

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

**1a. REVISIÓN DE ECONOMETRÍA II**  
**14 de mayo de 2015**

**Ejercicio (20 puntos)**

Se desea estudiar el efecto de estar sindicalizado en los salarios de los trabajadores. Para realizar dicho estudio se cuenta con una base con información sobre 545 trabajadores a través de 8 años, desde 1990 hasta 1997 (4360 observaciones). Dicha base incluye las siguientes variables:

Lwage: salario por hora (en logaritmos)  
Educ: años de educación  
Exper: años de experiencia y *expersq* años de experiencia al cuadrado  
Union: variable binaria que vale 1 si el trabajador está sindicalizado en 0 en caso contrario  
t: tendencia temporal, vale 1 para 1990, 2 para 1992, etc.  
Sect1 a Sect9: variables binarias que recogen el sector de actividad de la empresa.

En primer lugar se estima un modelo de regresión lineal por MCO. Los resultados son:

Salida 1

```
. reg lwage union t educ exper expersq
```

Source	SS	df	MS	Number of obs = 4360		
Model	210.020833	5	42.0041667	F( 5, 4354) = 178.16		
Residual	1026.50881	4354	.235762244	Prob > F = 0.0000		
Total	1236.52964	4359	.283672779	R-squared = 0.1698		
				Adj R-squared = 0.1689		
				Root MSE = .48555		

  

	lwage	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
union		.1771933	.0171781	10.32	0.000	.1435155	.2108712
t		.0242261	.0065419	3.70	0.000	.0114008	.0370515
educ		.0940819	.0051965	18.10	0.000	.0838942	.1042697
exper		.072108	.0124511	5.79	0.000	.0476974	.0965185
expersq		-.0024994	.0007346	-3.40	0.001	-.0039395	-.0010593
_cons		-47.89742	12.90961	-3.71	0.000	-73.20682	-22.58802

Un analista plantea que la variable *union* podría ser endógena en el modelo estimado. Debido a ello propone estimar el modelo utilizando el estimador de Mínimos Cuadrados en 2 etapas. Propone como instrumento para *union* utilizar el conjunto de dummies que captan el sector de actividad de las empresas.

Se estima un modelo de probabilidad lineal para la variable *union*. Los resultados son los siguientes:

Salida 2

```
reg union educ exper expersq sect1-sect8
```

Source	SS	df	MS	Number of obs = 4360		
Model	50.7519868	11	4.61381699	F( 11, 4348) = 26.62		
Residual	753.592967	4348	.17331945	Prob > F = 0.0000		
				R-squared = 0.0631		
				Adj R-squared = 0.0607		
				Root MSE = .41632		
Total	804.344954	4359	.18452511			

  

union	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
educ	.0170313	.0042733	3.99	0.000	.0086534	.0254092
exper	.0408176	.0086717	4.71	0.000	.0238166	.0578186
expersq	-.0025552	.000612	-4.18	0.000	-.003755	-.0013554
sect1	-.214857	.0276757	-7.76	0.000	-.2691154	-.1605986
sect2	-.2736854	.028324	-9.66	0.000	-.3292149	-.2181559
sect3	-.2904177	.0332551	-8.73	0.000	-.3556147	-.2252207
sect4	-.0684348	.0264896	-2.58	0.010	-.1203678	-.0165017
sect5	-.0855727	.0230936	-3.71	0.000	-.130848	-.0402975
sect6	.0223529	.0232509	0.96	0.336	-.0232308	.0679365
sect7	.0475308	.0278377	1.71	0.088	-.0070453	.1021069
sect8	-.2136662	.0552587	-3.87	0.000	-.3220013	-.1053311
_cons	-.0103404	.0588193	-0.18	0.860	-.1256562	.1049754

Se procede a realizar un contraste de la significación conjunta de las variables sect1 a sect8, los resultados son:

Salida 3

```
. test sect1 sect2 sect3 sect4 sect5 sect6 sect7 sect8
```

```

F( 8, 4348) = 34.71
Prob > F = 0.0000

```

A continuación se estima el modelo por MC2E:

Salida 4

```
. ivreg lwage educ exper expersq (union = sect1-sect8)
```

Instrumental variables (2SLS) regression

Source	SS	df	MS	Number of obs = 4360		
Model	33.8077034	4	8.45192586	F( 4, 4355) = 169.04		
Residual	1202.72194	4355	.276170365	Prob > F = 0.0000		
				R-squared = 0.0273		
				Adj R-squared = 0.0264		
				Root MSE = .52552		
Total	1236.52964	4359	.283672779			

  

lwage	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
union	-.29104	.0757557	-3.84	0.000	-.4395597	-.1425203
educ	.1008443	.0050128	20.12	0.000	.0910167	.110672
exper	.1146839	.0111579	10.28	0.000	.0928089	.136559
expersq	-.0042524	.0007889	-5.39	0.000	-.005799	-.0027057
_cons	.0008352	.0699136	0.01	0.990	-.136231	.1379015

  

