ECONOMETRÍA CON ESTIMACIONES PARA MÉXICO Guía para la solución de los ejemplos

Autora: María Delfina Ramírez Posgrado/ Facultad de Economía / UNAM

Esta guía contiene las instrucciones de cómputo para los resultados numéricos y gráficos de los ejemplos trabajados en el libro. El orden que se sigue en este documento es acorde con la descripción de las Tablas en el Apéndice C: 'Bases de Datos usadas en las Estimaciones' en la parte final del texto.

CONTENIDO

TABLA C.1 (Comandos para EViews 7)* BIRTH 3 (Comandos para STATA 12)* BIRTH 4 TABLA C.2 Generando resultados Cap. 3 sección 3.3 con SPSS 14 Generando resultados Cap. 3 sección 3.4 con SPSS 16 Generando resultados Cap. 6 sección 6.4 con SPSS 18 Generando resultados sección 6.5 con SPSS 20 Generando resultados Cap. 8, Sección 8.7 con SPSS 21 TABLA C.4 Generando resultados Cap. 7 sección 7.5 con SPSS 27 Generando resultados Cap. 7 sección 7.6 con EViews 29 TABLA C.6 Generando resultados Cap. 6 Sección 6.2 con EViews 30 TABLA C.8 Generando resultados Cap. 6 Sección 6.6.1 con SPSS 33 TABLA C.9 Generando resultados Cap. 9 Sección 9.1 con SPSS 35 Generando resultados Cap. 9 Sección 9.2 con SPSS 36 TABLA C.11 Generando resultados Cap. 9, Sección 9.3 con SPSS 37 Generando resultados Cap. 9, Sección 9.3 con STATA 12 41

Tabla C.1

BIRTH (Comandos para EViews 7)*

Archivo: Birth2006.wf1

Sample @all IF gnipc < 5000

Ecuación 1:

LS birth c gnipc mi View / Residual diagnostics / Histogram – Normality test View / Actual, Fitted, Residual / Residual Graph

Gráficas:

Seleccionar presionando la tecla de control: gnipc mi birth (abrir como grupo) View / Graph / Basic type / Distribution / Histogram / Multiple graphs View / Graph / Basic type / Scatter / Multiple series: Lower triangular matrix

Ventana de comandos: show birth @sqrt(gnipc) @sqrt(mi) View / Graph / Basic type / Distribution / Histogram / Multiple graphs View / Graph / Basic type / Scatter / Multiple series: Lower triangular matrix / Fit lines: Regression line

Ecuación con variables transformadas: LS birth c @sqrt(gnipc) @sqrt(mi)

Ecuación final: LS birth c @sqrt(mi) Forecast: birthf

Gráfica:

Seleccionar presionando la tecla de control: mi birth birthf (abrir como grupo) View / Graph / Basic type / Scatter / Single graph – First vs. All

*Los comandos generan las tablas y figuras en el orden en que aparecen en el texto.

BIRTH (Comandos para Stata 12)

Archivo: Birth.dta

Guardamos el archivo en Excel en formato .csv o lo importamos desde Stata

El comando *insheet* no se puede utilizar con el tipo de extensión .xlsx que tienen los archivos Excel. Pero es posible ingresar cualquier archivo de Excel a STATA desde el menú *File / Import / Excel spreadsheet*. Y al seleccionar el archivo necesario lo importa directamente a STATA. En comando sería: *import excel ''dirección del archivo'', sheet (''hoja deseada'') firstrow*.

La opción *sheet* nos permite decidir qué hoja de excel y la opción *firstrow* indica a STATA que el nombre de las variables está en la primer fila.

Alternativamente es posible utilizar desde Excel: *Guardar como*, y en tipo seleccionar *CSV*(delimitado por comas). Esto genera un archivo con extensión .csv que se puede introducir a STATA utilizando el comando *insheet*.

Una vez ingresado el archivo en Excel podemos renombrar las variables, en este caso:

rename v1 BIRTH

rename v2 GNIPC

rename v3 MI

rename v4 PAIS

Los datos que están en rojo, vistos a través del comando edit, señalan que STATA está considerando dichos datos como tipo string. Para modificarlos y ser manipulados utilizamos el comando:

destring NATALIDAD INBPC MI, replace force

destring NATALIDAD INBPC MI, replace ignore (NA)

Ahora es posible utilizar los datos para el análisis estadístico y por lo tanto realizamos la regresión lineal propuesta:

birth = f(c,gnipc, mi) con la restricción: gnipc<5000

Source	SS	df	MS		Number of obs	=	111
					F(2, 108)	=	141.97
Model	9117.3988	2	4558.6994		Prob > F	=	0.0000
Residual	3467.85976	108 3	2.1098126		R-squared	=	0.7245
					Adj R-squared	=	0.7193
Total	12585.2586	110 1	14.411441		Root MSE	=	5.6666
BIRTH	Coef.	Std. Er	r. t	P> t	[95% Conf.	In	terval]
GNIPC	0015162	.000555	1 -2.73	3 0.007	0026165		0004158
MI	.225037	.02287	9 9.84	1 0.000	.1796868		2703872
_cons	19.40718	1.98179	3 9.79	0.000	15.47892	2	3.33544

. reg BIRTH GNIPC MI if GNIPC<5000

Para obtener los test de normalidad se guardan los valores de los residuos de la estimación:

predict errors, residuals

A continuación obtenemos estadísticos descriptivos de los errores de estimación:

sum errors if GNIPC<5000, detail

		Residual	5	
	Percentiles	Smallest		
1%	-9.991594	-11.04855		
5%	-8.106346	-9.991594		
10%	-7.085493	-9.868693	Obs	111
25%	-4.464684	-9.558644	Sum of Wgt.	111
50%	1098966		Mean	1.28e-08
		Largest	Std. Dev.	5.614802
75%	3.891991	10.34041		
90%	7.582382	10.77389	Variance	31.526
95%	8.927711	13.74673	Skewness	.2807546
99%	13.74673	15.37404	Kurtosis	2.533499

Residuals

Y se obtiene un test de normalidad de Jarque-Bera:

sktest errors if GNIPC<5000

. sktest error	rs if GNIE	PC<5000			
	Ske	wness/Kurtosis	tests for Norm	ality	
Variable	Obs	Pr(Skewness)	Pr(Kurtosis)	adj chi2(2)	joint Prob>chi2
errors	111	0.2082	0.2963	2.74	0.2543

Adicionalmente podemos obtener el histograma:

histogram errors if GNIPC<5000, width(1) frequency normal



Y para la gráfica de residuos

line errors obs if GNIPC<5000, title(BIRTH RESIDUALS)



Para realizar análisis descriptivo pedimos los histogramas de cada variable con el comando

histogram variable, frequency normal

Para el caso de la variable BIRTH

histogram BIRTH if GNIPC<5000, width(2.5) start(7.5) frequency</pre>







histogram MI if GNIPC<5000, width(5) start(0) frequency</pre>



Para las gráficas de dispersión, se utiliza el comando:

graph matrix GNIPC MI BIRTH if GNIPC<5000, half



Se generan las variables *GNIPC_sr* y *MI_sr* como las raíces de las variables GNIPC y MI respectivamente y se realizan las mismas gráficas (histograma y dispersión) junto con la de la variable BIRTH





histogram GNIPC_sr if GNIPC<5000, width(2.5) start(10) frequency







graph matrix GNIPC_sr MI_sr BIRTH if GNIPC<5000, half



Ecuación con las variables transformadas

MCO: birth = f(c,gnipc_sr, mi_sr) con la restricción: gnipc<5000

reg BIRTH GNIPC_sr MI_sr if GNIPC<5000

. reg BIRTH GNIPC_sr MI_sr if GNIPC<5000

Source	SS	df	MS		Number of obs	=	111
					F(2, 108)	=	156.20
Model	9352.19658	2 4	676.09829		Prob > F	=	0.0000
Residual	3233.06197	108	29.935759		R-squared	=	0.7431
					Adj R-squared	=	0.7383
Total	12585.2586	110 1	14.411441		Root MSE	=	5.4714
BIRTH	Coef.	Std. Er	r. t	P> t	[95% Conf.	In	terval]
GNIPC_sr	0780424	.050442	4 -1.55	5 0.125	1780281		0219433
MI_sr	3.438766	.359461	9 9.5	7 0.000	2.72625	4	.151282
_cons	8.134675	4.07323	9 2.00	0.048	.0608086	1	6.20854

