



Universidad Carlos III de Madrid

ECONOMETRÍA

Guía Rápida de Gretl

Carlos J. Pérez

Índice

1. Introducción	1
2. Gestión de Datos	1
3. Estadística Descriptiva	3
3.1. Estadísticos Univariantes	3
3.2. Estadísticos Multivariantes	5
3.3. Gráficos	6
4. Añadir Variables	7
5. Trabajando con submuestras	7
6. Estimación por Mínimos Cuadrados Ordinarios	8
6.1. Estimación robusta de las varianzas y covarianzas de los coeficientes	10
6.2. Contrastes de hipótesis	11
7. Otros Modelos Lineales	12
7.1. Mínimos Cuadrados Ponderados	12
7.2. MCG Factibles	12
7.3. Variables Instrumentales y MC2E	13
8. Modelos no Lineales	14
9. Programación en Gretl	16
10. Comentarios adicionales	17

1. Introducción

Gretl es una aplicación diseñada para el análisis estadístico y la estimación de modelos econométricos. Es la herramienta fundamental de análisis empírico en la asignatura Econometría I y puede descargarse gratuitamente desde: <http://gretl.sourceforge.net/>. Esta guía contiene instrucciones de manejo básico de Gretl que son necesarias para el desarrollo de las prácticas de ordenador. En el menú **Ayuda** de la barra de herramientas de Gretl se encuentra la *Guía del Usuario* en formato **.pdf**, que contiene una visión mucho más detallada que la que se presenta aquí.

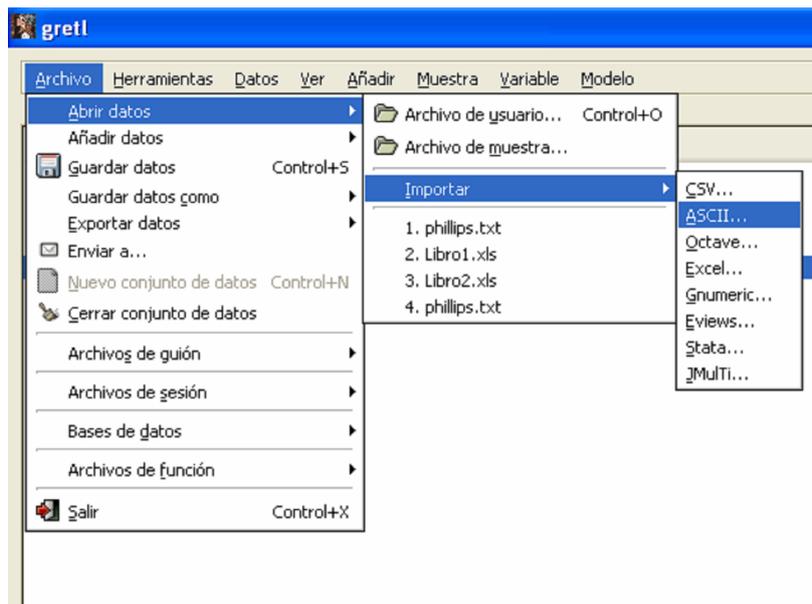
2. Gestión de Datos

Antes de empezar a trabajar con ellos, debemos introducir nuestros datos en Gretl. Gretl posee un formato propio para almacenar datos. Los archivos nativos (en formato propio) de Gretl tienen extensión **.gdt**, y se abren directamente mediante la secuencia **Archivo/Abrir Datos/Archivo de usuario**.

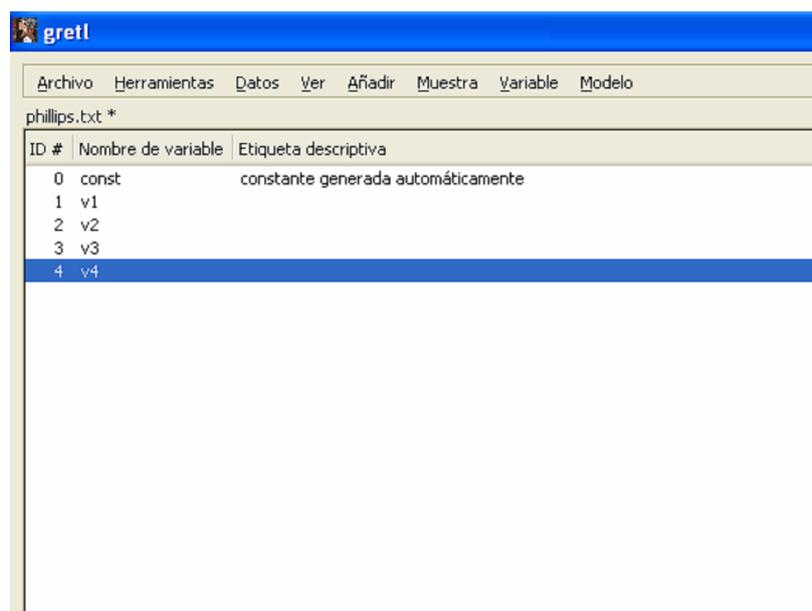
Además, Gretl puede importar archivos de distintos formatos como Stata, ASCII, Excel y Eviews Workfile, entre otros. En concreto, vamos a centrarnos en aprender cómo trabajar con los tres últimos. Tanto los archivos ASCII (archivos simples de texto **.txt**) como las hojas de Excel (**.xls**) suelen contener los datos dispuestos en columnas en cuya primera fila aparecen los nombres de las variables. Siempre que estén dispuestos de esta forma, Gretl importará los datos y dará esos nombres automáticamente a las variables. Si el archivo no contuviera una primera fila con los nombres de las variables, Gretl les asignará unos nombres por defecto. Si los datos estuvieran dispuestos en filas y no en columnas, deberíamos importarlos y después trasponerlos¹. Gretl lee también los archivos Eviews Workfiles (archivos nativos del paquete econométrico Econometric Views) y asigna directamente a las variables los nombres con los que fueron guardadas en la sesión de Eviews.

