

**UNIVERSIDAD DE LA REPÚBLICA**  
**FACULTAD DE CIENCIAS ECONÓMICAS Y DE ADMINISTRACIÓN**

**1ª REVISIÓN DE ECONOMETRÍA II**

**14 de mayo de 2014**

**EJERCICIO 1 – (25 puntos)**

Utilizando los datos de la Encuesta Continua de Hogares 2008 del INE, Uruguay se estima el efecto de la educación del padre en la educación de los hijos. Para ello se seleccionó una muestra de adolescentes de entre 16 y 20 años que convivían con su madre y con su padre en el momento de la encuesta.

Las variables utilizadas fueron:

**Aniosed:** los años de educación aprobados por el adolescente

**Edad:** la edad del adolescente

**AniosedPadre:** los años de educación aprobados por el padre

**EdadPadre:** la edad del padre

**EdadPadre2:** la edad del padre al cuadrado

**AniosedMadre:** los años de educación aprobados por la madre

Se realizaron un conjunto de estimaciones que se informan a continuación:

**Estimación 1: Estimación MCO de Aniosed sobre Edad, AniosedPadre, EdadPadre y EdadPadre2.**

```
. reg aniosed aniosedpadre edad edadpadre edadpadre2
```

Source	SS	df	MS	Number of obs =	5410
Model	30783.8432	4	7695.9608	F( 4, 5405) =	859.27
Residual	48409.1328	5405	8.95636129	Prob > F =	0.0000
				R-squared =	0.3887
				Adj R-squared =	0.3883
Total	79192.976	5409	14.6409643	Root MSE =	2.9927

aniosed	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
aniosedpadre	.59958	.010793	55.55	0.000	.5784277 .6207449
edad	.04082	.008853	4.61	0.000	.0234688 .0581812
edadpadre	.23145	.044578	5.19	0.000	.1440605 .3188437
edadpadre2	-.00236	.000434	-5.44	0.000	-.0032182 -.0015131
_cons	-3.30307	1.08061	-3.06	0.002	-5.421507 -1.184646

## Estimación 2: Estimación VI de Aniosed sobre Edad, AniosedPadre, EdadPadre y EdadPadre2 utilizando AniosedMadre como instrumento de AniosedPadre

```
. ivreg aniosed edad edadpadre edadpadre2 (aniosedpadre = aniosedmadre)
```

Instrumental variables (2SLS) regression

Source	SS	df	MS	Number of obs =	5410
Model	12256.4693	4	3064.11732	F( 4, 5405) =	363.82
Residual	66936.5067	5405	12.3841826	Prob > F	= 0.0000
				R-squared	= 0.1548
				Adj R-squared	= 0.1541
Total	79192.976	5409	14.6409643	Root MSE	= 3.5191

aniosed	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
aniosedpadre	1.090474	.0314594	34.66	0.000	1.028801 1.152147
edad	-.0002746	.010686	-0.03	0.980	-.0212233 .0206742
edadpadre	-.0451697	.0548719	-0.82	0.410	-.1527407 .0624012
edadpadre2	.0006075	.0005403	1.12	0.261	-.0004517 .0016666
_cons	.321502	1.288335	0.25	0.803	-2.204154 2.847158

Instrumented: aniosedpadre

Instruments: edad edadpadre edadpadre2 aniosedmadre

## Estimación 3: Estimación MCO de Aniosedpadre sobre Edad, Aniosedmadre, EdadPadre y EdadPadre2

```
reg aniosedpadre aniosedmadre edad edadpadre edadpadre2
```

Source	SS	df	MS	Number of obs =	5410
Model	15563.2493	4	3890.81232	F( 4, 5405) =	327.57
Residual	64198.6701	5405	11.8776448	Prob > F	= 0.0000
				R-squared	= 0.1951
				Adj R-squared	= 0.1945
Total	79761.9194	5409	14.7461489	Root MSE	= 3.4464

aniosedpadre	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
aniosedmadre	.7738269	.022475	34.43	0.000	.7297669 .8178869
edad	.0663299	.010084	6.58	0.000	.0465612 .0860985
edadpadre	-.2695273	.0362749	-7.43	0.000	-.3406407 -.1984139
edadpadre2	-.0005343	.0000819	-6.53	0.000	-.0006948 -.0003738
_cons	4.831708	.6492624	7.44	0.000	3.558892 6.104524

