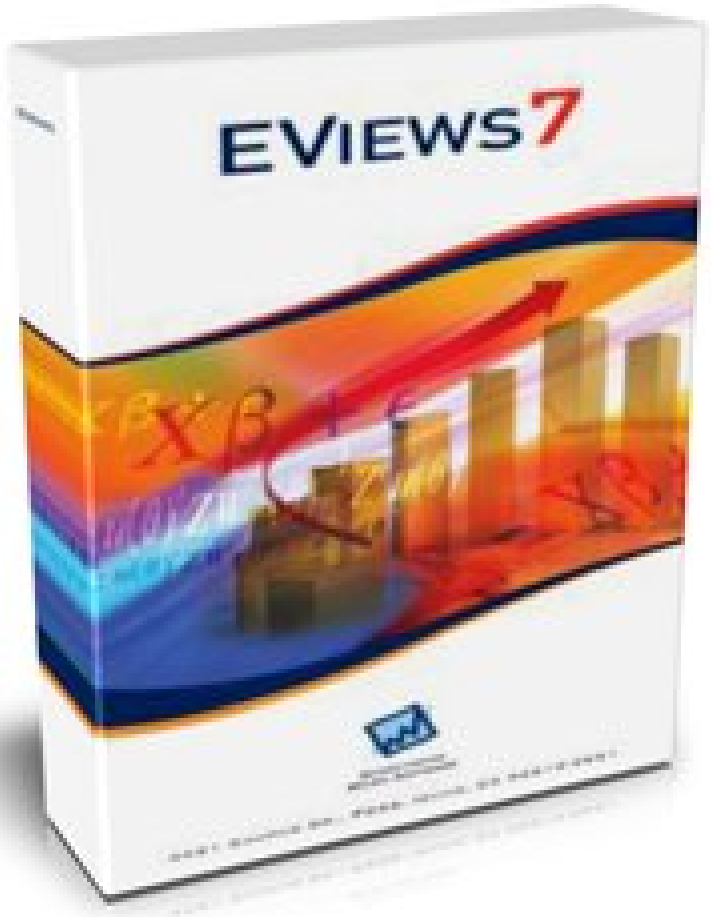
Apunte Introductorio de Eviews

Profesor: Rómulo Chumacero. Ayudantes: Adolfo Fuentes y Rodrigo Miranda.

1 de junio de 2014



Índice

1. Descarga e Instalación **3**

2. Manejo Básico de Eviews **6**
2.1. Cargar Bases de Datos 6
2.2. Regresión Simple por OLS (MCO) 12
2.3. Tests de Hipótesis asociados a los Coeficientes de la Regresión 14
2.3.1. Confidence Intervals 14
2.3.2. Wald Test - Coefficient Restrictions 15
2.4. Tests de Hipótesis asociados a los Residuos de la Regresión 16
2.4.1. Histogram - Normality Test 16
2.5. Tests de Hipótesis asociados a la Estabilidad de la Regresión 17
2.5.1. Chow Breakpoint Test 17
2.5.2. Ramsey RESET Test 18

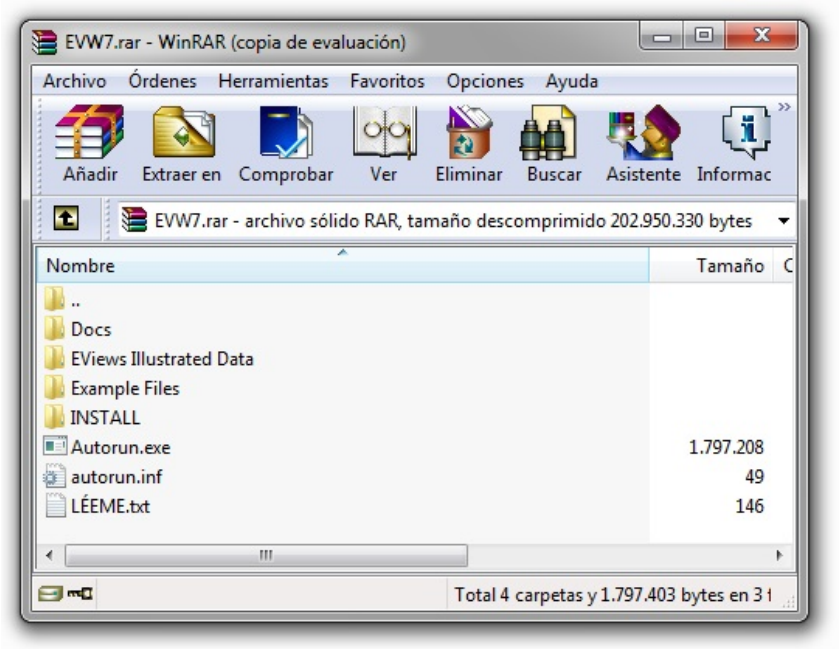
3. Tópicos de Econometría en Eviews **21**
3.1. Heterocedasticidad y Autocorrelación 21
3.1.1. Tests de Heterocedasticidad 21
3.1.2. ¿Cómo corregir heterocedasticidad y autocorrelación usando OLS? 24

4. Programación en Eviews **27**
4.1. Código que Realiza el Ejemplo Anterior 28
4.2. Explicación del Código 29

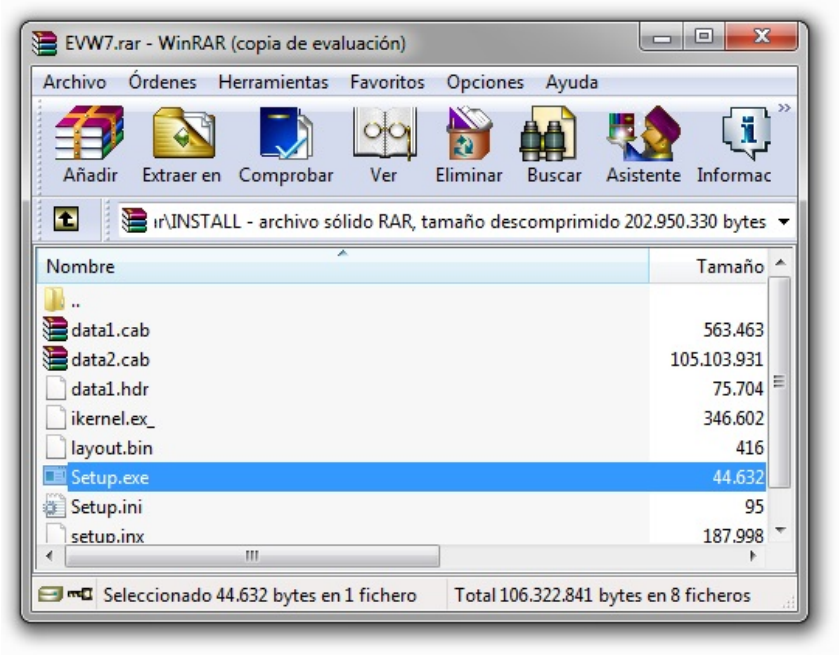
5. Tips Adicionales **32**
5.1. ¿Cómo Cargar Bases de Datos de Forma Eficiente? 32

1. Descarga e Instalación

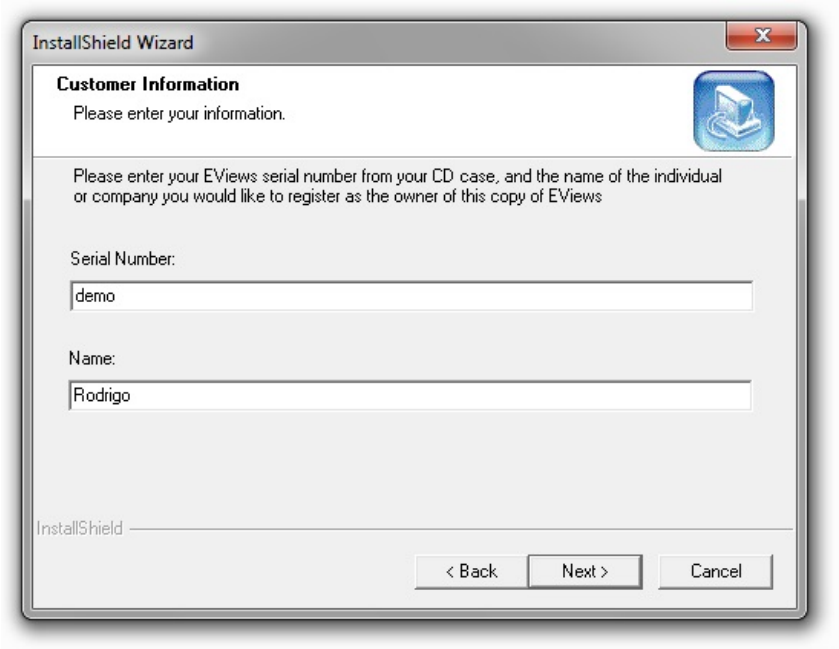
El software Eviews es un programa estadístico que requiere subscripción. Pese a ello existe una versión descargable de uso gratuito, que se puede descargar **aquí**. Si lo descargas aparecerá lo siguiente:



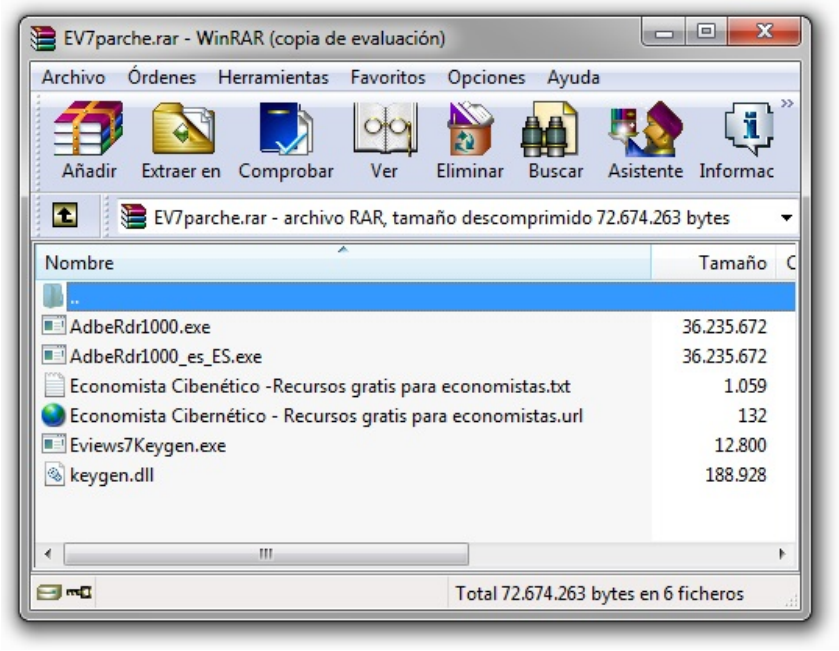
Selecciona la carpeta INSTALL y elige el archivo Setup.exe para instalar el programa.