Para abrir cualquiera de estos tipos de archivos con Gretl, hay que seguir la secuencia **Archivo/Abrir Datos/Importar** en la barra de herramientas, y a continuación seleccionar el tipo de archivo en la ventana correspondiente, tal y como aparece en la figura.

¹Esto se explica más adelante.



Una vez que los datos han sido importados con éxito, aparecerán en la pantalla las variables importadas



Cuando trabajamos con datos ASCII o Excel conviene informar a Gretl de si éstos son de sección cruzada o de series temporales. Para ello, hay que seguir la secuencia **Datos/Estructura de datos/de sección cruzada** o **Datos/Estructura de datos/Serie Temporal** en la barra de herramientas, según proceda. En el menú **Datos** podemos elegir, entre otras, las siguientes opciones:

- **Mostrar/Editar valores:** permitirá ver y modificar los datos de las variables seleccionadas.

- Leer/Editar información: permite escribir y leer comentarios acerca del conjunto de datos con el que se trabaja.
- Trasponer datos: traspone la matriz de datos original de forma que las variables pasan a ser observaciones y las observaciones, variables. Así se soluciona, por ejemplo, el problemas de que los datos originales estuvieran organizados en filas y no en columnas.

Podemos cambiar el nombre de las variables, introducir una descripción de las mismas, elegir el nombre con el que aparecerán en los gráficos o el grosor de la línea con la que Gretl las dibujará, siguiendo la secuencia; **Variable/Editar atributos** desde la barra de herramientas. También podemos ordenarlas por nombre o número (que es la opción por defecto) mediante **Herramientas/Ordenar variables**.

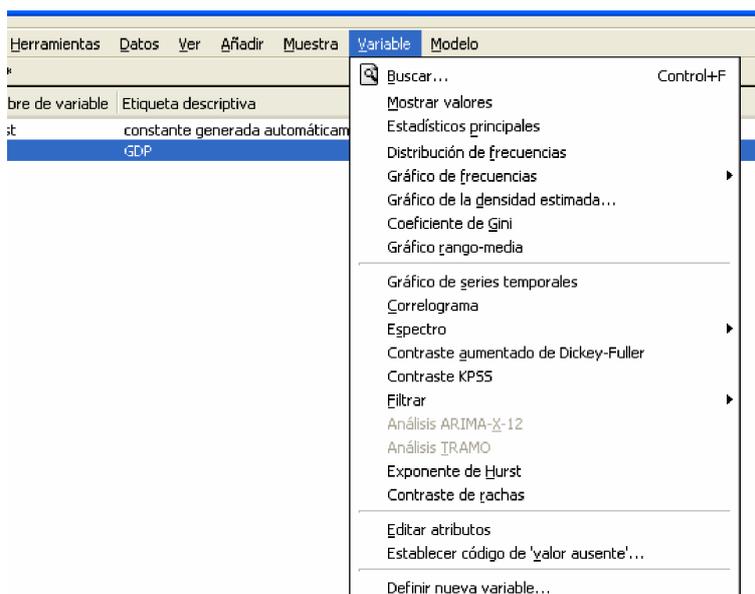
Gretl nos permite además declarar una variable como discreta desde **Variable/Editar atributos**. Al hacer esto, Gretl la interpretará como tal en el cálculo, por ejemplo, de su tabla de frecuencias. Para declarar una variable como discreta ésta debe tomar solamente valores enteros.

3. Estadística Descriptiva

Gretl proporciona información estadística acerca de los datos. Mediante los diferentes menús, podemos pedir a Gretl esa información, que será presentada en forma de ventana. Todas las ventanas de resultados de Gretl ofrecen la posibilidad de ser imprimidas, copiadas o guardadas en diferentes formatos (Word, Latex, ASCII).