Es posible observar que la variable GNIPC_sr es estadísticamente no significativa por lo que podemos eliminarla de la ecuación y a continuación presentamos la ecuación final:

reg BIRTH MI_sr if GNIPC<5000

Source	SS	df	MS		Number of obs	=	111
Model Residual	9280.53948 3304.71908	1 109	9280.53948 30.3185237		Prob > F R-squared Adj R-squared	= = =	0.0000
Total	12585.2586	110	114.411441		Root MSE	=	5.5062
BIRTH	Coef.	Std. E	rr. t	₽> t	[95% Conf.	In	terval]
MI_sr _cons	3.878268 2.299488	.22166 1.5481	91 17.50 99 1.49	0.000	3.438927 7689928	4 5	.317609 .367969

. reg BIRTH MI_sr if GNIPC<5000

Y obtenemos los valores pronosticados de la variable birth en base a la regresión anterior

predict BIRTH_f

Los valores obtenidos son utilizados para obtener una gráfica de dispersión que permita observar el ajuste de los valores pronosticados con los datos observados: *graph twoway* (*scatter BIRTH BIRTH_f MI, msize(small)*), *title("Valores observados y pronosticados"*)



TABLA C.2

Generando resultados Cap. 3 sección 3.3 con SPSS

- 1) Archivo de SPSS: *pea99bisred.sav*
- 2) Se usan datos no ponderados: *Datos / Ponderar casos /* elegir: *No ponderar casos / Aceptar*
- Aplicar filtro para usar casos seleccionados en la estimación: *Datos / Seleccionar casos /* elegir: Usar variable de filtro / Aceptar

A 10-11	Seleccionar
✓ d2ed	Todos los casos
🖉 d3ed	
🖉 d5ed	Si se satisface la condición
🐓 d6ed	Sila op
🔗 d4ed	
🔗 d1exp	
🔗 d2exp	Ejemplo
🔗 d3exp	Basándose en el rango del tiempo o de los casos
🔗 d4exp	
🔗 pea=1 & ttrvida<998 & ~ p	J3a=1 8
Unstandardized Predicted	I Value 🦉 💿 Usar variable de filtro:
🔗 d1ocu	
🔗 d2ocu	
🔗 dЗоси	
🔗 d4ocu	Resultado
🖉 d5ocu	
	Descartar casos no seleccionados
	Copiar casos seleccionados a un nuevo conjunto de datos
Alegou	Nombre de conjunto de datos:
✓ uooCu	
🖉 wa	Eliminar casos no seleccionados
stado actual: No filtrar casos	
0 contor	Perter Restablecer Canceler Avude

 4) Generar resultados separados para áreas 1 y áreas 2: Datos / Segmentar archivo / elegir: Organizar los resultados por grupos / Grupos basados en: áreas / Aceptar 5) Generar resultados separados para áreas 1 y áreas 2: Datos / Segmentar archivo / elegir: Organizar los resultados por grupos / Grupos basados en: áreas / Aceptar

Segmentar archivo				X
✓ var00003 p7d p8a p8b p8c p8d p9d	 △Analizar todos ○ Comparar los ○ Organizar los ○ Organizar los ○ Grupos ✓ área 	: los casos, no crea grupos resultados por gru basados en: s	ar grupos	
<pre></pre>	 Ordenar archi El archivo ya e 	vo <u>s</u> egún variables está ordenado	de agrupació	n
Estado actual: El análisis por grupos Aceptar Pegar	está desactivado. <u>R</u> establecer	Cancelar	Ayuc	ła

6) Analizar / Regresión / Lineal Variable dependiente

Independiente: Método:

inghordf añosedu Introducir

áreas = 2

	Variables introducidas/eliminadas ^{b,o}								
	Modelo	Variables introducidas	Variables eliminadas	Método					
	1	añoseduª		Introducir					
Ì	a. Todas las variables solicitadas introducidas.								
	b. áreas = 2								

c. Variable dependiente: inghordf

Resumen del modelo^b

Modelo	R	R cuadrado	R cuadrado corregida	Error típ. de la estimación				
1	.455ª	.207	.207	3.65808				
a. Variables predictoras: (Constante), añosedu								

b. áreas = 2

NOVA ^{b,c}	

ANOVA ^{b,o}							
Modelo		Suma de cuadrados	gl	Media cuadrática	F	Sig.	
1	Regresión	87099.985	1	87099.985	6508.947	.000ª	
	Residual	333241.467	24903	13.382			
	Total	420341.452	24904				

a. Variables predictoras: (Constante), añosedu b. áreas = 2

c. Variable dependiente: inghordf

Coeficientes^{a,b}

		Coeficientes no estandarizados		Coeficientes tipificados		
Modelo		В	Error típ.	Beta	t	Siq.
1	(Constante)	.564	.045		12.418	.000
	añosedu	.430	.005	.455	80.678	.000

a, áreas = 2

b. Variable dependiente: inghordf

Generando resultados Cap. 3 sección 3.4 con SPSS

7) Analizar / Regresión / Lineal Variable dependiente:

> Independientes: añosedu Inttrvida Inttactual

Método:

Introducir

lninhrdf

áreas = 2

Variables introducidas/eliminadas^b

Modelo	Variables introducidas	Variables eliminadas	Método
1	Inttactual, añosedu, Inttrvidaª		Introducir

a. Todas las variables solicitadas introducidas.

b. áreas = 2

Resumen del modelo^b

Modelo	R	R cuadrado	R cuadrado corregida	Error típ. de la estimación
1	.560ª	.314	.314	.62040

a. Variables predictoras: (Constante), Inttactual, añosedu, Inttrvida

b. áreas = 2

ANOVA^{b,c}

Modelo		Suma de cuadrados	gl	Media cuadrática	F	Sig.
1	Regresión	4380.741	3	1460.247	3793.917	.000ª
	Residual	9584.189	24901	.385		
	Total	13964.930	24904			

a. Variables predictoras: (Constante), Inttactual, añosedu, Inttrvida

b. áreas = 2

c. Variable dependiente: Ininhrdf

Coeficientes^{a,b}

		Coeficientes no	estandarizados	Coeficientes tipificados		
Model	lo	В	Error típ.	Beta	t	Siq.
1	(Constante)	415	.018		-23.437	.000
	añosedu	.093	.001	.539	100.194	.000
	Inttrvida	.134	.004	.232	37.132	.000
	Inttactual	.028	.003	.061	9.906	.000

a. áreas = 2

b. Variable dependiente: Ininhrdf

Resultados Cap. 4, sección 4.3

A las instrucciones de la regresión del apartado anterior se pide, en la opción de *Estadísticos*, la *Matriz de covarianzas*, para obtener las correlaciones de la Tabla 4.1 del texto:

1	Regresión lineal: Estadísticos
	Coeficientes de regresión Image: Ajuste del modelo Image: Estimaciones Image: Cambio en R cuadrado
	□ I <u>n</u> tervalos de confianza □ <u>D</u> escriptivos
	Nivel(%): 95 Correlaciones parciales y semiparciales
	Matriz de covarianzas Diagnósticos de colinealidad
	Residuos
	Durbin-Watson
	Diagnósti <u>c</u> os por caso
	O Valores atípicos fuera: 3 desviaciones típicas
	O Todos los c <u>a</u> sos
	Continuar Cancelar Ayuda

Generando resultados Cap. 6 sección 6.4 con SPSS

- 1) Archivo de SPSS: *pea99bisred.sav*
- 2) Se usan datos ponderados: *Datos / Ponderar casos /* elegir: *ponderar casos mediante: factormd / Aceptar*
- Aplicar filtro para usar casos seleccionados en la estimación: *Datos / Seleccionar casos /* elegir: *Usar variable de filtro / Aceptar*

😨 Seleccionar casos
Seleccionar casos
Aceptar <u>Pegar</u> <u>R</u> establecer Cancelar Ayuda

- 4) Generar resultados separados para áreas 1 y áreas 2: *Datos / Segmentar archivo /* elegir: Organizar los resultados por grupos / Grupos basados en: áreas / Aceptar
- 5) Generar las Tablas 6.2 y 6.3:

Analizar / Estadísticos descriptivos / Tablas de contingencia / elegir la variable DG para las Filas, la variable DSUP para las Columnas / elegir: casillas (como se ilustra) / continuar / aceptar

Tablas de continger	ncia			-	23
			Filas:		Estadísticos
id2ed2		^	🤣 DG		-stauisticos
🛷 rid2ed2ui2					Casillas
🖉 🔗 Unstandardized Re	sidual [rid3ed]				Eormato
id3ed2					
id3ed2ui2					
🖉 Unstandardized Re	sidual [rid4ed]				
id4ed2					
🛷 rid4ed2ui2					
🖉 Unstandardized Re	sidual [rid5ed]				
rid5ed2					
🛷 rid5ed2ui2					
🖉 Unstandardized Re	sidual [rid6ed]		0-h		
rid6ed2			Columnas:		
rid6ed2ui2			V DSOF		
Unstandardized Re	sidual [rid1exp]				
rid1exp2					
rid1exp2ui2					
Unstandardized Re	sidual [rid2exp]				
rid2exp2					
rid2exp2ui2					
Unstandardized Re	sidual [rid3exp]				
rid3exp2					
rid3exp2ui2					
Unstandardized Re	sidual [rid4exp]				
V rid4exp2		Capa 1	l de 1		
V rid4exp2ui2	Tablas de con	tingencia: M	ostrar en las casillas		
A neduc		general			
	-Recuentos-				
raizttrvida					
A raizttactual	Observado				
Vinttactual	Esperado				
Unstandardized Re Vinstandardized Re	Descenter		Devidence		
Vinstanuaruized Re	Porcentajes-		Residuos		
	Fila		No tipificado	s	
FDAD2	Columpa		Tinificados		
Predregsin			- npinoddo <u>s</u>		
Resreasin	[Tipificados o	orregidos	
Predregcon	Ponderacione	es no entera	\$		
Resregcon	- onder actorie	sa no entera	0		
	Redondear	recuentos de	casillas 🔘 Redondea	ar ponderaciones	de casos
Mostrar los gráfico	O Truncar rec	uentos de cas	ilļas 🔷 Truncar p	onderaciones de	casos
Suprimir tablas	O No efectuar	correcciones			
Aceptar		Continue		Aunala	
	1	Continuar	Cancelar	Ayuda	

6) Generando Tablas 6.4 y 6.5 y coeficiente Chi-cuadrado de Pearson: *Analizar / Estadísticos descriptivos / Tablas de contingencia / casillas / Recuentos / Observado / Esperado / Continuar*



Tablas de contingencia / Estadísticos / Chi-cuadrada / Continuar / Aceptar

Generando resultados sección 6.5 con SPSS

- Generar resultados separados para áreas 1 y áreas 2: *Datos / Segmentar archivo /* elegir: comparar los grupos / Grupos basados en: áreas / Aceptar
- 2) Analizar / Regresión / Lineal:

A R	Dependientes:	Estadísticos
R ▲ ♣ controlb ♣ hogar ▶ hogarnud ♣ sexo ♣ edad ♣ ugarnac ♣ dackvill ♣ numfijos ♣ escola ♣ numéscol ✔ afössedu ♣ educat	Piloque 1 de 1 Anterior Independientes: Image: Siguiegte Image: Siguieg	Gráficos Gugrdar Opciones
dicatnum condresi	Método: Introducir 💌	
💑 migracio 💑 informan	Variabl <u>e</u> de selección: Regla	
💑 p3	Etiquetas de <u>c</u> aso:	
nocupa 💦 ocupabis	Ponderación MCP:	

Aceptar

Los resultados de la regresión permiten construir las tablas 6.7 y 6.8.

Generando resultados Cap. 8, Sección 8.7 con SPSS

- 1) Archivo de SPSS: *pea99bisred.sav*
- 2) Se usan datos ponderados: *Datos / Ponderar casos /* elegir: *ponderar casos mediante: factormd / Aceptar*
- 3) Aplicar filtro para usar casos seleccionados en la estimación:
- Datos / Seleccionar casos / elegir: Usar variable de filtro (como se ilustra) /

Aceptar

Seleccionar casos	Seleccionar Todos los casos Si se satisface la condición St la op Muestra aleatoria de casos Eemplo Basándose en el rango del tiempo o de los casos Rango Sultar variable de filtro: Secultado Descartar casos no seleccionados
 asocu disocu disocu disocu disocu disocu disocu disocu disocu Estado actual: No filtrar casos 	O Descartar casos no seleccionados Cogpiar casos seleccionados a un nuevo conjunto de datos Nombre de conjunto de datog; Eliminar casos no seleccionados ar Restablecer Cancelar Ayuda

- 4) Generar resultados separados para áreas 1 y áreas 2: *Datos / Segmentar archivo /* elegir: Organizar los resultados por grupos / Grupos basados en: áreas / Aceptar
- 5) Analizar / Regresión / Lineal. Seleccionar como variable dependiente: lninhrdf y como regresores: 6 variables dummy de educación y las 4 dummy de experiencia. En la opción: Estadísticos pedir Diagnósticos de colinealidad / Continuar. En la opción: Guardar pedir Valores pronosticados No tipificados y Residuos No tipificados / Continuar / Aceptar.