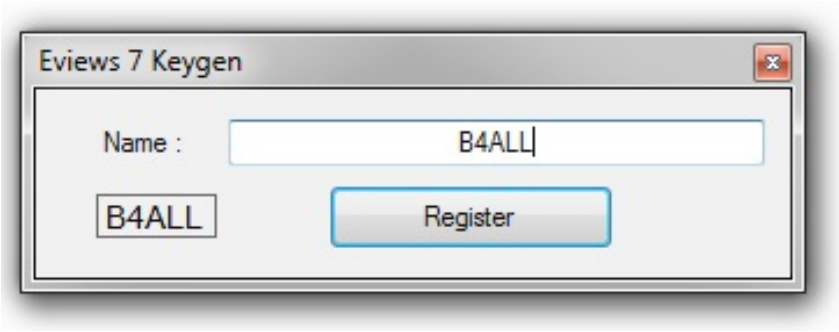
Durante el proceso de instalación, aparecerá una ventana que pedirá el Serial Number (número de serie) del programa. Para seguir la instalación debes poner como Serial Number lo siguiente: demo.



Sin embargo, para que el programa funcione, es necesario bajar un parche **aquí**. Si lo abres aparecerá lo siguiente:



Lo que debes hacer es extraer todos los archivos en la carpeta donde se instaló Eviews. Y despues seleccionar el archivo Eviews7Keygen.exe. Saldrá la siguiente ventana:

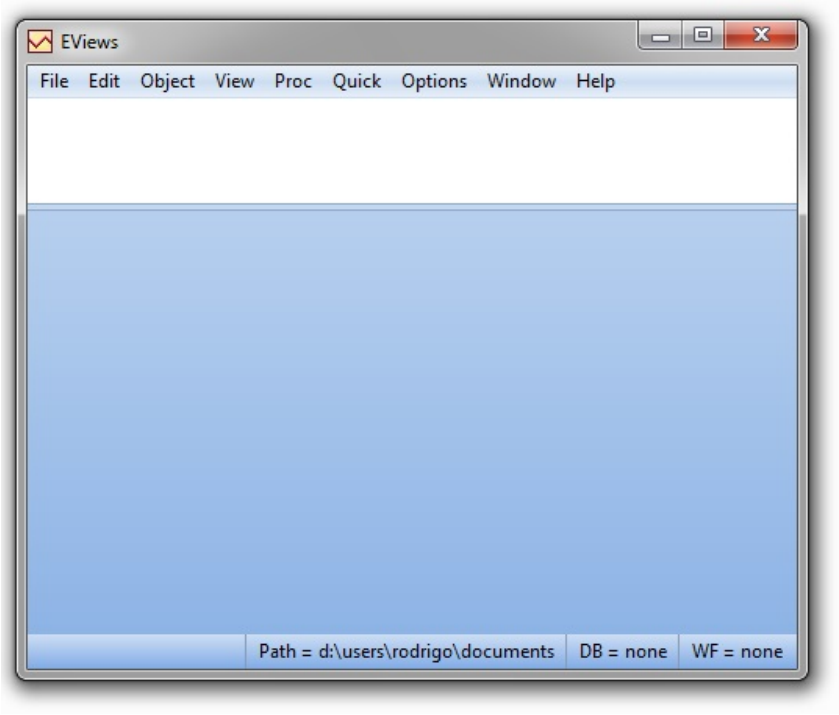


Donde dice Name se debe introducir: B4ALL. Luego apretar Register y listo: ya podemos empezar a usar Eviews7.

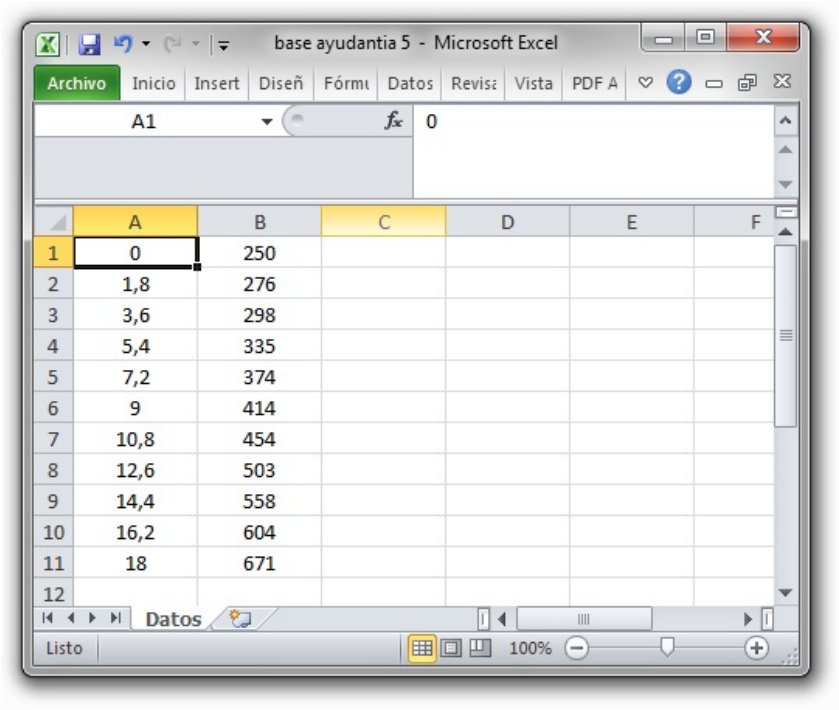
2. Manejo Básico de Eviews

2.1. Cargar Bases de Datos

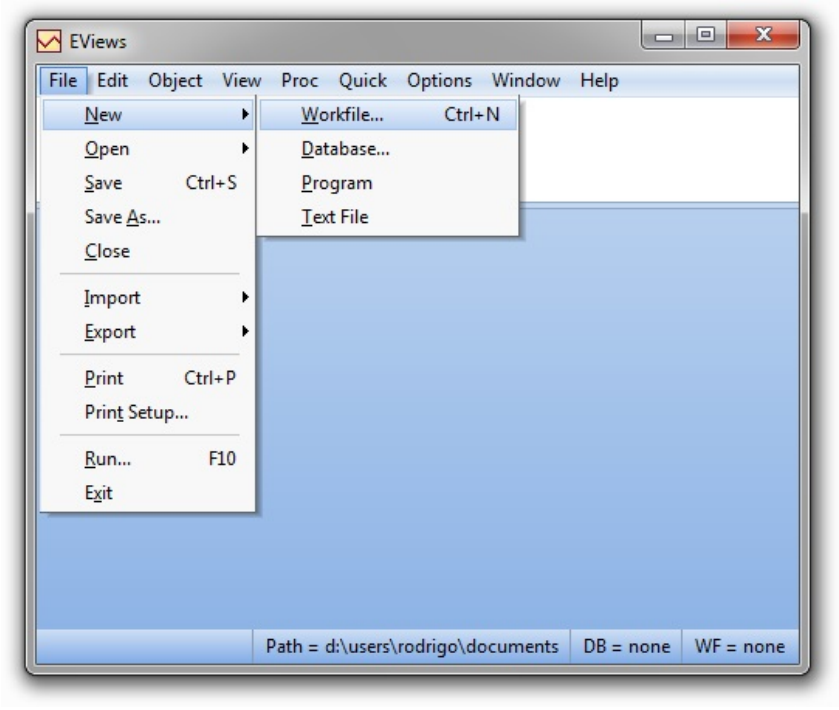
Al abrir el programa sale lo siguiente:



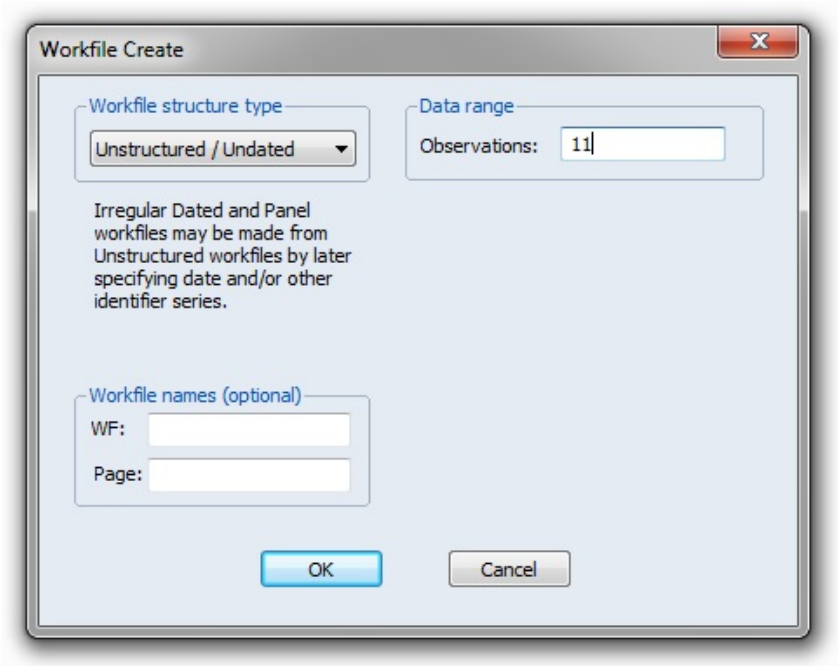
Para empezar a trabajar necesitamos una base de datos, como la que aparece abajo:



Primero, dado que la base tiene 2 variables y 11 muestras, con datos de corte transversal, debemos seleccionar **File-New-Workfile**:

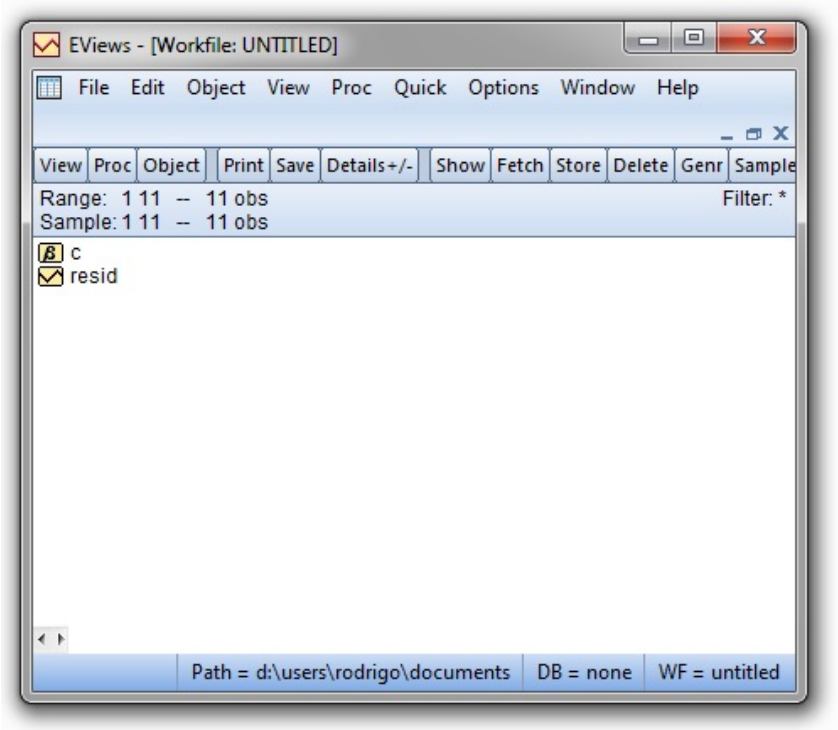


A continuación aparecerá la siguiente ventana:



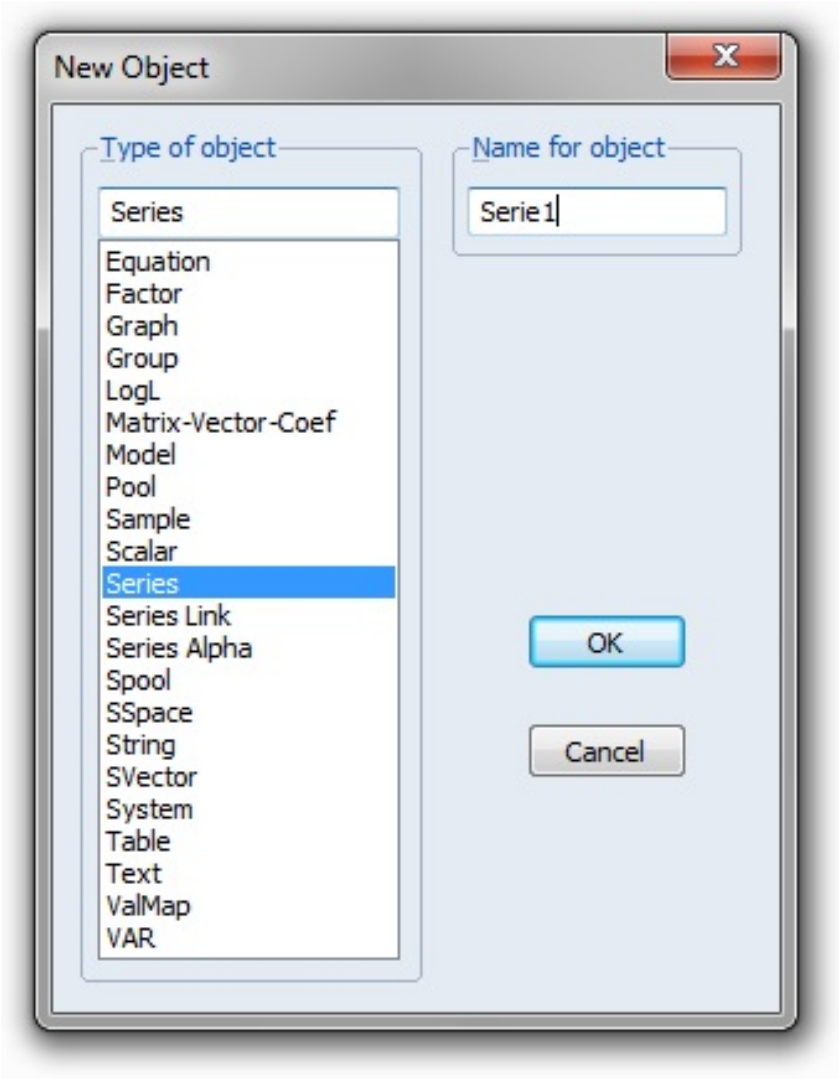
En Workfile structure type deben elegir el tipo de datos (corte transversal, serie de tiempo o panel de datos). Como trabajaremos con datos de corte transversal, elegiremos esa opción (Unstructured/Undated) y en Date range pondremos el número de observaciones (11).

Una vez elegido el tipo de datos, aparece lo siguiente:

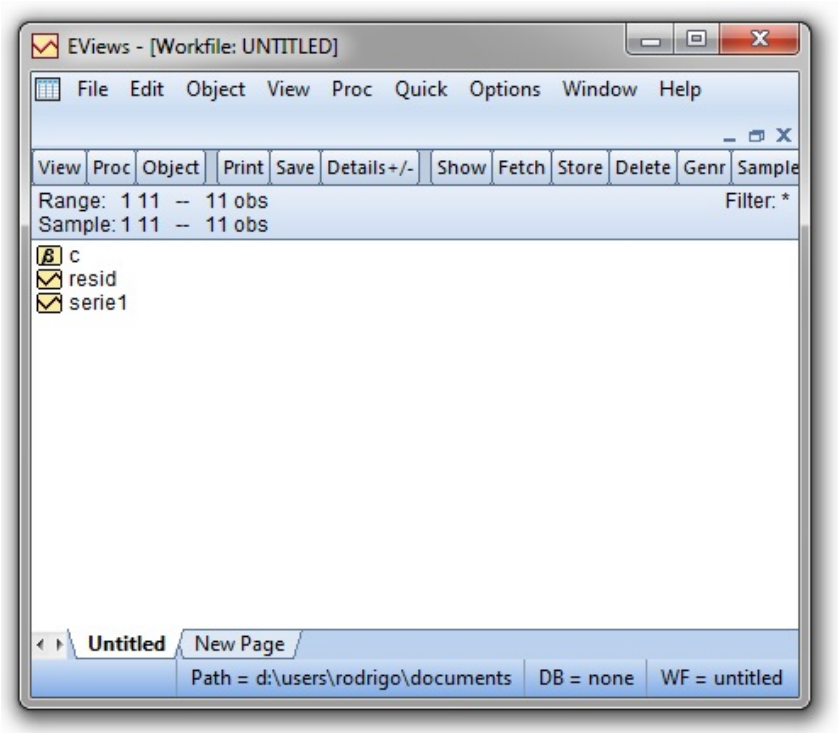


El elemento c contiene los estimadores (pero como no se ha estimado nada, contiene puros ceros) y el elemento resid los residuos o errores estimados (pero como no se ha estimado nada, contiene puros ceros).

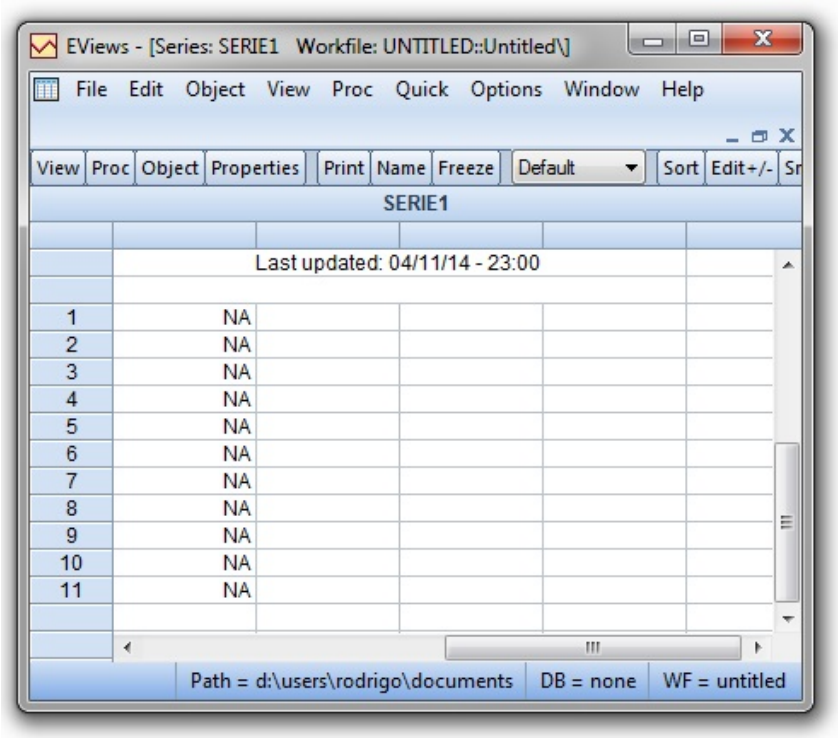
Ahora procederemos a importar los datos. Para ello debemos seleccionar la opción **Object - New Object** y aparecerá lo siguiente:



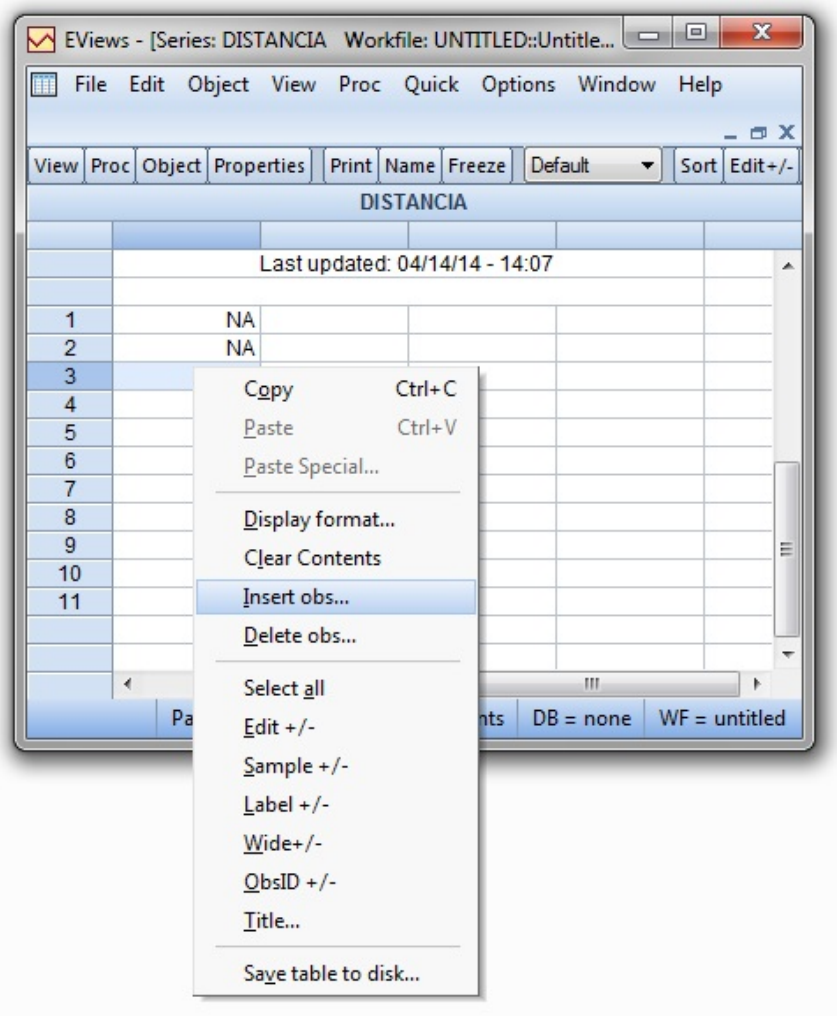
Elegimos la opción Serie y lo llamaremos Serie1. Aparecerá lo siguiente:



Si seleccionamos el elemento creado, serie1, aparecerá una planilla sin datos (NA), donde pondremos una de las dos series:



Para editar la planilla y pegar los datos, primero debemos apretar el botón derecho del ratón y elegir la opción Insert obs e indicar cuantas celdas deseamos editar (como deseamos editar todas pondremos el tamaño de la muestra). Solo entonces podemos pegar los datos desde la planilla Excel.