3.1. Estadísticos Univariantes

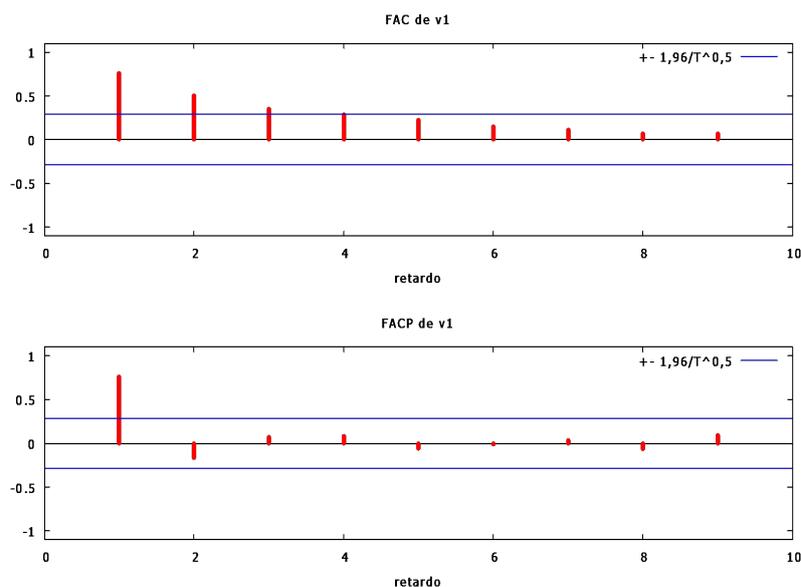
La primera parte del menú **Variable** contiene todas las opciones de análisis univariante para datos de sección cruzada. Este menú se encuentra activo cuando tenemos seleccionada una sola variable



La opción de **Estadísticos principales** ofrece una tabla con información básica sobre la variable seleccionada, como la media, las desviación típica, el coeficiente de asimetría, etc. La opción **Distribución de frecuencias** muestra una tabla resumen de frecuencias que incluye un estadístico de Jarque-Bera para contrastar normalidad. Mediante **Gráfico de frecuencias** existe la posibilidad de ver el clásico diagrama de barras frente a los gráficos de las distribuciones teóricas normal y gamma.

Para el análisis de series temporales, tenemos diversas opciones en la segunda parte del menú **Variable**. El **Correlograma** muestra las funciones de autocorrelación y autocorrelación parcial de la variable que hayamos seleccionado así como sus gráficos. Podemos elegir el número de retardos que Gretl incluirá en sus cálculos. En la tabla de salida, además del correlograma, aparecen los valores (y p-valores) del estadístico Q de Ljung-Box-Pierce cuya hipótesis nula es la ausencia de autocorrelación hasta el orden indicado por el retardo correspondiente. Las líneas azules en el gráfico de abajo representan bandas de confianza de ancho $\pm \frac{1,96}{\sqrt{T}}$ donde T es el tamaño muestral y 1,96 es el valor que deja un 5% de probabilidad en las colas de una normal estándar. Las dos figuras que siguen contienen la tabla y el gráfico del correlograma de la variable $V1$ del archivo `phillips.txt`.

Función de autocorrelación para $V1$						
RETARDO	FAC		FACP		Estad-Q	[valor p]
1	0,7591	***	0,7591	***	27,6482	[0,000]
2	0,5066	***	-0,1644		40,2223	[0,000]
3	0,3529	**	0,0726		46,4561	[0,000]
4	0,2897		0,0836		50,7510	[0,000]
5	0,2219		-0,0596		53,3274	[0,000]



3.2. Estadísticos Multivariantes

En primer lugar, debemos seleccionar un conjunto de variables con el botón izquierdo del ratón, bien arrastrándolo o bien manteniendo pulsado el botón control del teclado. En el menú **Ver**, además de la opción de **Estadísticos principales**, que es la versión multivariante de la descrita arriba, podemos obtener además una tabla que contiene la matriz de correlaciones de las variables seleccionadas. Por ejemplo, la matriz de correlaciones de las variables $V1$, $V2$ y $V3$ del archivo `ali.txt` es:

Coefficientes de correlación, usando las observaciones 1 - 965
 valor crítico al 5% (a dos colas) = 0,0631 para $n = 965$

V1	V2	V3	
1,0000	0,4889	0,2418	V1
	1,0000	0,6032	V2
		1,0000	V3

También en el menú **Ver**, la opción **Distancias de mahalnobis** proporciona la matriz de covarianzas y la inversa de la matriz de covarianzas de las variables seleccionadas. Para las variables $V1$, $V2$ y $V3$ de `ali.txt` tenemos las siguiente matriz de covarianzas:

Matriz de covarianzas

V1	V2	V3	
4,3003e+006	9,2850e+006	3,7721e+006	V1
9,2850e+006	8,3882e+007	4,1553e+007	V2
3,7721e+006	4,1553e+007	5,6578e+007	V3

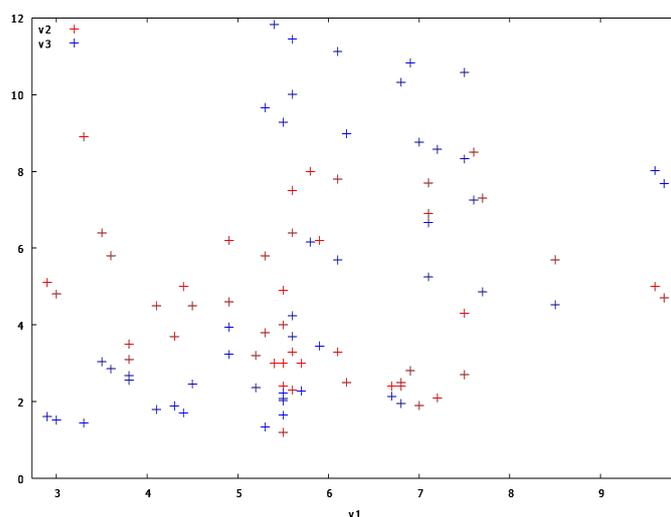
y su inversa:

Inversa de la matriz de covarianzas

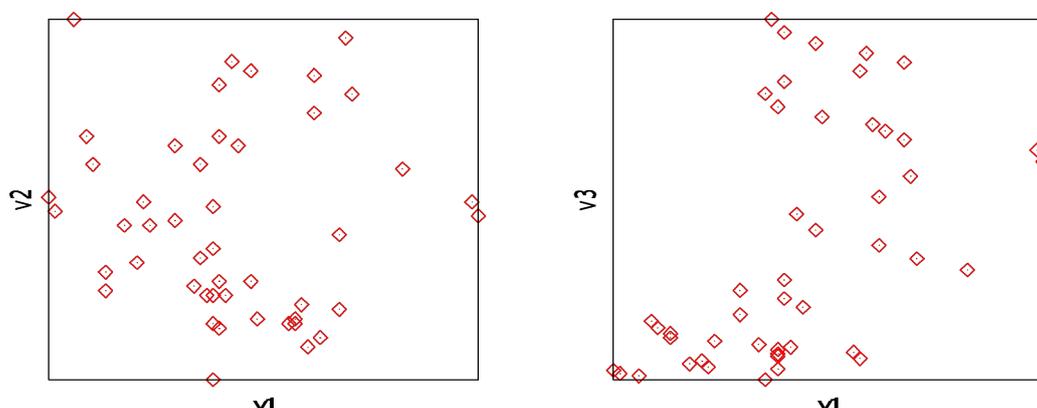
V1	V2	V3	
3,0736e-007	-3,7522e-008	7,0660e-009	V1
-3,7522e-008	2,3320e-008	-1,4626e-008	V2
7,0660e-009	-1,4626e-008	2,7945e-008	V3

3.3. Gráficos

Gretl permite visualizar gráficos de las variables. Las opciones de gráficos se encuentran en el menú **Ver**. Podemos ver gráficos de las series temporales de datos, así como ver gráficos de unas variables frente a otras. Esto es, podemos elegir qué variable tomará valores en el eje X y qué variable o variables en el eje Y , y así inspeccionar visualmente qué tipo de relación existe entre ellas. El siguiente gráfico es uno de tipo scatter con las variables v2 y v3 en el eje Y , y la variable v1 en el X .



También tienes la opción de visualizar las variables por separado en lugar de todas juntas mediante **Gráficos múltiples**. La siguiente figura muestra una versión de la anterior desagregada en dos gráficos



Los gráficos se pueden imprimir, copiar y guardar en diferentes formatos. Sólo hay que hacer click sobre el gráfico con el botón derecho del ratón y elegir la opción deseada.

4. Añadir Variables

Para trabajar con transformaciones de las variables originales, desde el menú **Añadir** podemos generar, entre otras; logaritmos, cuadrados, retardos y primeras diferencias de las variables que hayamos seleccionado previamente. Estas variables generadas aparecerán en la ventana de trabajo y podremos trabajar con ellas del mismo modo que con las originales. Además, Gretl incluirá una etiqueta descriptiva para que recordemos cómo se generó nuestra nueva variable.

Si quisiéramos utilizar otro tipo de transformación, como por ejemplo el producto de dos variables originales, que no aparece entre las opciones del menú, podemos generarla nosotros mismos mediante la opción **Definir una nueva variable** que también se encuentra en el menú **Añadir**. Sólo tenemos que escribir la fórmula de la nueva variable; para generar una nueva variable z como resultado del producto entre x e y bastaría con escribir

$$z = x * y$$

en la ventana correspondiente. Otro ejemplo, para obtener x_{t-5} , el quinto retardo de x_t y llamarlo $V5$ haríamos

$$V5 = x(5)$$

5. Trabajando con submuestras

Desde el menú **Muestra** de la barra de herramientas tenemos diversas opciones para determinar qué observaciones van a ser objeto de nuestro análisis. Todas ellas nos permiten trabajar con una submuestra del conjunto de datos original.

- **Establecer Rango:** Podemos cambiar el rango de los datos, seleccionando la primera y la última observación de la muestra original a considerar. Esto es muy útil en el análisis de series temporales, puesto que permite analizar aisladamente la submuestra correspondiente al periodo que se elija. Para volver a trabajar con la muestra completa, basta con marcar **Recuperar el rango completo**.
- **Definir la submuestra a partir de una variable ficticia:** Permite elegir una variable ficticia (que toma sólo los valores 0 y 1) para que actúe como criterio de selección muestral. Gretl construye la submuestra incluyendo sólo las observaciones en las que la variable ficticia seleccionada toma el valor 1.