Resumen del modelo^b

Modelo 1	R .557ª	R cuadrado .311	R cuadrado corregida .311	Error típ. de la estimación .61534			
a. Variables predictoras: (Constante), d4exp, d5ed, d4ed, d6ed, d3exp, d3ed, d2exp, d1ed, d1exp, d2ed							

b. áreas = 2

ANOVA^{b,o}

	lodelo	Suma de cuadrados	gl	Media cuadrática	F	Sig.
1	Regresión	1697952.255	10	169795.225	448425.742	.000ª
	Residual	3769397.032	9954901	.379		
	Total	5467349.287	9954911			

a. Variables predictoras: (Constante), d4exp, d5ed, d4ed, d6ed, d3exp, d3ed, d2exp, d1ed, d1exp, d2ed

b. áreas = 2

c. Variable dependiente: Ininhrdf

Coeficientes^{a,b}

		Coeficientes no	estandarizados	Coeficientes tipificados			Estadísticos d	e colinealidad
Modelo		В	Error típ.	Beta	t	Siq.	Tolerancia	FIV
1	(Constante)	.376	.000		786.780	.000		
	d1ed	.214	.001	.122	386.503	.000	.696	1.437
	d2ed	.472	.001	.290	872.776	.000	.629	1.589
	d3ed	.639	.001	.223	765.015	.000	.816	1.226
	d4ed	.869	.002	.151	558.134	.000	.945	1.058
	d5ed	1.137	.002	.188	696.486	.000	.953	1.049
	d6ed	1.579	.001	.524	1839.069	.000	.853	1.172
	d1exp	.237	.000	.138	478.266	.000	.834	1.199
	d2exp	.351	.001	.178	615.832	.000	.830	1.205
	d3exp	.408	.001	.156	545.029	.000	.849	1.178
	d4exp	.273	.001	.094	323.083	.000	.820	1.220
a	áreas = 2	•						

b. Variable dependiente: Ininhrdf

Aparecen dos nuevas variables en el archivo: 'Unstandardized predicted values' y 'Unstandardized residuals'

6) Con el comando *Gráficos* se genera un *Diagrama de dispersión simple* con los *Residuos no estandarizados en el eje Y* y las *Predicciones no estandarizadas en el eje X*:



Casos ponderados por FACTORMDEF

7) Con el comando Transformar se calcula la variable de residuos no estandarizados al cuadrado: Transformar / Calcular variable / CuadradoResiduos = RES_1**2

Con el mismo procedimiento del inciso (6) se genera la gráfica de los Residuos al cuadrado contra la Predicción no estandarizada.

8) Prueba Breusch-Pagan. Se estima una regresión lineal con el cuadrado de los residuos como variable dependiente y como regresores todas las variables explicativas del modelo inicial:



9) Regresión del ln(Salario Real) en función de las variables de educación, experiencia y ocupación para las áreas menos urbanizadas: Analizar / Regresión / Lineal. Seleccionar como variable dependiente: lninhrdf y como regresores: 6 variables dummy de educación, las 4 dummy de experiencia y las 20 variables dummy de ocupación.

En la opción: Estadísticos pedir Diagnósticos de colinealidad.

En la opción: Guardar pedir Valores pronosticados No tipificados y Residuos No tipificados / Continuar / Aceptar.

🛃 Regresión lineal		×
▶ P4_1	Dependientes: Biopue 1 de 1 Anterior Siguiegte Independientes:	Egtadísticos Gráficos Gugrdar Qpciones
	Método: Introducir	
d18ocu d19ocu d19ocu d20ocu d20ocu Unstandardized Predicted Value [pre_2] Unstandardized Predicted Value [RE_1cocuysexo] Unstandardized Residual [RES_1cocuysexo]	Variatie de selección: Regla Eliquetas de gaso: Ponderación MCP:	
Aceptar	Pegar Restablecer Cancelar Ayuda	

10) Se generan las gráficas de diagrama de dispersión (a) de los residuos no estandarizados contra el valor de predicción no estandarizado, (b) de los residuos no estandarizados al cuadrado contra la predicción no estandarizada, (c) del valor absoluto de los residuos no estandarizados contra la predicción no estandarizada, con el mismo procedimiento señalado en el inciso (6). Previamente debe calcularse el cuadrado de los residuos así como el valor absoluto de los residuos utilizando el comando *Transformar / Calcular variable*.

Figura 8.6



Casos ponderados por FACTORMDEF

TABLA C.4

Generando resultados Cap. 7 sección 7.5 con SPSS

- 1) Archivo de SPSS: *imar-tmi-2005.sav*
- 2) Crear una variable X1=1 para la constante.
- 3) Analizar / Regresión / Lineal Variable dependiente: tmi Independientes: % de Población analfabeta de 15 años o más % de Población sin primaria completa de 15 años o más % de Población sin primaria completa de 15 años o más % Ocupantes en vivienda sin drenaje ni servicio sanitario % Ocupantes en vivienda sin energía eléctrica % Ocupantes en vivienda sin agua entubada % Viviendas con algún nivel de hacinamiento % Ocupantes en vivienda con piso de tierra % Población en localidades con menos de 5000 habs. % Población ocupada con ingreso de hasta 2 s. mínimos

Método:

Introducir (incluye la constante de la ecuación)

Estadísticos:

Coeficientes de regresión, Ajuste del modelo, Diagnósticos de colinealidad / Continuar



Resumen del modelo						
Modelo	R	R cuadrado	R cuadrado corregida	Error típ. de la estimación		
1	.855ª	.732	.731	4.21973		
- Venichles and distances (Constants), as simple 2 and similar data a						

a. Variables predictoras: (Constante), pocingh2sm, vivsindrena, vivsinagua, vivsinener, plocmen5000, vivconhacin, analf15omas, vivcpisotie, sinprim15om

ANUAP	

	Modelo	Suma de cuadrados	gi	Media cuadrática	F	Sig.
	1 Regresión	118570.177	9	13174.464	739.885	.000ª
×	Residual	43518.122	2444	17.806		
	Total	162088.299	2453			

a. Variables predictoras: (Constante), pocingh2sm, vivsindrena, vivsinagua, vivsinener, plocmen5000, vivconhacin, analf15omas, vivcpisotie, sinprim15om

b. Variable dependiente: tmi

	Coeficientes ^a								
		Coeficientes no	Coeficientes tipificados			Estadísticos d	e colinealidad		
Modelo		В	Error típ.	Beta	t	Siq.	Tolerancia	FIV	
1	(Constante)	6.731	.464		14.517	.000			
	analf15omas	024	.019	033	-1.312	.189	.172	5.828	
	sinprim15om	.256	.015	.446	17.516	.000	.169	5.914	
	vivsindrena	.084	.008	.130	10.755	.000	.750	1.333	
	vivsinener	.241	.014	.241	17.661	.000	.591	1.692	
	vivsinagua	.014	.005	.035	2.667	.008	.636	1.573	
	vivconhacin	.183	.010	.312	18.845	.000	.400	2.499	
	vivcpisotie	.031	.007	.087	4.627	.000	.313	3.190	
	plocmen5000	.057	.003	.242	16.752	.000	.526	1.899	
	pocingh2sm	150	.009	333	-17.636	.000	.308	3.252	

a. Variable dependiente: tmi

					Proporciones de la varianza								
Modelo	Dimensión	Autovalores	Índice de condición	(Constante)	analf15omas	sinprim15om	vivsindrena	vivsinener	vivsinaqua	vivconhacin	vivcpisotie	plocmen5000	pocingh2sm
1	1	7.993	1.000	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00
	2	.706	3.364	.01	.00	.00	.00	.27	.10	.00	.01	.01	.00
	3	.510	3.957	.00	.00	.00	.76	.00	.03	.00	.02	.00	.00
	4	.294	5.218	.00	.00	.00	.04	.45	.76	.00	.00	.00	.00
	5	.237	5.804	.02	.05	.00	.02	.20	.08	.00	.22	.01	.00
	6	.115	8.349	.03	.01	.00	.00	.01	.00	.03	.00	.63	.00
	7	.090	9.442	.01	.25	.02	.11	.03	.02	.01	.57	.00	.00
	8	.024	18.153	.31	.04	.18	.03	.01	.00	.51	.16	.12	.02
	9	.019	20.779	.34	.17	.08	.01	.01	.01	.05	.01	.23	.67
	10	.012	25.798	.28	.47	.72	.02	.01	.00	.39	.00	.00	.30

Diagnósticos de colinealidad^a

a. Variable dependiente: tmi

Generando resultados Cap. 7 sección 7.6 con EViews

4) La matriz de correlación Rho_x:

Archivo de EViews: *imar-tmi-2005_bis.f1*

Se eligen las nueve variables de rezago socioeconómico para crear un grupo (G1X) de variables. Visualizando G1X y con la opción: *View / Covariance Analysis / Correlation / Method: Ordinary / OK*, se obtiene la matriz de correlaciones.