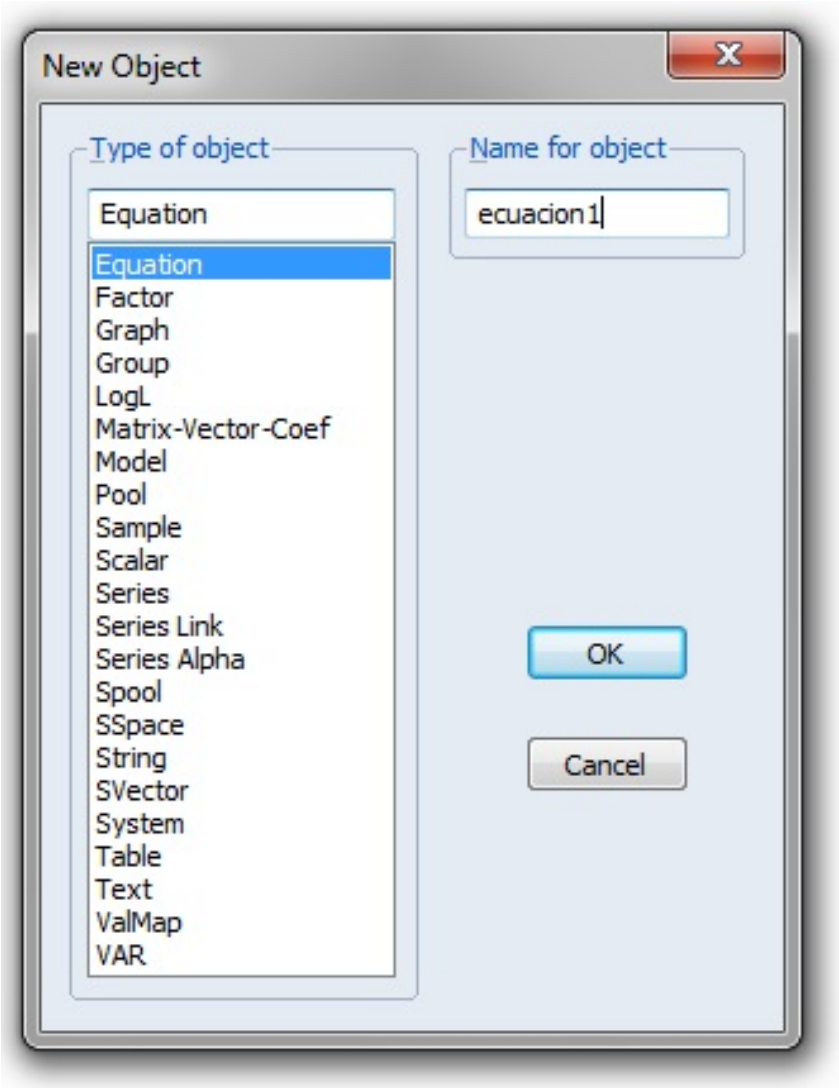
Haremos lo mismo con la otra serie. En particular, la base de datos define la primera serie como distancia (entre 0 y 18) y la segunda como dureza.

2.2. Regresión Simple por OLS (MCO)

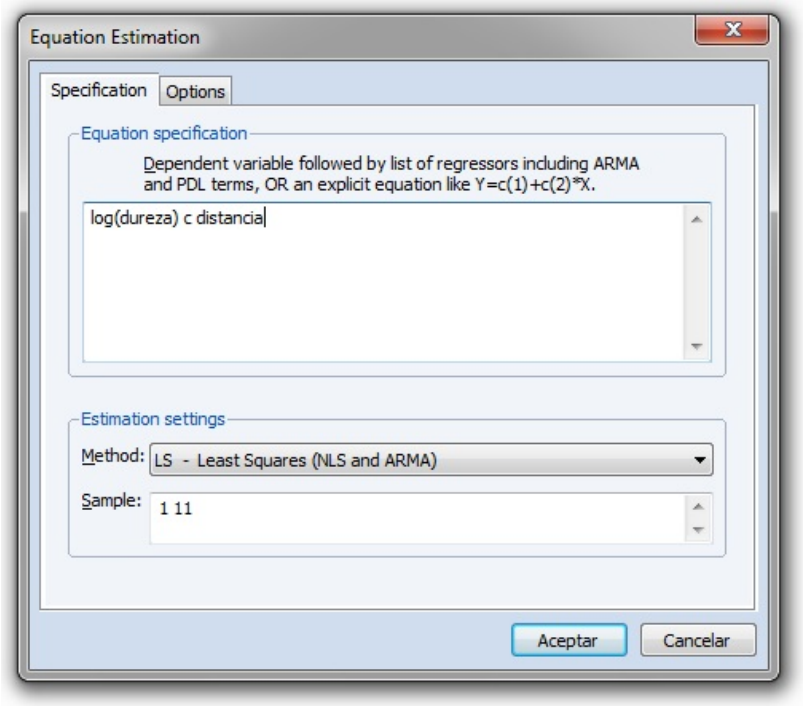
Ahora procederemos a estimar la siguiente relación por OLS:

$$\ln(dureza_i) = \beta_0 + \beta_1 distancia + \varepsilon_i$$

¿Cómo lo hacemos? En la pestaña **Object - New Object**, elegimos el objeto equation y le ponemos un nombre (en este caso ecuacion1):

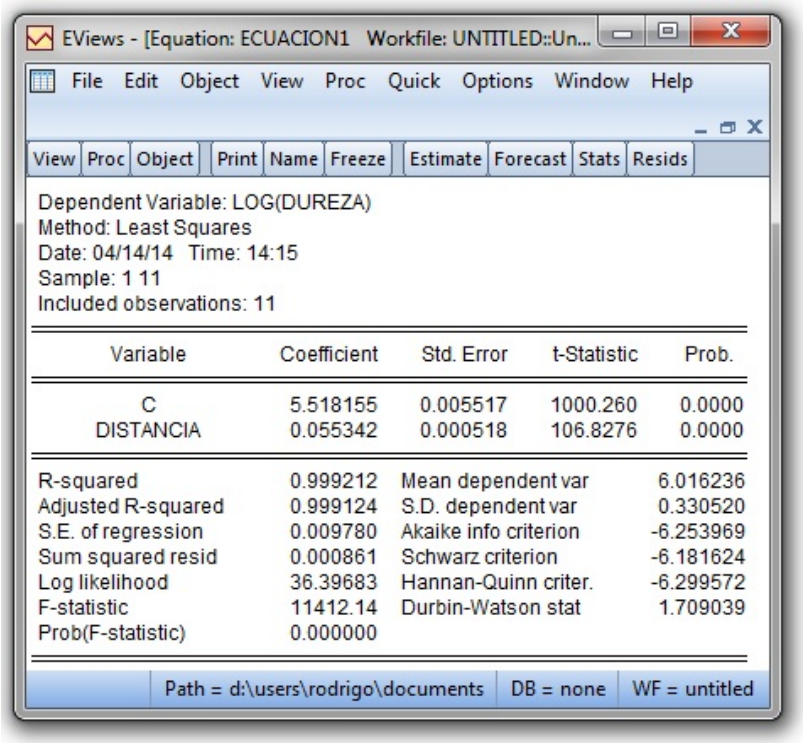


Eligiendo esta opción aparece lo siguiente:



Para escribir la ecuación que se plantea regresionar debemos poner las variables, añadiendo la letra c si deseamos poner una constante.

Al hacer esto se crea el objeto ecuacion1, que contiene el resultado de la regresión:



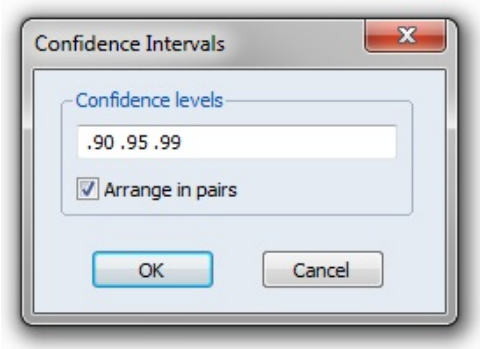
En el output de la regresión, obtenemos los estimadores OLS, los errores estándar de los estimadores, el estadístico del test t (y su respectivo p-value), el R^2 y R^2 ajustado, el test F de significancia conjunta y los criterios de información del modelo regresionado (Akaike o AIC, Schwarz o BIC y Hannan-Quinn o HQ).

2.3. Tests de Hipótesis asociados a los Coeficientes de la Regresión

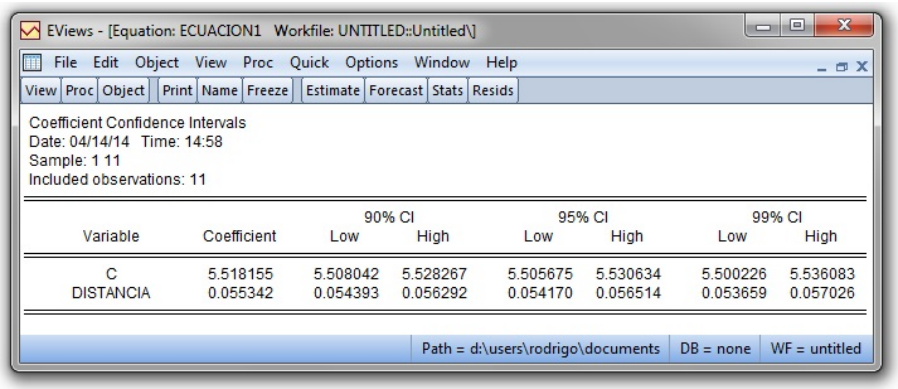
En el output de la regresión podemos elegir las pestañas **View - Coefficient Diagnostics**, donde podremos encontrar los siguientes tests (solo revisaremos los tests relevantes):

2.3.1. Confidence Intervals

Esta opción encuentra los intervalos de confianza de los parámetros. Por ejemplo, si deseamos encontrar los intervalos de confianza al 90 %, 95 % y 99 %, debemos escribir lo siguiente en la pantalla que aparece al seleccionar este test:



Y aparecerá los siguientes resultados, para ambos coeficientes:



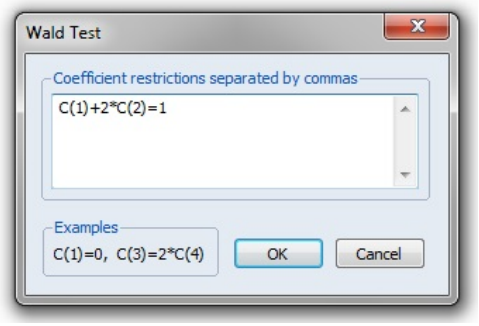
		90% CI		95% CI		99% CI	
Variable	Coefficient	Low	High	Low	High	Low	High
C	5.518155	5.508042	5.528267	5.505675	5.530634	5.500226	5.536083
DISTANCIA	0.055342	0.054393	0.056292	0.054170	0.056514	0.053659	0.057026

2.3.2. Wald Test - Coefficient Restrictions

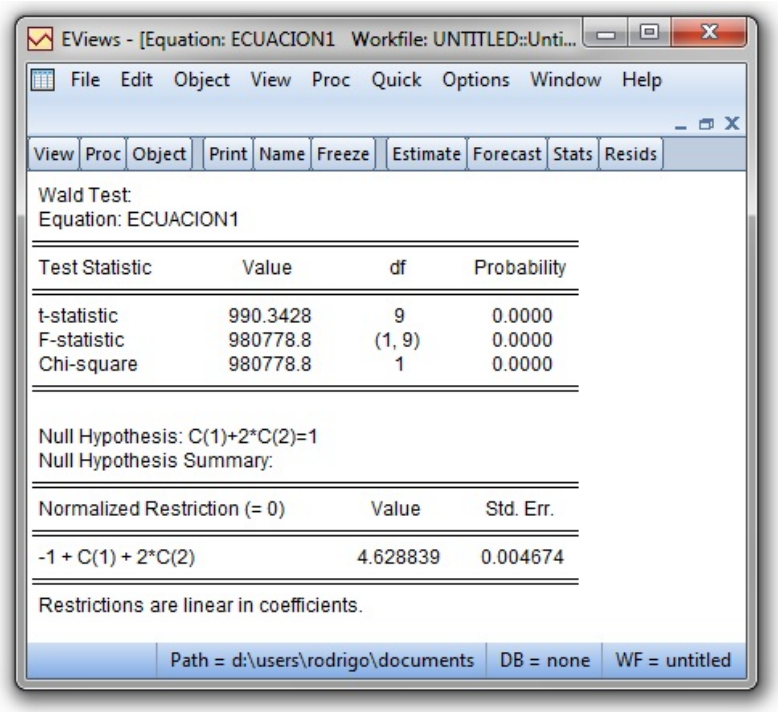
Supongamos que deseamos testear la siguiente hipótesis nula:

$$H_0 : \beta_0 + 2\beta_1 = 1$$

Al seleccionar esta opción aparecerá la siguiente pantalla, que debemos llenar así para testear la hipótesis nula:



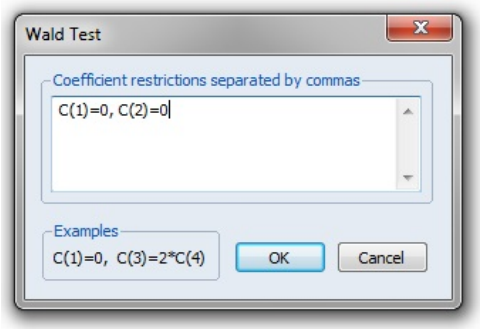
Aparecerá el siguiente resultado:



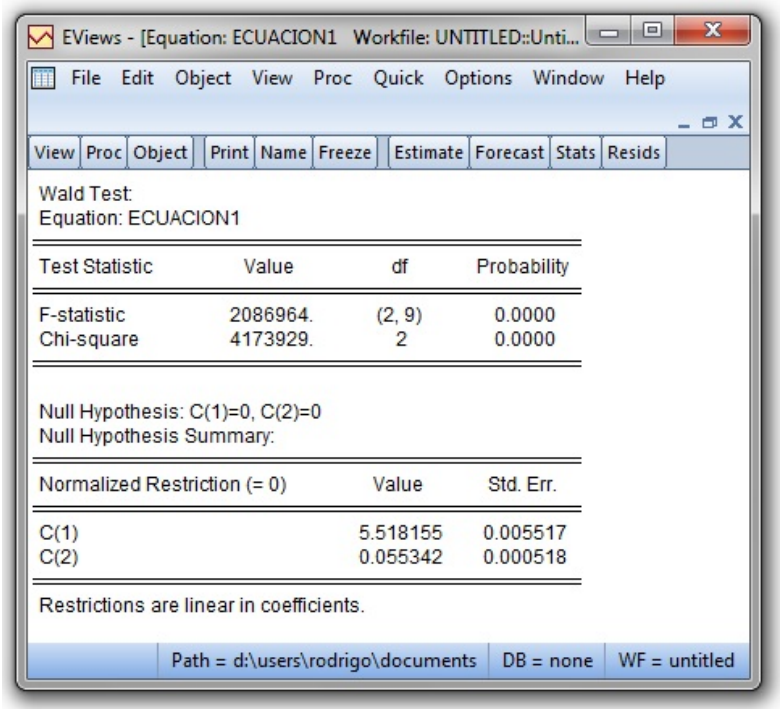
Ahora supongamos que deseamos testear la siguiente hipótesis nula conjunta:

$$H_0 : \beta_0 = 0 \wedge \beta_1 = 0$$

Para hacerlo, elegimos la misma opción que antes, pero de la siguiente forma:



Aparecerá el siguiente resultado:

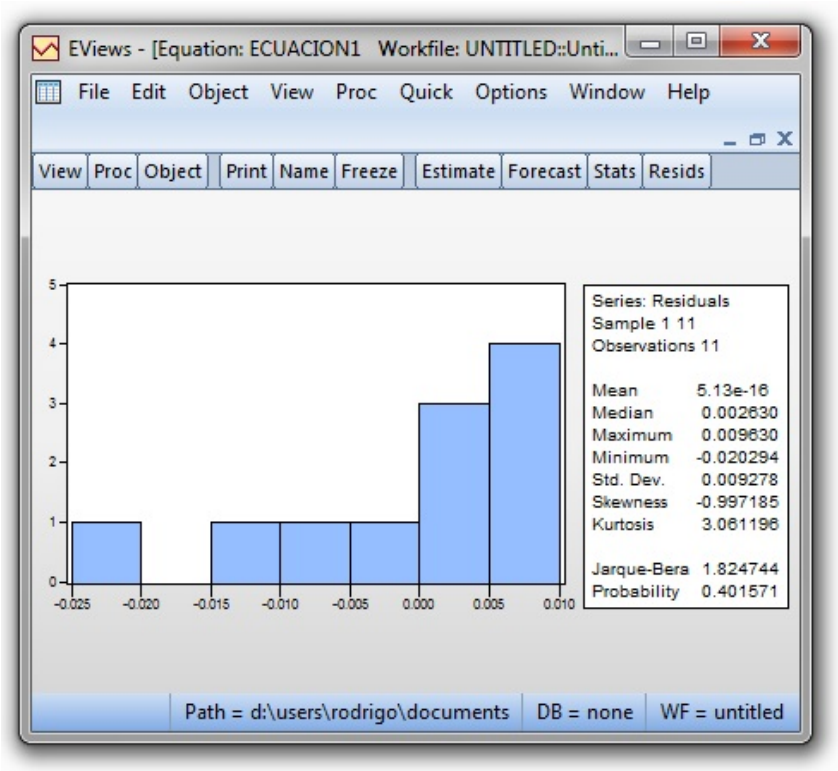


2.4. Tests de Hipótesis asociados a los Residuos de la Regresión

En el output de la regresión podemos elegir las pestañas **View - Residual Diagnostics**, donde podremos encontrar los siguientes tests (solo revisaremos los tests relevantes):

2.4.1. Histogram - Normality Test

Al elegir esta opción aparece lo siguiente:



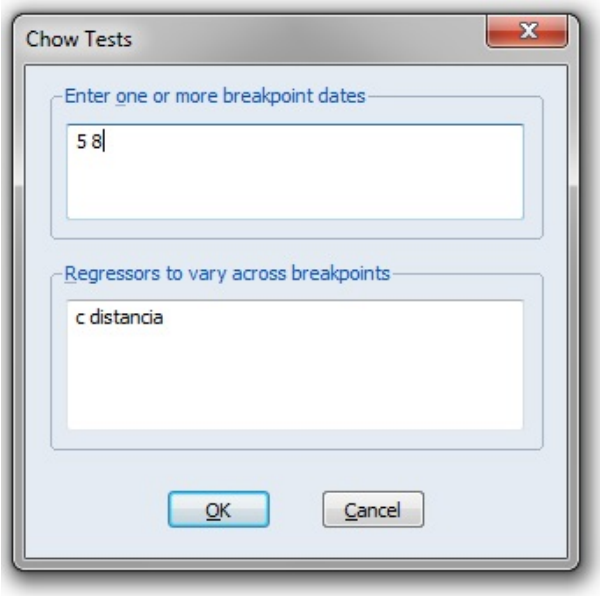
Este test arroja el histograma de los errores (lo que permite ver al ojo si se parece a una distribución normal o no), algunos estadísticos de los errores (media, mediana, máximo, mínimo, desviación estándar, coeficiente de simetría y de curtosis). Tambíen arroja el valor del estadístico y el p-value del test de normalidad Jarque-Bera.

2.5. Tests de Hipótesis asociados a la Estabilidad de la Regresión

En el output de la regresión podemos elegir las pestañas **View - Stability Diagnostics**, donde podremos encontrar los siguientes tests (solo revisaremos los tests relevantes):

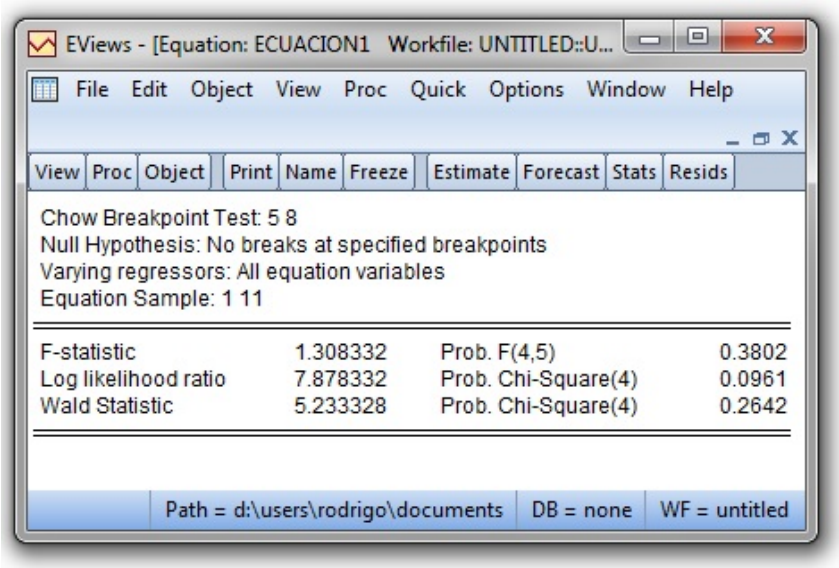
2.5.1. Chow Breakpoint Test

Este opción permite realizar el test de quiebre estructural o de Chow, visto en clases. Al seleccionarlo aparece lo siguiente:



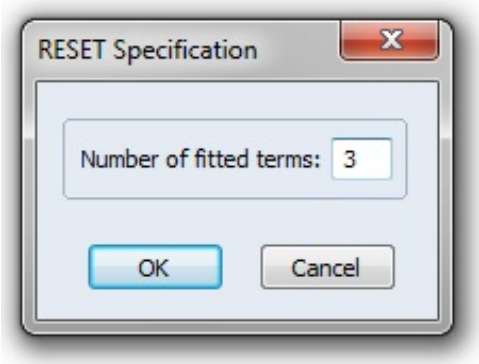
Supongamos que sospechamos que hay quiebre estructural en la muestra 5 y 8, eso lo testeamos introduciendo las observaciones donde hay quiebre estructural.