- Restringir de acuerdo con un criterio lógico: Permite introducir una expresión lógica como criterio de selección muestral. Si, por ejemplo, queremos seleccionar sólo las observaciones para las que la variable $V1$ del archivo `ali.txt` toma valores superiores a 3000, hay que escribir $V1 > 3000$ en la ventana correspondiente:



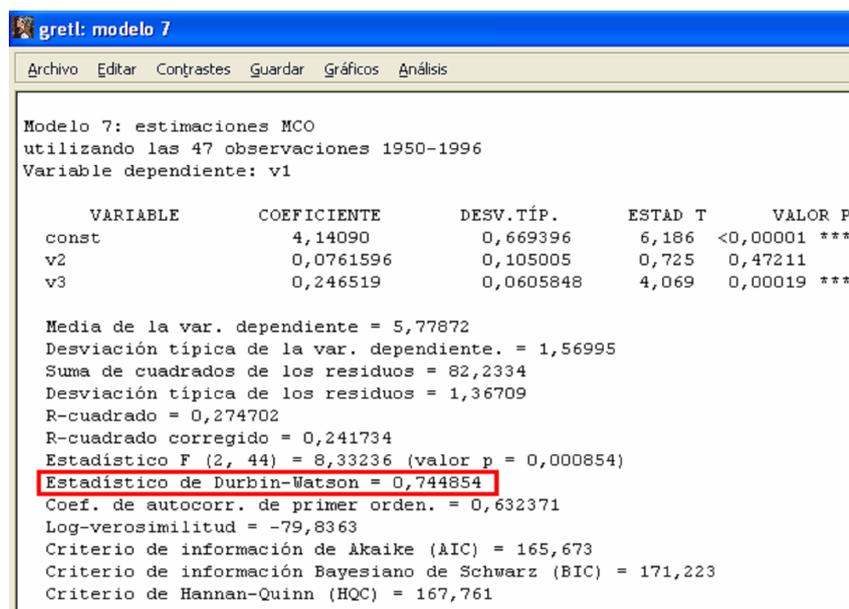
Las opciones **añadir a/Reemplazar la selección actual** permiten considerar o no restricciones impuestas con anterioridad. Para encadenar un conjunto de restricciones, seleccionaremos **añadir a la selección actual** cada vez que introduzcamos una nueva condición.

6. Estimación por Mínimos Cuadrados Ordinarios

La estimación de cualquier tipo de modelo se hace desde el menú **Modelo**. Para estimar un modelo por MCO debemos seleccionar esta opción y a continuación seleccionar la variable dependiente y la(s) explicativa(s) o independiente(s) dentro de la ventana que se muestra abajo. Para que las variables aparezcan en el lado derecho tenemos que presionar los botones **Elegir** y **Añadir**.



Al estimar el modelo, Gretl devuelve una ventana como ésta:



La tabla contiene la estimación de los coeficientes así como de sus desviaciones típicas. La tercera y cuarta columnas son el valor del estadístico t de significación individual (cuya hipótesis nula es que el coeficiente en cuestión es cero) y su p-valor asociado. Gretl calcula automáticamente el estadístico de Durbin-Watson cuando trabajamos con series temporales.

Gretl permite guardar información relevante acerca del modelo. Desde el menú **Guardar** podemos conservar los valores ajustados (\hat{v}_1), los residuos del modelo y los cuadrados de los residuos para luego trabajar con ellos. Estas nuevas variables se guardan en la ventana

de trabajo y se pueden utilizar del mismo modo que las variables originales. También se puede guardar parte de la información que aparece automáticamente con la estimación.

Tras estimar un modelo en Gretl (sea cual sea el procedimiento de estimación), podemos guardar la ventana de resultados en diferentes formatos, del mismo modo que lo hicimos en la sección de estadísticos descriptivos. Desde el menú **Archivo** de la ventana del modelo y seleccionando **Guardar como**, aparece una ventana como ésta



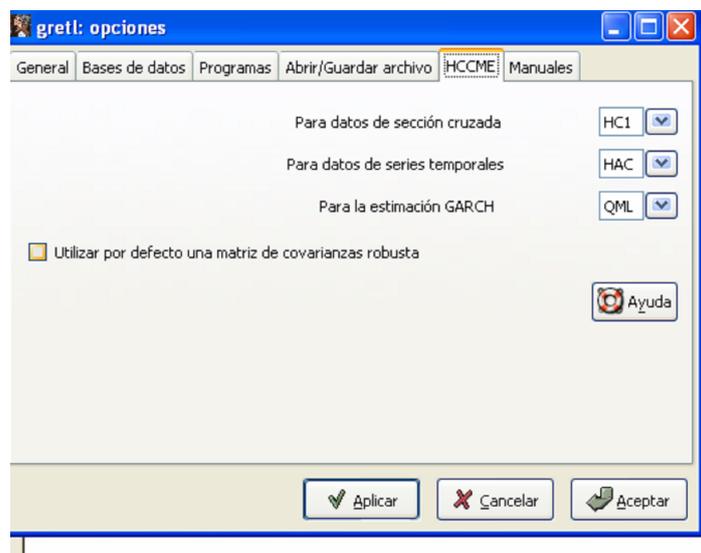
en la que podemos elegir el formato en el que guardar la información. La opción texto plano nos proporciona un archivo `.txt` mientras que la RTF (MS Word) genera un archivo de texto enriquecido compatible con Microsoft Word. Los archivos LaTeX son específicos del sistema de edición de textos científicos que lleva el mismo nombre (en concreto, esta Guía Rápida ha sido compuesta en LaTeX).