- 5) En la ventana que muestra la matriz G1X se activan los comandos *View / Principal Components / Table*, para obtener los Eigenvalues (valores propios), los Eigenvectors (vectores propios) y la matriz de correlaciones nuevamente, que se muestran en la Tabla 7.4.
- 6) En la ventana que muestra la matriz G1X se activan los comandos View / Principal Components / Display: Eigenvalues Plots / Display Graphs of: Eigenvalues (Scree plot) para visualizar la Gráfica de Quiebre de la Figura 7.1.
- 7) En la ventana que muestra la matriz G1X se activan los comandos *Proc / Make Principal Components /* en la ventana de *Save results* se elige la opción *Scaling: Normalize scores /* en la opción *Scores series names: f1 f2 / loadings matrix: loadmx / Aceptar.*

En *'loadmx'* se obtienen los 'factor loadings' de la Tabla 7.5. En f1 se obtiene el índice de marginación de CONAPO, el cual resulta de la multiplicación: Z * ci1, donde ci1 es el vector propio 1 dividido por la raíz cuadrada del valor propio 1, y Z es la matriz de variables socioeconómicas estandarizadas.

TABLA C.6

Generando resultados Cap. 6 Sección 6.2 con EViews

Archivo:

totmex.wf1

- La serie de términos de intercambio 1970.01 a 2012.04 'totmex' se grafica como aparece en la figura 2 del texto. La gráfica se debe congelar para poder usar la opción de insertar una línea vertical en el punto correspondiente a Julio de 1986: *freeze / line-shade / Orientation: vertical – bottom axis / Position: 1986M07.*
- 2) Se genera la variable de tendencia con la instrucción en la ventana de comandos:
 genr t = @trend(1969:12)
- 3) Se estima la regresión inicial: LS log(totmex) c t xpetxtot
- 4) En la ventana de resultados de la ecuación anterior se usa el comando *forecast* para pedir el valor de predicción de la variable totmex. El programa nombra a la predicción automáticamente: *Forecast name: totmexf*

Fo	recast	X
	Forecast equation EQ01	
	Series to forecast TOTMEX O LOG	(TOTMEX)
	S.E. (optional): GARCH(optional):	Method Static forecast (no dynamics in equation) Structural (ignore ARMA) Coef uncertainty in S.E. calc
	Forecast sample 1980m01 2012m04	Output Forecast graph Forecast evaluation
	✓ Insert actuals for out-of-sample obse	rvations Cancel

Presionando la tecla 'Ctrl' se señalan las variables *totmex totmexf* y se abren como grupo. En la ventana que muestra los datos de estas variables se genera la gráfica de la Figura 6.4: *View / Graph / Basic type / Line & symbol / Multiple series: single graph.*

5) Generando la variable dummy DU. Para cada uno de los sub-períodos, mediante el comando *Sample* que aparece en la parte superior del archivo de trabajo se establece el rango muestral como sigue:

Sample	
Sample range pairs (or sample object to copy) 1980m01 1986m07 OK IF condition (optional) Cancel	Se indica el primer período. Una vez establecido, en la ventana de comandos se escribe la instrucción: genr DU = 0
Sample X Sample range pairs (or sample object to copy) 1986m08 2012m04 OK IF condition (optional) Cancel	Se indica el segundo período. Una vez establecido, en la ventana de comandos se escribe la instrucción: genr DU = 1

- 6) Se vuelve a establecer el rango muestral 1980:01 a 2012:04 para estimar la ecuación (6.7) con la instrucción en la ventana de comandos: *LS log(totmex) c du t du*t xpetxtot du*xpetxtot*.
- 7) Nuevamente se genera el valor de la predicción con el comando *Forecast* de la ventana de resultados de la ecuación. El programa automáticamente asigna como nombre de la serie *totmexf*, pero se debe modificar para que no borre la predicción anterior, por ejemplo nombrando a esta segunda predicción como *totmexff*

-Forecast equation EQ02	
Series to forecast TOTMEX OLOG(FOTMEX)
Series names Forecast name: totmexff S.E. (optional): GARCH(optional):	Method Static forecast (no dynamics in equation) Structural (ignore ARMA) V Coef uncertainty in S.E, calc
Forecast sample 1980m01 2012m04	Output Forecast graph Forecast evaluation
✓ Insert actuals for out-of-sample obser	Cancel

- 8) Una vez más, presionando la tecla 'Ctrl' se señalan las variables totmex totmexff y se abren como grupo. En la ventana que muestra los datos de estas variables se genera la gráfica de la Figura 6.5: View / Graph / Basic type / Line & symbol / Multiple series: single graph.
- 9) En la ventana de resultados de la ecuación anterior se usa la opción: *Proc / Make residual series* y se guardan con el nombre apropiado:

Make Residuals	×
Residual type Ordinary Standardized	ОК
Name for resid series	Cancel
resid02	

- 10) En la ventana de datos de los residuos se genera el correlograma de los residuos con la opción *View / Correlogram / Level / Lags to include: 36 / OK*.
- Se estima el modelo restringido: LS log(totmex) c du t du*t, que junto con los resultados de la segunda ecuación permite construir la prueba F para probar dos restricciones.

TABLA C.8

Generando resultados Cap. 6 Sección 6.6.1 con SPSS

Archivo:

EmisionesCO2_21 Abril Datos2007.sav

 El archivo contiene las ocho variables dummy de las regiones consideradas omitiendo a la región D3 por ser la región base. La instrucción para generar la regresión inicial con los resultados que se presentan en la Tabla 6.9: *Analizar / Regresión / lineal*

ſ	Regresión lineal			×
Γ		Dependientes:		
	🐣 País	Miles de to	neladas métricas de CO2 ICO	Estadísticos
	A Miles de toneladas mét			Gráficos
	💑 Miles de toneladas mét	Bioque 1 de 1		Guardar
	💑 Miles de toneladas mét	Anterior	Siguiente	Gu <u>a</u> i uai
l	💑 Miles de toneladas mét		\$	Opciones
	💑 Miles de toneladas mét	Досн		
ł	💑 Miles de toneladas mét	🖉 DUS		
	💑 pcapemrate	🖉 DIN		
	💑 Miles de toneladas mét	🔗 DJA		
	💑 Miles de toneladas mét	🔗 D1		
	💑 country	🔗 DAUS		
	GDP2007ConstantUSd	DUE 🛷 DUE		
1	GDP2007CurrentUSdls	🕈 🔗 D2		
	🕹 VAINDporcpibtotWBdata			
	VAIND2007CurrentUS			
	🔗 VAIND2007ConstantMi			
	🛷 región			
	🛷 DCH			
	🛷 DJA			
	n 🖉 DIN			
	🔗 DUS	<u>M</u> ét	odo: Introducir 🔻	
l	🔗 D1			
1	✓ D2	Variable de sei		
	🔗 D3		Regia	
	🔗 DUE	Etiquetas de ca	S0:	
	🔗 DAUS			
	👍 VAR00002	Ponderación M	CP:	
	\$CASENUM ~= 35 and	₩		
	Aceptar	Pegar <u>R</u> establ	ecer Cancelar	Ayuda
L				

Generando resultados Cap. 6 Sección 6.6.2 con SPSS

- 2) Del menú principal se usa el comando para generar la gráfica en la Figura 6.7: Gráficos / Cuadros de diálogo antiguos /Dispersión puntos / Dispersión simple / definir / Se elige para el eje Y: Miles de toneladas métricas de CO2 / Se elige para el eje X: GDP 2007 MillonesUSdlsCorr / Aceptar.
- Del menú principal se usa el comando para generar la gráfica en la Figura 6.8: Gráficos / Cuadros de diálogo antiguos /Dispersión puntos / Dispersión simple / definir / Se elige para el eje Y: logemisionesco2 / Se elige para el eje X:

loggdpcorr / Etiquetar casos mediante País / Opciones /Mostrar el gráfico con las etiquetas del caso / Aceptar.