Instrumented: union  
Instruments: educ exper expersq sect1 sect2 sect3 sect4 sect5 sect6 sect7  
sect8

Por último se realiza un contraste de Hausman a partir de las estimaciones MCO y MC2E. El valor del estadístico de Hausman es 12.63. En la tabla siguiente se informan los valores del p-valor de este estadístico para distintas distribuciones

<i>Distribución</i>	<i>p-valor correspondiente a 12.63</i>
F(1,4360)	0.00
F(4,4360)	0.00
F(8,4360)	0.0004
CHI2(1)	0.004
CHI2(4)	0.013
CHI2(8)	0.125

Se pide:

- 1) Analice los resultados de la estimación MCO. En particular, interprete los resultados correspondientes a las variables *union* y *t*.
- 2) Razone las posibles causas de endogeneidad de la variable *union* en el modelo.
- 3) En base al razonamiento realizado en 2) y la evidencia disponible ¿son las dummies por sector buenos instrumentos para *union*?
- 4) Analice los resultados de la estimación MC2E y compárelos con los presentados en 1).
- 5) ¿Cuál de los estimadores MCO o MC2E considera más apropiado en el caso bajo estudio? Justifique rigurosamente.

**Pregunta 1** (10 puntos)

Se desea estudiar el impacto, en la capacidad exportadora de las empresas, de un programa de exoneraciones fiscales a la inversión. Todas las empresas del país pueden presentarse al programa de exoneración. No obstante, a la hora de la decisión de otorgar exoneraciones o no a una empresa se valoran aspectos tales como la creación de puestos de trabajo de buena calidad y la potencialidad exportadora de la empresa. Proponga una estrategia para realizar dicho estudio:

1. Indique que variables incluiría en el análisis, que modelo(s) estimaría señalando la variable dependiente y los regresores a utilizar y los métodos que utilizaría para proceder con la estimación de dicho(s) modelo(s).
2. ¿Reconoce alguna(s) limitación(es) a la estrategia que propone? ¿Cuál(es)? ¿Conoce alguna alternativa que pueda colaborar a lidiar con dichas limitaciones?

**Pregunta 2** (10 puntos)

Se cuenta con los datos de una muestra de adolescentes de entre 14 y 16 años a los que se les preguntó si asistían a la educación y su género. En el cuadro siguiente se informan las frecuencias según género,

	Varón	Mujer	Total
Asiste	218	263	481
No Asiste	730	774	1504
Total	948	1037	1985

Considere el siguiente modelo Logit:

$$\Pr(\text{Asiste}=1|\text{Sexo})=\Lambda(\beta_0 + \beta_1 \text{Mujer}) = \frac{\exp(\beta_0 + \beta_1 \text{Mujer})}{1 + \exp(\beta_0 + \beta_1 \text{Mujer})} = \frac{1}{1 + \exp(-\beta_0 - \beta_1 \text{Mujer})}$$

1. Obtenga estimaciones en el modelo Logit para Beta0 y Beta1
2. Someta a prueba la hipótesis H0: Beta1=0

**Valores críticos para un nivel de significación del 5%:**

Normal estándar	1.96
F(1,1984)	3.84
F(2,1983)	3.00
CHI2(1)	3.84
CHI2(2)	5.99

**Pregunta 3** (10 puntos)

El modelo de regresión no lineal de Poisson es útil para tratar problemas en donde la variable dependiente observada es un recuento del número de ocurrencias de cierto evento, como ser fenómenos de elección individual o eventos de naturaleza poco frecuentes. Sin embargo, debido a que dicho modelo supone que la esperanza y varianza coinciden, no trata en forma adecuada el problema habitual de sobre-dispersión que poseen este tipo de fenómenos. Es por ello que una forma de mejorar la especificación del modelo es utilizar el denominado Modelo Binomial Negativo. El mismo establece que, dado el siguiente modelo no lineal para la esperanza condicional,

$$E(y_i|x_i) = e^{x_i'\beta}$$

la función de densidad se define por:

$$f(y|x) = g(y_i, \theta) \left( \frac{e^{x_i'\beta}}{e^{x_i'\beta} + \theta^{-1}} \right)^{y_i} \left( \frac{\theta^{-1}}{e^{x_i'\beta} + \theta^{-1}} \right)^{\frac{1}{\theta}}$$

Donde  $g(y_i, \theta)$  es una función continuamente diferenciable y  $\theta > 0$  un parámetro que mide el grado de sobre-dispersión de la variable dependiente (cuanto mayor el  $\theta$  más es la sobre-dispersión).

Se pide:

1. Plantee la función de verosimilitud,  $L(\beta, \theta)$ , para una muestra de tamaño  $n$  de observaciones i.i.d.
2. Formule la log-verosimilitud, plantee el problema de maximización y encuentre las condiciones de primer orden para hallar el estimador máximo verosímil de  $\beta$ . Es posible obtener una solución analítica cerrada para este estimador? Fundamente.
3. Derive el efecto parcial asociado respecto de un regresor  $x_j$ . Qué alternativas conoce para evaluar dicho efecto?