La estimación de la matriz de varianzas-covarianzas de los coeficientes se obtiene en **Análisis/Matriz de covarianzas de los coeficientes**.

6.1. Estimación robusta de las varianzas y covarianzas de los coeficientes

Se pueden hacer estimaciones robustas de la matriz de covarianzas de los estimadores con simplemente marcar la opción "Desviaciones típicas robustas.^{en} la parte inferior de la ventana. Aparecen diversas opciones para estimar la matriz de covarianzas. La opción HC0 produce los estimadores de White. Las demás opciones (HC1, HC2, HC3 y HC3a) son sucesivas variaciones que pretenden obtener mejores resultados (los detalles se pueden consultar en Mackinnon y White (1985) *Journal of Econometrics*). La opción HAC disponible para datos de series temporales incluye además retardos. Para modificar el número de retardos por defecto, puede consultarse el manual original de Gretl.

Si presionamos **Configurar** se abrirá la siguiente ventana, en la que podremos elegir el método de estimación para la matriz de covarianzas:



6.2. Contrastes de hipótesis

En el menú **Contrastes** de la ventana del modelo estimado se puede elegir entre un buen grupo de contrastes de especificación y sobre los residuos. El manejo de estos contrastes es inmediato si se conoce bien la teoría, por lo que no los enumeraremos aquí. No obstante, veremos cómo se introducen en Gretl las restricciones lineales sobre los parámetros. Gretl llama a los coeficientes bi , donde i es el orden que cada coeficiente ocupa en la regresión. Así, por ejemplo, si en el modelo

$$V1 = \beta_0 + \beta_1 V2 + \beta_2 V3 + \beta_3 V4$$

se quiere contrastar la hipótesis lineal de que

$$\beta_1 + 5\beta_2 = 0$$

$$\beta_3 = 4$$

se debe escribir

$$b1 + 5 * b2 = 0$$

$$b3 = 4$$

en la ventana correspondiente. Al realizar algunos de los contrastes, Gretl añadirá esta información a la ventana del modelo para que no haya que volver a calcularlos y se pueda exportar toda la información de una sola vez.

Por supuesto, Gretl basa los estadísticos en la opción robusta de la matriz de varianzas y covarianzas si es ésta la que se utiliza en la estimación.

7. Otros Modelos Lineales

Dentro del menú **Modelos** se accede a otros modelos lineales.

7.1. Mínimos Cuadrados Ponderados

Este método pondera cada observación por el valor de la variable que hayas seleccionado como ponderación. Si para estimar el modelo:

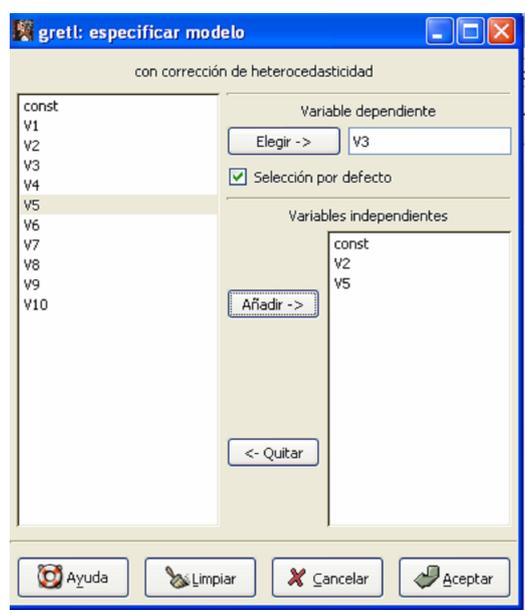
$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 \quad (1)$$

elegimo ponderar mediante la variable x_4 , entonces Gretl multiplicará y , x_1 , y x_2 por x_4 y efectuará la estimación de

$$(yx_4) = \delta_0 + \delta_1 (x_1x_4) + \delta_2 (x_2x_4)$$

7.2. MCG Factibles

Se accede a ellos mediante **Modelo/Otros modelos lineales/Corrección de heterocedasticidad**. La ventana emergente es idéntica a la de MCO



Sin embargo, en este caso Gretl efectúa primero la regresión MCO que se indique y guarda los residuos. Entonces regresa el logaritmo del cuadrado de esos residuos contra todas las variables independientes y sus cuadrados. Por último, usa los valores ajustados de esta segunda regresión para construir las ponderaciones con las que estimará (por mínimos cuadrados ponderados) el modelo original. La variable de ponderación la construye como $1/\sqrt{\exp(\hat{z})}$, donde \hat{z} son los valores ajustados en la segunda estimación.