El resultado será el siguiente:



2.5.2. Ramsey RESET Test

Esta opción nos permite realizar el test de Ramsey, que detecta la omisión de no linealidades de las variables presentes en el modelo. Al elegir esta opción aparece lo siguiente:



Se recomienda no poner un número superior a 3. Poniendo 3, considera variables a lo más elevadas a la cuarta. Pero con números mayores arroja multicolinealidad perfecta. La intuición es que variables elevadas a un número par se parecen mucho a variables al cuadrado, y variables elevadas por un número impar se parecen mucho a variables elevadas al cubo.

Al realizar el test nos arroja los resultados (detallados):

EViews - [Equation: ECUACION1 Workfile: UNTITLED::Unti...

File Edit Object View Proc Quick Options Window Help

View Proc Object Print Name Freeze Estimate Forecast Stats Resids

Ramsey RESET Test
Equation: ECUACION1
Specification: LOG(DUREZA) C DISTANCIA
Omitted Variables: Powers of fitted values from 2 to 4

	Value	df	Probability
F-statistic	2.008283	(3, 6)	0.2144
Likelihood ratio	7.647372	3	0.0539

F-test summary:

	Sum of Sq.	df	Mean Squares
Test SSR	0.000431	3	0.000144
Restricted SSR	0.000861	9	9.57E-05
Unrestricted SSR	0.000430	6	7.16E-05
Unrestricted SSR	0.000430	6	7.16E-05

LR test summary:

	Value	df
Restricted LogL	36.39683	9
Unrestricted LogL	40.22051	6

Unrestricted Test Equation:
Dependent Variable: LOG(DUREZA)
Method: Least Squares
Date: 04/14/14 Time: 15:56
Sample: 1 11
Included observations: 11

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-2115.186	1477.672	-1.431431	0.2023
DISTANCIA	-29.27546	20.34896	-1.438671	0.2003
FITTED^2	130.8328	91.85874	1.424283	0.2042
FITTED^3	-14.33235	10.18887	-1.406667	0.2092
FITTED^4	0.587878	0.423366	1.388579	0.2143

R-squared	0.999607	Mean dependent var	6.016236
Adjusted R-squared	0.999345	S.D. dependent var	0.330520
S.E. of regression	0.008461	Akaike info criterion	-6.403730
Sum squared resid	0.000430	Schwarz criterion	-6.222868
Log likelihood	40.22051	Hannan-Quinn criter.	-6.517738
F-statistic	3813.428	Durbin-Watson stat	2.992606
Prob(F-statistic)	0.000000		

Path = d:\users\rodrigo\documents DB = none WF = untitled

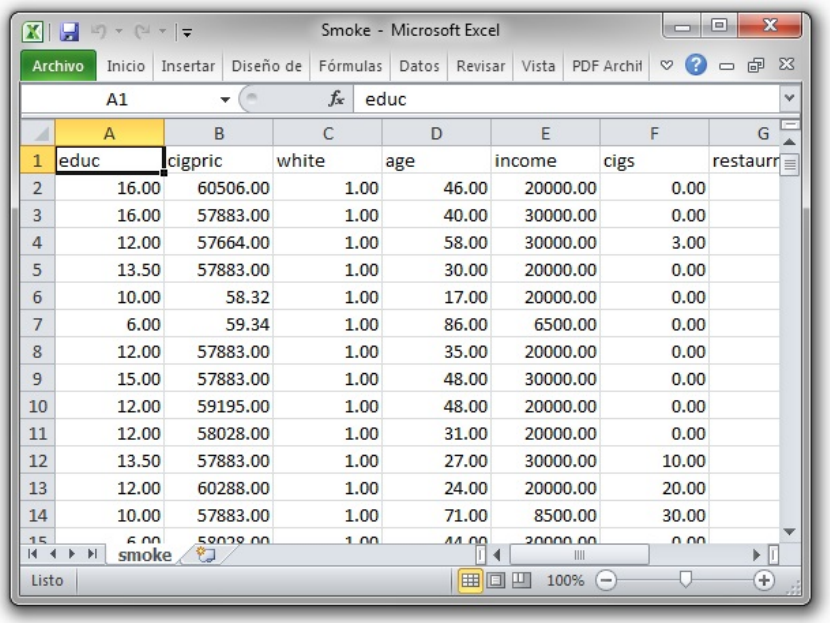
3. Tópicos de Econometría en Eviews

Acá explicaremos prontamente, a medida que avance la materia del curso, de como aplicar contenidos más avanzados (como por ejemplo Heterocedasticidad) en el software Eviews.

3.1. Heterocedasticidad y Autocorrelación

3.1.1. Tests de Heterocedasticidad

Supongamos que deseamos trabajar con datos de la base Smoke.xlsx:

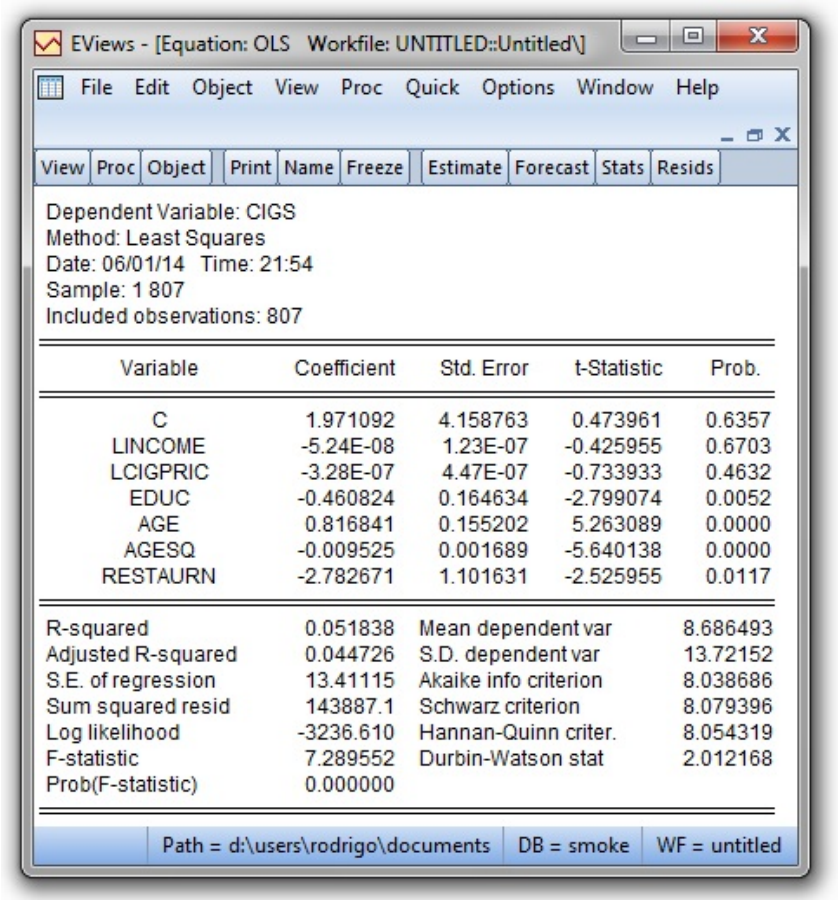


	A	B	C	D	E	F	G
1	educ	cigpric	white	age	income	cigs	restaurr
2	16.00	60506.00	1.00	46.00	20000.00	0.00	
3	16.00	57883.00	1.00	40.00	30000.00	0.00	
4	12.00	57664.00	1.00	58.00	30000.00	3.00	
5	13.50	57883.00	1.00	30.00	20000.00	0.00	
6	10.00	58.32	1.00	17.00	20000.00	0.00	
7	6.00	59.34	1.00	86.00	6500.00	0.00	
8	12.00	57883.00	1.00	35.00	20000.00	0.00	
9	15.00	57883.00	1.00	48.00	30000.00	0.00	
10	12.00	59195.00	1.00	48.00	20000.00	0.00	
11	12.00	58028.00	1.00	31.00	20000.00	0.00	
12	13.50	57883.00	1.00	27.00	30000.00	10.00	
13	12.00	60288.00	1.00	24.00	20000.00	20.00	
14	10.00	57883.00	1.00	71.00	8500.00	30.00	

Y deseamos correr la siguiente regresión por OLS:

$$cigs = \beta_0 + \beta_1 \ln(income) + \beta_2 \ln(cigpric) + \beta_3 educ + \beta_4 age + \beta_5 age^2 + \beta_6 restaurn + u$$

El resultado con Ewiews es el siguiente:



EViews - [Equation: OLS Workfile: UNTITLED::Untitled\]