## Contraste 1: Contraste de Hausman sobre la endogeneidad de la variable **aniosedpadre**

. hausman IVPADRE MCOPADRE

	---- Coefficients ----			
	(b)	(B)	(b-B)	sqrt(diag(V_b-V_B))
	IVPADRE	MCOPADRE	Difference	S.E.
aniosedpadre	1.090474	.5995863	.4908877	.0295501
edad	-.0002746	.040825	-.0410996	.0059839
edadpadre	-.0451697	.2314521	-.2766219	.0319952
edadpadre2	.0006075	-.0023656	.0029731	.0003206

b = consistent under Ho and Ha; obtained from ivreg  
 B = inconsistent under Ha, efficient under Ho; obtained from regress

Valor del estadístico de hausman= 16.74

Los p-valores correspondientes a 16.74 para distintas distribuciones

	p-valor
Chi_cuadrado con 1 grado de libertad	0.000042870
F-Fisher con 1 grado de libertad en el numerador y 5408 en el denominador	0.00004442
Chi_cuadrado con 4 grados de libertad	0.00217117
F-Fisher con 4 grado de libertad en el numerador y 5408 en el denominador	0.0000004
Chi-cuadrado con 10 grados de libertad	0.08031876

### Se pide:

1. Interprete los resultados de la estimación MCO (Estimación 1)
2. Explique en qué consiste el estimador de Variables Instrumentales (VI) y en qué circunstancias es preferible al estimador MCO.
3. Señale brevemente cómo se obtiene el estimador de Mínimos Cuadrados en 2 Etapas (MC2E). Indique la relación entre este estimador y el estimador por Variables Instrumentales.
4. Indique cómo cambian los resultados cuando se estima el modelo por MC2E (Estimación 2) respecto a las estimaciones MCO.
5. ¿Qué características tiene que poseer una variable para ser un buen instrumento?  
 ¿Considera que la educación de la madre es un buen instrumento para la educación del padre en un modelo que pretende estimar el efecto de la educación del padre en la educación de los adolescentes? Justifique su opinión utilizando tanto argumentos teóricos como la información que dispone a partir de las estimaciones y contrastes realizados. Si existen aspectos que le hacen dudar respecto a la validez del instrumento, señálelos.
6. Analice la información correspondiente al contraste de Hausman. Indique la hipótesis nula y la alternativa, la forma del estadístico de contraste y su distribución asintótica. Concluya indicando el criterio de decisión y el p-valor de la prueba en el caso analizado. ¿Cuál es la utilidad de dicho contraste? ¿Qué implicación tiene la conclusión a la que arribó en el contraste de Hausman en el caso bajo estudio?

## EJERCICIO 2 – (25 puntos)

Una compañía financiera lleva adelante una campaña de marketing para ofrecer un nuevo producto, enviando un correo promocional a cada uno de sus clientes. Finalizado el período de la promoción, se desea analizar la respuesta de los clientes, utilizando una muestra de 925 casos. En primer lugar interesa analizar los factores que explican si los clientes invierten o no (variable binaria *invierte* que toma el valor 1 en caso de que invierta y el valor 0 en caso contrario).

### Parte I: Modelos MPL y Logit

De campañas anteriores, se sabe que el sexo y la edad son dos variables explicativas a tener en cuenta, y que la respuesta también parece estar asociada a si el cliente ya es un inversor activo en la firma (es decir que tiene inversiones en otros productos). El sexo y la condición de activo se recogen en las variables binarias *varon* y *activo*. Las variables *edad* y *edad\_2* recogen la edad en años cumplidos y su cuadrado, respectivamente. Los estadísticos descriptivos de esas variables se presentan a continuación:

Variable	Obs	Mean	Std. Dev.	Min	Max
varon	925	.7254054	.4465512	0	1
activo	925	.1881081	.3910099	0	1
edad	925	50.68108	13.39844	11	93
invierte	925	.5081081	.5002047	0	1

Como primera aproximación se estima un Modelo de Probabilidad Lineal (por Mínimos Cuadrados Ponderados), y luego se estima un modelo Logit. Las salidas se presentan a continuación:

#### Modelo de Probabilidad Lineal (primera etapa: MCO)

Source	SS	df	MS	Number of obs =	925
Model	18.8515333	4	4.71288332	F( 4, 920) =	20.42
Residual	212.337656	920	.2308018	Prob > F =	0.0000
Total	231.189189	924	.25020475	R-squared =	0.0815
				Adj R-squared =	0.0775
				Root MSE =	.48042

invierte	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
varon	.224002	.0358086	6.26	0.000	.1537259 .294278
activo	.2082684	.0406694	5.12	0.000	.1284528 .2880839
edad	.0154942	.0078609	1.97	0.049	.0000669 .0309216
edad_2	-.0001521	.0000751	-2.03	0.043	-.0002994 -4.77e-06
_cons	-.0608878	.1959057	-0.31	0.756	-.4453617 .3235862