7.3. Variables Instrumentales y MC2E

Al seleccionar la opción de estimación por mínimos cuadrados en dos etapas, Gretl nos pedirá que elijamos la variable dependiente, las independientes y los instrumentos. Como "variables independientes", Gretl se refiere a las variables del lado derecho de la regresión, sean estas exógenas o endógenas. Por otra parte, la lista de instrumentos debe incluir TODAS las variables exógenas que se quiera emplear como instrumentos. Es decir, si una de las variables explicativas del modelo es exógena, debe incluirse en ambas listas.

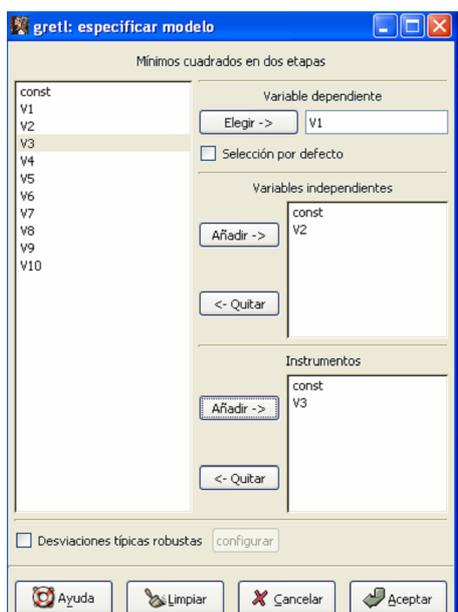
Por ejemplo, en el conjunto de datos `ali.gdt` se encuentran las variables

$V1 =$ Gasto en alimentación
 $V2 =$ Gasto total
 $V3 =$ Renta total

Queremos estimar la ecuación

$$V1 = \beta_0 + \beta_1 V2 + \varepsilon \quad (2)$$

pero sospechamos que $V2$ pudiera ser endógena. Podemos estimar (2) mediante MC2E usando $V3$ como instrumento. Basta con marcar en la ventana



para obtener una salida como la que sigue:

The screenshot shows the Gretl software window titled 'gretl: modelo 1'. The main area displays the following regression results:

VARIABLE	COEFICIENTE	DES.V.TÍP.	ESTAD T	VALOR P
const	2818,88	193,441	14,572	<0,00001 ***
V2	0,0907775	0,0106057	8,559	<0,00001 ***

Below the table, the following statistics are listed:

- Media de la var. dependiente = 4396,91
- Desviación típica de la var. dependiente. = 2073,71
- Suma de cuadrados de los residuos = 3,18678e+009
- Desviación típica de los residuos = 1819,13
- R-cuadrado = 0,238998
- R-cuadrado corregido = 0,238207
- Grados de libertad = 963
- Criterio de información de Akaike (AIC) = 17227,3
- Criterio de información Bayesiano de Schwarz (BIC) = 17237,1
- Criterio de Hannan-Quinn (HQC) = 17231,1

El R-cuadrado se calcula como el cuadrado de la correlación entre los valores observado y estimado de la variable dependiente.

Contraste de Hausman -
 Hipótesis nula: Los estimadores de MCO son consistentes
 Estadístico de contraste asintótico: Chi-cuadrado(1) = 5,64213
 con valor p = 0,0175339

Estadístico F de la primera etapa (1, 963) = 550,743
 Un valor < 10 puede indicar que los instrumentos son débiles

Cerrar

Gretl muestra automáticamente el estadístico y el p-valor asociado para el contraste de Hausman.

8. Modelos no Lineales

Seleccionando **Modelo/Modelos no lineales** tenemos la opción de estimar modelos logit y probit. No hay más que elegir la variable (binaria) dependiente y la lista de explicativas. El archivo `pfe.gdt` contiene datos de sección cruzada con 865 observaciones de mujeres españolas casadas con edades entre 25 y 60 años, seleccionadas aleatoriamente de la *Encuesta de Población Activa* correspondiente al segundo trimestre de 1990, elaborada por el Instituto Nacional de Estadística. Contiene las siguientes variables: $V1$ = participación (que toma el valor 1 si la mujer trabaja y 0 en caso contrario), $V2$ = edad (en años), $V3$ = Número de hijos menores de 6 años a cargo, $V4$ = Número de hijos entre 6 y 18 años a cargo, $V5$ = estudios universitarios (que toma el valor 1 si la mujer tiene titulación universitaria y 0 en caso contrario)

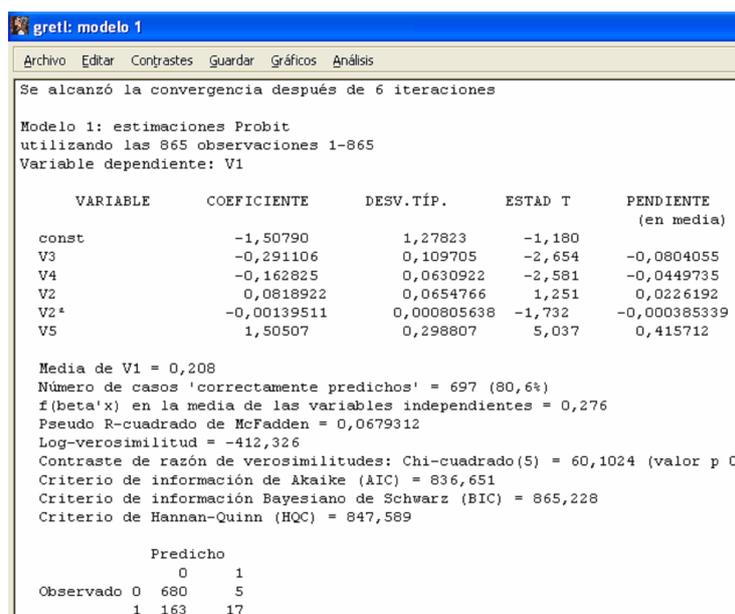
Queremos saber cómo afecta la composición familiar, el nivel de estudios y la edad (como aproximación de la experiencia laboral de la mujer) a la decisión de participación. Para ello podemos estimar el modelo