File Edit Object View Proc Quick Options Window Help

View Proc Object Print Name Freeze Estimate Forecast Stats Resids

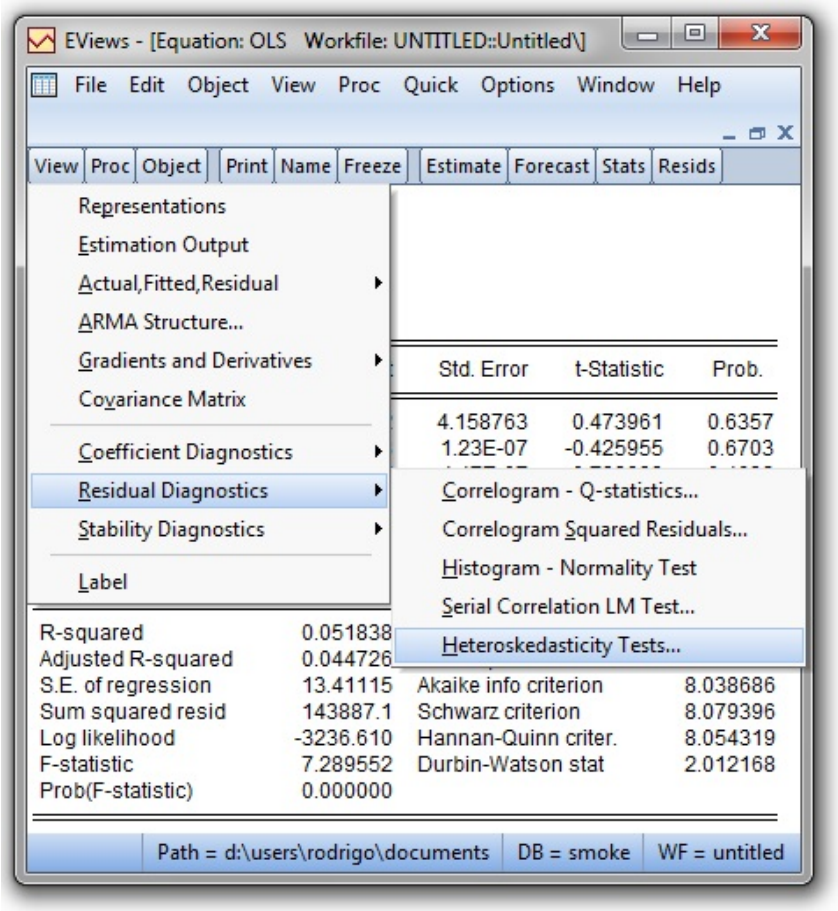
Dependent Variable: CIGS
Method: Least Squares
Date: 06/01/14 Time: 21:54
Sample: 1 807
Included observations: 807

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	1.971092	4.158763	0.473961	0.6357
LINCOME	-5.24E-08	1.23E-07	-0.425955	0.6703
LCIGPRIC	-3.28E-07	4.47E-07	-0.733933	0.4632
EDUC	-0.460824	0.164634	-2.799074	0.0052
AGE	0.816841	0.155202	5.263089	0.0000
AGESQ	-0.009525	0.001689	-5.640138	0.0000
RESTAURN	-2.782671	1.101631	-2.525955	0.0117

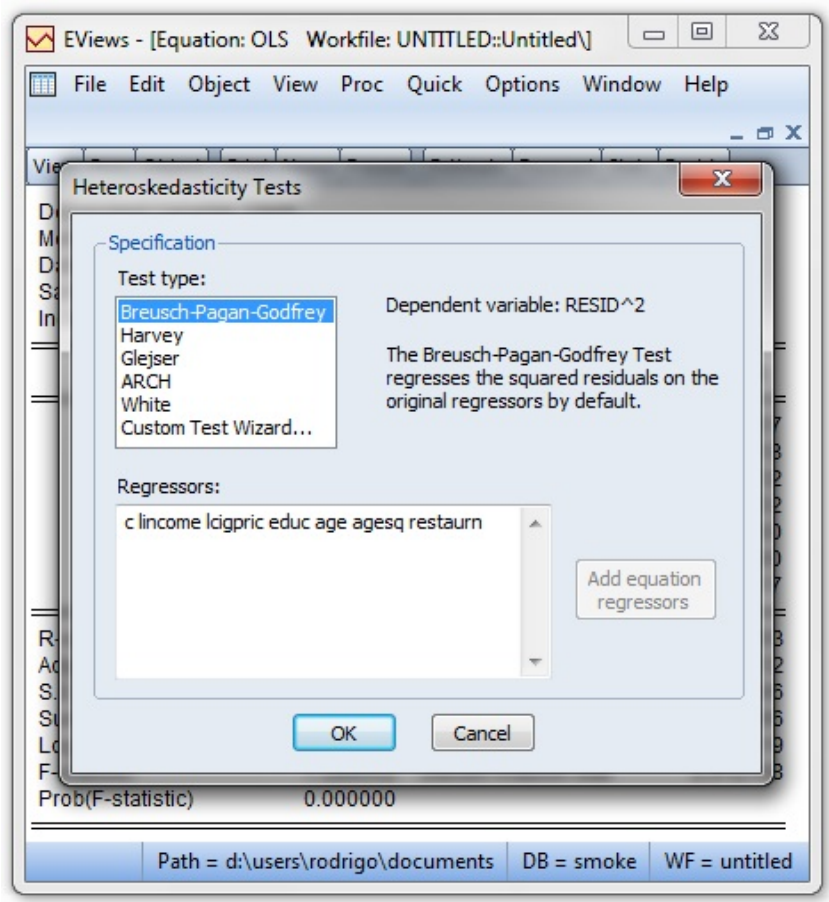
R-squared	0.051838	Mean dependent var	8.686493
Adjusted R-squared	0.044726	S.D. dependent var	13.72152
S.E. of regression	13.41115	Akaike info criterion	8.038686
Sum squared resid	143887.1	Schwarz criterion	8.079396
Log likelihood	-3236.610	Hannan-Quinn criter.	8.054319
F-statistic	7.289552	Durbin-Watson stat	2.012168
Prob(F-statistic)	0.000000		

Path = d:\users\rodrigo\documents DB = smoke WF = untitled

¿Como realizamos los tests de heterocedasticidad? En primer lugar, los tests que realiza Eviews son **Breusch-Pagan**, **Glejser** y **White**. Para hacerlo, en la hoja de resultados seleccionamos *View-Residual Diagnostics-Heteroskedasticity Tests*:



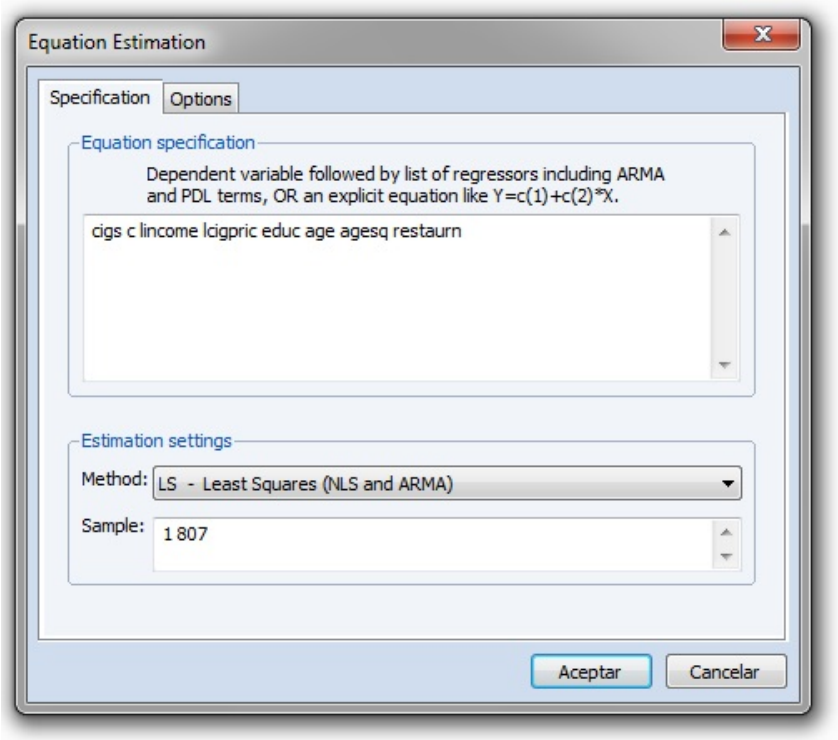
Aparecerá lo siguiente:



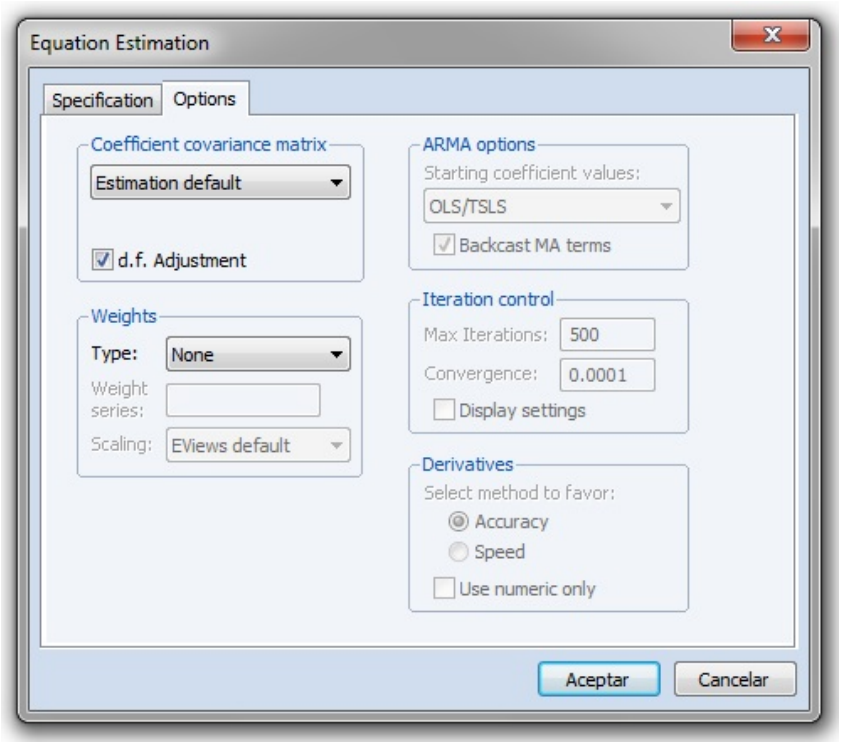
Acá basta con seleccionar el test a realizar y presionar OK, con lo que aparecerá el resultado del test seleccionado.

3.1.2. ¿Cómo corregir heterocedasticidad y autocorrelación usando OLS?

