 \\ &= \delta_{20} + \gamma_{21}\left\{\delta_{10} + \delta_{11}\log(pcinc) + u_1\right\} + \delta_{21}\log(pcinc) + \delta_{22}\log(land) + u_2 \\ &= \delta_{20} + \gamma_{21}\delta_{10} + \left\{\gamma_{21}\delta_{11} + \delta_{21}\right\}\log(pcinc) + \delta_{22}\log(land) + u_2 + \gamma_{21}u_1 \end{aligned}$$

y se comprueba que *open* sigue estando correlada con u_1 porque $\gamma_{21} \neq 0$ y por tanto es endógena en la primera ecuación y los correspondientes estimadores MCO serán inconsistentes.