4) Se estima la regresión con resultados que se presentan en la Tabla 6.10: Analizar / Regresión / lineal / Observe que la dummy D3 aparece entre las variables explicatorias debido a que la región de referencia es ahora la región representada con D1 y se agrega el logaritmo natural del pib como variable explicatoria, en tanto que la variable dependiente es el logaritmo natural de las emisiones de CO2.

Regresión lineal	Callman Link M	
	Dependientes:	Estadísticos
🞸 Miles de toneladas m 🔺	V logemisionesCO2	Lausticus
💑 Miles de toneladas m	-Bloque 1 de 1	Gráficos.
💑 Miles de toneladas m		Guardar
💑 Miles de toneladas m	Anterior Siguiente	
💑 Miles de toneladas m	Independientes:	Opciones.
💑 Miles de toneladas m	🖉 DCH	
💑 Miles de toneladas m	DUS	
💑 pcapemrate	DIN .	
💑 Miles de toneladas m	DJA	
💑 Miles de toneladas m	Ø DAUS	
🚜 country	<i>I</i> → DUE	
뤚 GDP2007ConstantU	₩ D2	
🔗 GDP2007CurrentUSdls	A D3	
💑 VAINDporcpibtot/VB 🔛		
🔗 VAIND2007CurrentU		-
🔗 VAIND2007Constant		
🔗 región		
🔗 DCH		
🔗 DJA		
🔗 DIN		
🔗 DUS		
🔗 D1		
🔗 D2		
🔗 D3	Método: Introducir	•
🔗 DUE		
🔗 DAUS	Variable de selección:	
🚜 VAR00002	Regla	
ళ \$CASENUM ~= 35 a	Etiquetas de caso:	
🔗 Kg de Carbono por d		
🔗 Kg de Carbono por d	Ponderación MCP:	
🖋 Kg emitido de Carbo 👻	· · · · · · · · · · · · · · · · · · ·	

TABLA C.9

Generando resultados Cap. 9 Sección 9.1 con SPSS

- 1) Archivo: combilbis_texto.sav
- 2) Ponderar y Filtrar las observaciones en el archivo usando las instrucciones:
 (1) Datos / Ponderar casos / ponderar casos mediante: factorexp / Aceptar.
 (2) Datos / Seleccionar casos / Usar variable de filtro / Aceptar.

Seleccionar casos	×
	Seleccionar
💊 cuántcopas	Todos los casos
💰 frecincoomás	
\delta numundía	
💰 menosefecto	Si la op
💰 necesicopa	🔘 Muestra aleatoria <u>d</u> e casos
🕹 enfermó	Fiemplo
💑 irritable	Ljonpo
💑 alucinó	O Basándose en el rango del tiempo o de los casos
💰 convulsionó	Rango
💰 deseocopa	O Llear veriable de filtro:
💰 apesarsalud	
💰 apesarpsicol	p7edad >= 18 & p7edad <= 65 & p15d1ingre
🛷 factorexp	
🧳 patconsal 🛛 🖉	Resultado
🛷 ingmes	Tusunduo
🛷 Iningmes	 Descartar casos no seleccionados
🛷 Escola	Copiar casos seleccionados a un nuevo conjunto de datos
bebedor con proble	Nombre de conjunto de datos:
🧳 Jóvenes	
🖋 Jóvenes2 🔹	 Ljiminar casos no seleccionados
Estado actual: No filtrar casos	3
Aceptar <u>P</u> e	gar <u>R</u> establecer Cancelar Ayuda

Para construir la tabla 9.1 se genera una tabla de contingencia con las variables categóricas siguientes: *p5género* (=1 hombre, =0 mujer); *bebeprob* (=1 bebedor con problema, =0 no bebedor con problema), así como la variable con 7 niveles de educación *nivel 12_65*. Comandos en SPSS: *Analizar / Estadísticos descriptivos / Tablas de contingencia / Filas: bebeprob / Columnas: p5género / Capa 1 de 1: Nivel12_65 / Casillas / Recuentos: Observado / Porcentajes: Columna / Continuar / Aceptar.*

Generando los boxplots de las Figuras 9.1 y 9.2

- 4) Usar archivo de EViews *bebedores.wf1*
- 5) Este archivo contiene las 7 proporciones de 'Bebedor con Problema' por género dadas en la Tabla 9.2: *phombres, pmujeres*. Estas dos variables se abren conjuntamente como grupo de variables y se generan los boxplots de la Figura 9.1

con los comandos siguientes. En la ventana en que se muestran las dos variables: *View / Graph / Basic type / Boxplot / Multiple series: Single graph / OK.*

6) Con las proporciones se crean los odds para hombres y mujeres: *oddsh, oddsm*, y sus logits mediante el logaritmo natural de los odds: *logith, logitm*. Se abren conjuntamente *logith, logitm como grupo de variables* y en la ventana que muestra sus datos se genera la Figura 9.2: *View / Graph / Basic type / Boxplot / Multiple series: Single graph / OK.*

Generando resultados Cap. 9 Sección 9.2 con SPSS (archivo

combi1bis_texto.sav)

Desuente

- 7) Para construir la tabla 9.3 se genera una tabla de contingencia con las variables categóricas: p5género, bebeprob, así como la variable nivel 12_65bis con 5 niveles de educación. Comandos en SPSS: Analizar / Estadísticos descriptivos / Tablas de contingencia / Filas: nivel12_65bis / Columnas: p5género / Capa 1 de 1: bebeprob / Casillas / Recuentos: Observado / Porcentajes: Columna / Continuar / Aceptar.
- 8) Tabla 9.4: Analizar / Estadísticos descriptivos / Tablas de contingencia / Filas: CATEDUC / Columnas: género / Capa 1 de 1: bebeprob / Casillas / Recuentos Observado / Aceptar.

Tablas de contingencia

		Casos						
	Váli	dos	Per	didos	Total			
	N	Porcentaje	N	Porcentaje	N Porcer			
CATEDUC * género * bebedor con problema	25118863	100.0%	0	.0%	25118863	100.0%		

Resumen del procesamiento de los casos

Tabla de contingencia CATEDUC * género * bebedor con problema

<u>necu</u>	CHLO				
			géne		
bebedor con problema		hombre	mujer	Total	
0	CATEDUC	masedu	4722920	3475843	8198763
		menosedu	7644822	5911391	13556213
	Total		12367742	9387234	21754976
1	CATEDUC	masedu	1006472	89287	1095759
		menosedu	2182085	86043	2268128
	Total		3188557	175330	3363887

9) Las Tablas 9.5 y 9.6 resultan de cálculos sencillos generados con Excel con base en la Tabla 9.4.

TABLA C.11Generando resultados Cap. 9, Sección 9.3 con SPSS

- 1) Archivo: COE1SDEM_208_TEXTO.sav
- Aplicar la variable de filtro para trabajar con la sub-muestra analizada en el texto: *Datos / Seleccionar casos / Usar variable de filtro / Descartar casos no seleccionados / Aceptar*.