### Modelo de Probabilidad Lineal (segunda etapa: MCP)

Source	SS	df	MS	Number of obs = 925		
				F( 4, 920) = 24.57		
Model	99.1914851	4	24.7978713	Prob > F = 0.0000		
Residual	928.667992	920	1.00942173	R-squared = 0.0965		
				Adj R-squared = 0.0926		
Total	1027.85948	924	1.11240203	Root MSE = 1.0047		

  

invierte_sd	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
varon_sd	.2253788	.03695	6.10	0.000	.1528628	.2978948
activo_sd	.1998861	.0439127	4.55	0.000	.1137055	.2860667
edad_sd	.0162957	.0079604	2.05	0.041	.000673	.0319184
edad_2_sd	-.0001594	.0000767	-2.08	0.038	-.0003098	-8.95e-06
_cons	-.1683939	.418044	-0.40	0.687	-.9888244	.6520366

### Modelo LOGIT

```
Iteration 0: log likelihood = -641.03952
Iteration 1: log likelihood = -601.93102
Iteration 2: log likelihood = -601.86236
Iteration 3: log likelihood = -601.86236
```

Logistic regression	Number of obs	=	925
	LR chi2(4)	=	78.35
	Prob > chi2	=	0.0000
Log likelihood = -601.86236	Pseudo R2	=	0.0611

invierte	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
varon	.9536943	.1581827	6.03	0.000	.643662	1.263727
activo	.9137479	.1847788	4.95	0.000	.551588	1.275908
edad	.0699453	.0356054	1.96	0.049	.0001599	.1397306
edad_2	-.0006869	.000341	-2.01	0.044	-.0013552	-.0000186
_cons	-2.488358	.8899918	-2.80	0.005	-4.232709	-.7440058

```
Marginal effects after logit
      y = Pr(invierte) (predict)
      = .50816352
```

variable	dy/dx	Std. Err.	z	P> z	[	95% C.I.	]	X
varon*	.2321776	.03637	6.38	0.000	.160887	.303468		.725405
activo*	.2192487	.04097	5.35	0.000	.138952	.299545		.188108
edad	.0174817	.0089	1.96	0.049	.000039	.034924		50.6811
edad_2	-.0001717	.00009	-2.01	0.044	-.000339	-4.7e-06		2747.9

(\*)  $dy/dx$  is for discrete change of dummy variable from 0 to 1

**Se pide:**

1. Explique por qué el estimador MCO es ineficiente en el caso de variable dependiente binaria. Indique brevemente los pasos que componen la estimación de modelos binarios en dos etapas por Mínimos Cuadrados Ponderados (MCP).
2. ¿Cuáles son las principales ventajas y desventajas de la estimación del modelo Logit en comparación con el MPL?
3. Escriba la verosimilitud de la muestra para el modelo Logit y obtenga las condiciones de primer orden que dan lugar al estimador Máximo Verosímil.
4. Obtenga el efecto parcial asociado a la condición de activo en el modelo Logit, interprete el resultado y compárelo con el efecto parcial obtenido por MPL. Realice los cálculos en los valores promedio de los restantes regresores. En todos los casos presente las fórmulas de cálculo utilizadas.
5. Obtenga el efecto parcial de un año de edad adicional para una mujer de 30 años que no es cliente activa, e interprete el resultado. Presente la fórmula de cálculo utilizada.

**Parte II: Mínimos Cuadrados no lineales**

Considere ahora la posibilidad de estimar el mismo modelo pero utilizando el estimador de Mínimos Cuadrados No Lineales. La ecuación a estimar sería la siguiente:

$$y = \Lambda(x'\beta) + u$$

Donde  $\Lambda(\ )$  representa la distribución de probabilidad acumulada de una variable aleatoria logística estándar.

**Se pide:**

1. ¿Cómo se define el estimador de Mínimos Cuadrados no Lineales?
2. Obtenga las condiciones de primer orden del estimador MCNL en el modelo anterior.
3. Compare dichas condiciones con las que surgen en la estimación por Máxima Verosimilitud del modelo Logit.