$$\Pr(V1 = 1 | V2, V3, V4, V5) = \Phi(\beta_0 + \beta_1 V3 + \beta_2 V4 + \beta_3 V2 + \beta_4 V2^2 + \beta_5 V5)$$

a través de la opción **Probit**. Si se seleccionan las variables tal y como aparecen en la figura



Gretl devolverá la siguiente ventana:



Como los efectos marginales (“pendientes” en la tabla) en estos modelos dependen del valor que tomen las últimas, Gretl calcula su valor en la media de las variables explicativas (sean éstas continuas o discretas). En la parte inferior de la ventana aparece un cuadro resumen de las coincidencias y divergencias entre las estimaciones del modelo y los valores que toma la variable endógena. Por ejemplo, de las 190 observaciones de mujeres que no trabajan vemos que el modelo probit predice que sí lo hacen 163 de ellas (utilizando el umbral $\widehat{V}_1 > 0,5$).

9. Programación en Gretl

Además de dar órdenes a Gretl a través de los diferentes menús, también podemos trabajar mediante comandos. Gretl acepta instrucciones directas (de una en una a través de una consola y también por lotes) y las ejecuta.

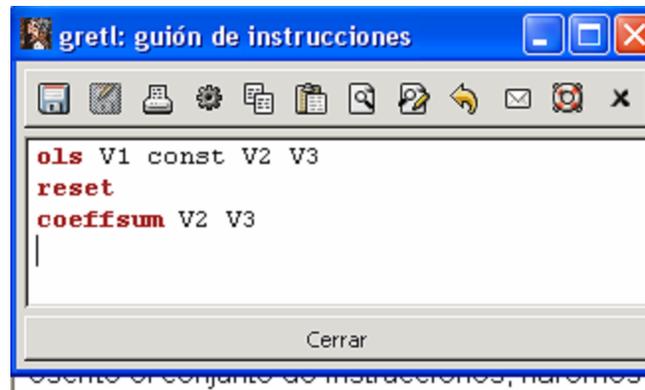
Para introducir lotes de instrucciones debemos seguir la ruta **Archivo/Archivos de guión/Nuevo guión** desde la barra de herramientas. Una vez hecho ésto se abrirá una ventana



en la que debemos escribir el conjunto de instrucciones que queremos que Gretl ejecute. Los dos primeros botones en la parte superior de esta ventana nos van a permitir guardar el lote de instrucciones que previamente hayamos introducido en un archivo por lotes. Gretl guarda estos archivos por lotes con la extensión `.inp`. Después de guardarlos, podremos abrirlos en otra ocasión mediante **Archivo/Archivos de guión/Archivo de usuario..y** volver a ejecutarlos. Una vez escrito el conjunto de instrucciones, haremos que Gretl las ejecute presionando el cuarto botón de la ventana (también podemos escribirlas en otro programa editor de archivos ASCII, como el *Block de Notas* de Windows, y copiarlas a ésta).

A diferencia del modo de trabajo mediante ventanas, todos los resultados correspondientes a un lote de instrucciones se muestran conjuntamente. Esto es muy útil en el caso de que queramos analizar múltiples resultados de una sola vez. Por otra parte, si vamos a hacer un uso continuado de transformaciones de variables suele ser más cómodo trabajar de este modo. También podemos ejecutar las instrucciones de una en una mediante una consola que se abre desde **Herramientas/Consola de Gretl**. Basta con introducir la instrucción deseada y Gretl la ejecutará al presionar Intro.

Veamos un ejemplo, el siguiente lote estima por MCO la regresión de $V1$ contra $V2$, $V3$ y una constante, y realiza el contraste RESET de Ramsey. Después calcula la suma de los coeficientes de $V2$ y $V3$, y el estadístico de contraste y p – valor para la hipótesis nula de que dicha suma es igual a cero:



```
ols V1 const V2 V3
reset
coeffsum V2 V3
```

10. Comentarios adicionales

Gretl dispone de bastantes alternativas para el análisis tanto con datos de series temporales como con datos longitudinales (datos de panel).

Además, pueden descargarse desde <http://gretl.sourceforge.net/win32/> programas complementarios para el análisis con series temporales para luego ser ejecutados desde Gretl. En particular, destacan los programas X-12-ARIMA y TRAMO/SEATS, para ajuste estacional en modelos ARIMA.

En esa misma página, pueden descargarse archivos de datos de conocidos libros de texto de econometría.