- 3) Ponderar casos: Datos / Ponderar casos / Ponderar casos mediante (variable de frecuencia): FAC / Aceptar.
- 4) Figura 9.4: Gráficos / Cuadros de diálogo antiguos / Diagramas de caja / Simple / Los datos del gráfico son: Resúmenes para distintas variables / Definir / Las cajas representan: ANIOS_ESC EDUC_PAREJA / Aceptar.



5) Instrucciones para generar el resultado del modelo lineal de probabilidad en la Tabla 9.7: Analizar / Regresión / Lineal / Dependientes: MOADEC / Independientes: ANIOS_ESC EDAD EDAD2 ANTIGU HMEN_nul HMAY_nul EDUC_PAREJA / Método: Introducir / Estadísticos: Estimaciones, Ajuste del modelo, Diagnósticos de colinealidad / Continuar / Aceptar.

		Resumen de	el modelo	
Modelo	R	R cuadrado	R cuadrado corregida	Error típ. de la estimación
1	.576ª	.332	.332	.384
a. Vari HMAY	ables predic _nu1, EDAD,	toras: (Constar , ANTIGU, HMEI	nte), EDUC_PARE N_nu1, ANIOS_ES	JA, IC, edad2

ANOVA^b

Modelo	Suma de cuadrados	gl	Media cuadrática	F	Sig.
1 Regresió	n 223594.065	7	31942.009	216714.480	.000a
Residual	450887.713	3059103	.147		
Total	674481.778	3059110			
a Variablea pr	distance: (Constants			EDED INTICU	1

a. Variables predictoras: (Constante), EDUC_PAREJA, HMAY_nu1, EDAD, ANTIGU, HMEN_nu1, ANIOS_ESC, edad2 b. Variable dependiente: MOADEC

Coeficientes^a Coeficientes tipificados Coeficientes no estandarizados Estadísticos de colinealidad Tolerancia Error típ, Beta Sia. EIV B (Constante) -36.377 .000 -.115 .003 ANIOS_ESC .040 .000. .393 609.222 .000 .526 1.901 EDAD .000. .021 48.668 .016 .321 98.296 .000 edad2 .000 .000 -.375 -115.476 .000 .021 48.384 ANTIGU .012 .000. .218 385.744 .000 .684 1.461 HMEN_nu1 -.014 .000 -.021 -38.106 .000 .726 1.378 HMAY_nu1 -.040 .000 -.094 -177.499 .000 .782 1.278 EDUC_PAREJA .007 .000. .073 118.932 .000. .582 1.717

a. Variable dependiente: MOADEC

6) Comandos para estimar la regresión logística, Tablas 9.8, 9.9, 9.10, 9.11: Analizar / Regresión / Logística binaria / Dependientes: MOADEC / Covariables: EDAD EDAD2 ANIOS_ESC ANTIGU HMEN_nu1 HMAY_nu1 EDUC_PAREJA / Método: Adelante RV (Razón de Verosimilitud hacia adelante ó forward likelihood ratio).

Bloque 1: Método = Por pasos hacia adelante (Razón de verosimilitud)

		Chi cuadrado	gl	Sig.
Paso 1	Paso	984006.615	1	.000
	Bloque	984006.615	1	.000
	Modelo	984006.615	1	.000
Paso 2	Paso	207301.338	1	.000
	Bloque	1191307.953	2	.000
	Modelo	1191307.953	2	.000
Paso 3	Paso	14555.097	1	.000
	Bloque	1205863.050	3	.000
	Modelo	1205863.050	3	.000
Paso 4	Paso	10590.410	1	.000
	Bloque	1216453.460	4	.000
	Modelo	1216453.460	4	.000
Paso 5	Paso	1666.912	1	.000
	Bloque	1218120.372	5	.000
	Modelo	1218120.372	5	.000
Paso 6	Paso	9333.093	1	.000
	Bloque	1227453.465	6	.000
	Modelo	1227453.465	6	.000
Paso 7	Paso	615.604	1	.000
	Bloque	1228069.069	7	.000
	Modelo	1228069.069	7	.000

Pruebas omnibus sobre los coeficientes del modelo

Resumen del modelo

	Paso	-2 log de la verosimilitud	R cuadrado de Cox y Snell	R cuadrado de Nagelkerke
	1	2.888E6	.275	.383
	2	2.681E6	.323	.449
	3	2.666E6	.326	.454
	4	2.656E6	.328	.457
•	5	2.654E6	.328	.457
	6	2.645E6	.331	.460
	7	2.644E6	.331	.461

		Va	riables en la	a ecuación			
		В	E.T.	Wald	gl	Sig.	Exp(B)
Paso 1ª	ANIOS_ESC	.330	.000	615257.439	1	.000	1.391
	Constante	-2.454	.004	366050.967	1	.000	.086
Paso 2 ^b	ANIOS_ESC	.306	.000	501459.097	1	.000	1.358
	ANTIGU	.105	.000	161553.576	1	.000	1.111
	Constante	-2.843	.004	429626.070	1	.000	.058
Paso 3°	ANIOS_ESC	.272	.001	283544.032	1	.000	1.312
	ANTIGU	.106	.000	162358.103	1	.000	1.112
	EDUC_PAREJA	.049	.000	14645.705	1	.000	1.050
	Constante	-2.958	.004	437388.872	1	.000	.052
Paso 4 ^d	ANIOS_ESC	.268	.001	273147.666	1	.000	1.307
	ANTIGU	.104	.000	156555.378	1	.000	1.110
	HMAY_nu1	140	.001	10510.604	1	.000	.869
	EDUC_PAREJA	.050	.000	15054.698	1	.000	1.051
	Constante	-2.762	.005	329905.485	1	.000	.063
Paso 5°	edad2	.000	.000	1662.325	1	.000	1.000
	ANIOS_ESC	.264	.001	260548.586	1	.000	1.303
	ANTIGU	.108	.000	147475.716	1	.000	1.114
	HMAY_nu1	145	.001	11092.215	1	.000	.865
	EDUC_PAREJA	.048	.000	13384.406	1	.000	1.049
	Constante	-2.590	.006	165114.903	1	.000	.075
Paso 6 ^f	EDAD	.104	.001	9204.683	1	.000	1.110
	edad2	001	.000	10168.304	1	.000	.999
	ANIOS_ESC	.261	.001	252162.576	1	.000	1.298
	ANTIGU	.107	.000	142506.287	1	.000	1.113
	HMAY_nu1	212	.002	18827.733	1	.000	.809
	EDUC_PAREJA	.045	.000	11951.033	1	.000	1.046
	Constante	-4.415	.020	48172.093	1	.000	.012
Paso 7 ⁹	EDAD	.100	.001	8318.968	1	.000	1.105
	edad2	001	.000	9818.928	1	.000	.999
	ANIOS_ESC	.261	.001	252223.985	1	.000	1.298
	ANTIGU	.107	.000	142803.203	1	.000	1.113
	HMEN_nu1	065	.003	615.780	1	.000	.937
	HMAY_nu1	218	.002	19405.606	1	.000	.804
	EDUC_PAREJA	.045	.000	11804.466	1	.000	1.046
	Constante	-4.259	.021	40818.016	1	.000	.014

Tabla de clasificaciónª

	Pronosticado				
	MOADEC				
Observado	0 1 Porcentaje				
Paso1 MOADEC 0	546984 457000 54.5				
1	246278 1808849 88.0				
Porcentaje global	77.0				
Paso 2 MOADEC 0	627970 376014 62.5				
1	276219 1778908 86.6				
Porcentaje global	78.7				
Paso 3 MOADEC 0	629386 374598 62.7				
1	284510 1770617 86.2				
Porcentaje global	78.5				
Paso 4 MOADEC 0	624178 379806 62.2				
1	279014 1776113 86.4				
Porcentaje global	78.5				
Paso 5 MOADEC 0	616722 387262 61.4				
1	276221 1778906 86.6				
Porcentaje global	78.3				
Paso 6 MOADEC 0	619821 384163 61.7				
1	279494 1775633 86.4				
Porcentaje global	78.3				
Paso 7 MOADEC 0	621565 382419 61.9				
1	277222 1777905 86.5				
Porcentaje global	78.4				

a. El valor de corte es .500

Generando resultados Cap. 9, Sección 9.3 con STATA 12

1) Archivo: COE1SDEM_208_TEXTO.dta

use C:\Users\Maria\Documents\ COE1SDEM_208_TEXTO.dta

2) Método forward likelihood ratio

logit MOADEC ANIOS_ESC ANTIGU EDAD edad2 HMEN_nu HMAY_nu EDUC_PAREJA if filter_==1 [fweight = FAC], iterate(10)

. logit MOADEC ANIOS_ESC ANTIGU EDAD edad2 HMEN_nu HMAY_nu EDUC_PAREJA if filter_==1 [fweight =
> FAC], iterate(10)

Iteration	0:	log	likelihood	=	-1936089.2
Iteration	1:	log	likelihood	=	-1366508.1
Iteration	2:	log	likelihood	=	-1322993
Iteration	3:	log	likelihood	=	-1322055.8
Iteration	4:	log	likelihood	=	-1322054.7
Iteration	5:	log	likelihood	=	-1322054.7

9111
9.07
0000
3172
2

MOADEC	Coef.	Std. Err.	z	₽> z	[95% Conf.	Interval]
ANIOS_ESC ANTIGU EDAD edad2 HMEN_nu HMAY_nu EDUC_PAREJA cons	.2607382 .1069783 .1001926 0013664 0647427 2182186 .0448111 -4.258861	.0005192 .0002831 .0010985 .0000138 .002609 .0015665 .0004124 .0210798	502.22 377.89 91.21 -99.09 -24.81 -139.30 108.65 -202.03	0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000	.2597207 .1064234 .0980396 0013935 0698563 2212889 .0440027 -4.300176	.2617558 .1075331 .1023456 0013394 0596291 2151483 .0456194 -4.217545

3) Tabla de clasificación: estat class

78.44%

. estat clas	S			
Logistic mod	el for MOADEC			
	True -			
Classified	D	~D	Total	
+	1777905	382419	2160324	
-	277222	621565	898787	
Total	2055127	1003984	3059111	
True D defin	ed as MOADEC != (0		
Sensitivity		Pr(+ D)	86.51%	
Specificity		Pr(- ~D)	61.91%	
Positive pre	dictive value	Pr(D +)	82.30%	
Negative pre	dictive value	Pr(~D −)	69.16%	
False + rate	for true ~D	Pr(+ ∼D)	38.09%	
False - rate	for true D	Pr(- D)	13.49%	
False + rate	for classified -	+ Pr(~D +)	17.70%	
False - rate	for classified -	- Pr(D -)	30.84%	

4) Estimación de la razón de 'odds' (odds ratio): logit, or

Correctly classified

. logit, or							
Logistic regression				Numbe	r of obs	=	3059111
				LR ch	i2(7)	=	1228069.07
				Prob	> chi2	=	0.0000
Log likelihood = -1322054.7				Pseud	lo R2	=	0.3172
MOADEC	Odds Ratio	Std. Err.	Z	₽> z	[95% Co	onf.	Interval]
ANIOS_ESC	1.297888	.0006738	502.22	0.000	1.2965	68	1.299209
ANTIGU	1.11291	.0003151	377.89	0.000	1.1122	93	1.113528
EDAD	1.105384	.0012143	91.21	0.000	1.1030	06	1.107766
edad2	.9986345	.0000138	-99.09	0.000	. 99860	75	.9986615
HMEN_nu	.9373086	.0024455	-24.81	0.000	.93252	78	.9421139
HMAY_nu	.8039497	.0012594	-139.30	0.000	.80148	51	.8064218
EDUC_PAREJA	1.04583	.0004313	108.65	0.000	1.0449	85	1.046676
_cons	.0141384	.000298	-202.03	0.000	.01356	62	.0147348

5) Predicción de probabilidades:

a. Primero se reestima el modelo logit y se pide la predicción de probabilidades: *quietly logit MOADEC ANIOS_ESC ANTIGU EDAD edad2 HMEN_nu HMAY_nu EDUC_PAREJA if filter_==1 [fweight = FAC], iterate(10)*

Predict proba_moadec

. quietly logit MOADEC ANIOS_ESC ANTIGU EDAD edad2 HMEN_nu HMAY_nu EDUC_PAREJA if filter_==1 [fwei
> ght = FAC], iterate(10)
. predict proba_moadec

```
(option pr assumed; Pr(MOADEC))
(394580 missing values generated)
```

b. El resultado indica que hay valores faltantes en las variables del modelo debido a los casos que no tienen observaciones. Pediremos al programa que nos genere una variable de probabilidad pronosticada para los casos que fueron utilizados en el modelo. Para ello usamos la opción e(sample)==1. Se estima de nuevo la regresión y se pide la predicción, con un nombre distinto, para los casos utilizados:

quietly logit MOADEC ANIOS_ESC ANTIGU EDAD edad2 HMEN_nu HMAY_nu EDUC_PAREJA if filter_==1 [fweight > = FAC], iterate(10).

predict proba_casos if e(sample)==1

. quietly logit MOADEC ANIOS_ESC ANTIGU EDAD edad2 HMEN_nu HMAY_nu EDUC_PAREJA if filter_==1 [fweig > ht = FAC], iterate(10)

```
. predict proba_casos if e(sample)==1
(option pr assumed; Pr(MOADEC))
(394690 missing values generated)
```

- c. Se observa que los valores faltantes en proba_MOADEC (394580) y los faltantes con proba_casos (394690), son diferentes. Para encontrar la diferencia usamos el comando sum: *sum proba_casos proba_moadec*
 - . sum proba_moadec proba_casos

Variable	Obs	Mean	Std. Dev.	Min	Max
proba_moadec	13280	.7072447	.2730236	.0027707	1
proba_casos	13170	.7066878	.2732017	.0027707	.9993545

El número de casos usados para la estimación es menor que el número de predicciones para proba_moadec. Generaremos una variable dummy llamada *casos_verdad = 1* para los casos utilizados en la estimación.

. gen casos_verdad=1 if proba_casos!=.
(394690 missing values generated)

Esta instrucción pide: generar casos_verdad=1 si proba_casos es diferente de "." Este punto lo pone Stata cuando no hay un valor de predicción de la probabilidad. d. Ahora generamos casos_verdad=0 si no se utilizaron esos casos para la estimación del modelo: *replace casos_verdad=0 if casos_verdad!=1*

. replace casos_verdad=0 if casos_verdad!=1
(394690 real changes made)

Esta instrucción pide que casos_verdad sea cero cuando casos_verdad es diferente de 1, de manera que todas las observaciones de casos verdad sean 1 o 0 dependiendo si fueron usados o no en la estimación.

Filter y casos_verdad tienen el mismo número de observaciones iguales a 1.

e. Ahora se pide la lista con las predicciones de la probabilidad para los casos que sí fueron utilizados en la estimación:

```
list MOADEC proba_casos if casos_verdad==1
```

f. Se estima el modelo logit pero en lugar de usar filter==1 utilizamos la variable casos_verdad==1. No hay cambios en los coeficientes, pero así nos aseguramos que los casos sean los mismos para la estimación y para la predicción de probabilidades.