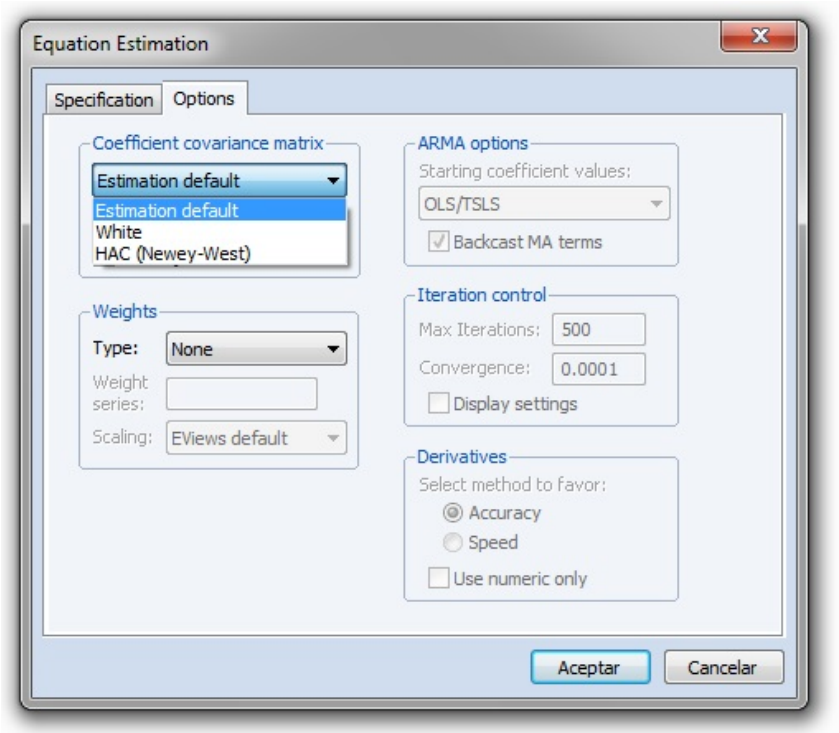
Al igual que en STATA, se permite corregir los errores estándar de los coeficientes estimados ante heterocedasticidad y autocorrelación. Para ello, seleccionemos la pestaña Estimate, y aparecerá lo siguiente:



Elijamos la pestaña Options, y aparecerá lo siguiente:



Seleccionemos la pestaña Estimarion Default del recuadro Coefficient covariance matrix y se desplegará la pestaña:



Entonces se desplegarán 3 opciones:

- **Estimation Default:** estima con OLS tradicional.
- **White:** estima con OLS pero corrigiendo los errores estándar con el estimador de White. Se usa cuando hay heterocedasticidad.
- **HAC (Newey-West):** estima con OLS pero corrigiendo los errores estándar con el estimador de Newey-West. Se usa cuando hay autocorrelación.

Después de elegir la forma de estimación, se presiona aceptar y sale un nuevo resultado de estimación pero considerando la corrección de errores estándar elegida.

4. Programación en Eviews

Supongamos que deseamos realizar remuestreos (simulaciones de Montecarlo, Bootstrap, etc.) con el fin de verificar los resultados asintóticos obtenidos en una regresión con una muestra demasiado pequeña. Supongamos, en particular, que deseamos realizar 1000 experimentos con el método de Montecarlo la siguiente ecuación:

$$\ln(dureza_i) = \beta_0 + \beta_1 distancia + \varepsilon_i$$

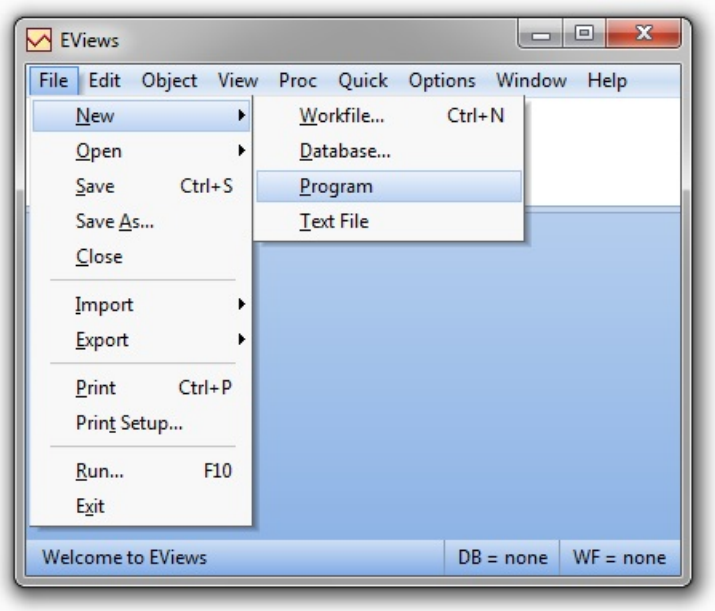
Supongamos que cada experimento generará una muestra aleatoria de tamaño $n = 100$, de la siguiente forma:

- Se generará un vector de errores ε de forma aleatoria, a partir de una distribución $N(0, 1)$.
- Se generará un vector de distancias de forma aleatoria, a partir de una distribución uniforme continua entre 0 y 18.
- Fijaremos valores para β_0 y β_1. En particular, serán los valores obtenidos por OLS en la sección 2.
- Generaremos vectores de $\ln(dureza) = \beta_0 + \beta_1 \cdot distancia + \varepsilon$.
- Regresionaremos $\ln(dureza)$ contra una constante y *distancia*. Para este caso, solo obtendremos los estimadores de la regresión para cada experimento para promediarlos[1].

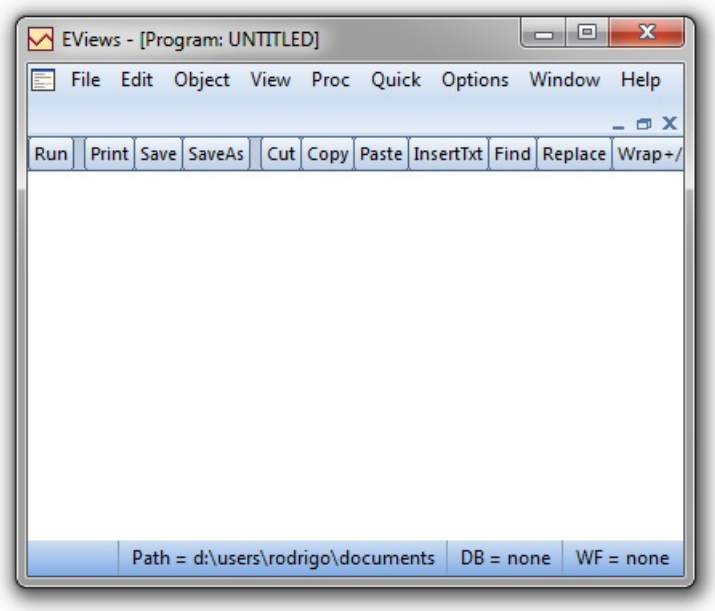
Una opción es realizar mil experimentos de la forma vista en las secciones anteriores. Suponiendo que cada regresión toma 1 minuto, requeriríamos una semana (trabajando 8 horas de corrido) para realizar esto, lo cual no es muy eficiente. Pero la programación en Eviews permite realizar esto de forma automática en pocos segundos.

[1] El promedio de estos estimadores es el estimador obtenido por la simulación, que es consistente a medida que la cantidad de experimentos tiende a infinito.

Para crear un programa, que repita lo que le pidamos 1000 veces, necesitamos crear un programa en blanco así:



Aparecerá el siguiente programa en blanco:



4.1. Código que Realiza el Ejemplo Anterior

Ejemplo realizado en ayudantía 6, del día 9 de Abril de 2014, del curso Econometría I del Prof. Rómulo Chumacero.

Este es el código que permite realizar la simulación antes descrita, para encontrar un estimador de β_1:

Código

```
create u 1 100
scalar alpha=5.52
salar beta=0.06
vector(1000) B=0
for !i = 1 to 1000
genr errores = nrnd
genr explicat = rnd*18
genr y = 0
y=alpha+beta*explicat+errores
equation eq1.ls y c explicat
B(!i) = eq1.@coef(2)
next
```

Y para estimar β_0:

Código

```
create u 1 100
scalar alpha=5.52
salar beta=0.06
vector(1000) B=0
for !i = 1 to 1000
genr errores = nrnd
genr explicat = rnd*18
genr y = 0
y=alpha+beta*explicat+errores
equation eq1.ls y c explicat
B(!i) = eq1.@coef(1)
next
```

4.2. Explicación del Código

Código

```
create u 1 100
```

Este comando crea la base de datos de 100 muestras y de corte transversal (eso lo indica la letra u).

Código

```
scalar alpha=5.52
salar beta=0.06
```

Este comando define los escalares alpha (β_0) y beta (β_1), que utilizará el código como insumo para generar series.

Código

```
vector(1000) B=0
```

Crea un vector B, de 1000 datos, todos ceros. La idea es poner acá los estimadores generados en cada experimento, reemplazando cada cero por el respectivo valor del estimador.

Código

```
for !i = 1 to 1000
next
```

Crea un subprograma, donde itera cada vez, desde 1 a 1000, los comandos que se encuentran entre el for y el next. Next indica en que punto debe pasarse a la siguiente iteración.

Código

```
genr errores = nrnd
genr explicat = rnd*18
genr y = 0
```

Genera variables. En el caso de errores, nrnd genera un vector de tamaño 100 (por el tamaño de muestra que se definió) con números aleatorios a partir de una distribución normal estándar. En el caso de explicat, rnd genera un vector de tamaño 100 con números aleatorios a partir de una distribución uniforme contínua entre 0 y 1, pero al multiplicarla por 18, los genera a partir de una distribución uniforme contínua entre 0 y 18. Para y se genera un vector de ceros.

Código

```
y=alpha+beta*explicat+errores
```

Sustituye los ceros del vector y, por lo señalado en la fórmula.

Código

```
equation eq1.ls y c explicat
```

Genera una ecuación, que regresiona OLS la variable y contra explicat y una constante.

Código

```
B(!i) = eq1.@coef(1)
```

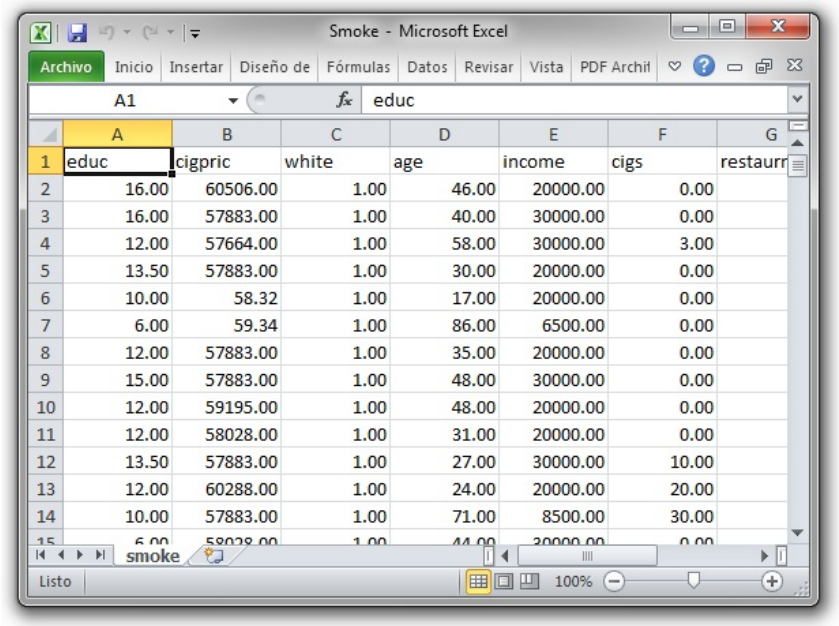
Rellena el elemento $!i$ (que con el for itera de 1 a 1000) del vector B, colocando el primer coeficiente (en este caso, por el modelo señalado, el estimador de la constante) obtenido de la regresión registrada en la ecuación creada anteriormente.

5. Tips Adicionales

5.1. ¿Cómo Cargar Bases de Datos de Forma Eficiente?

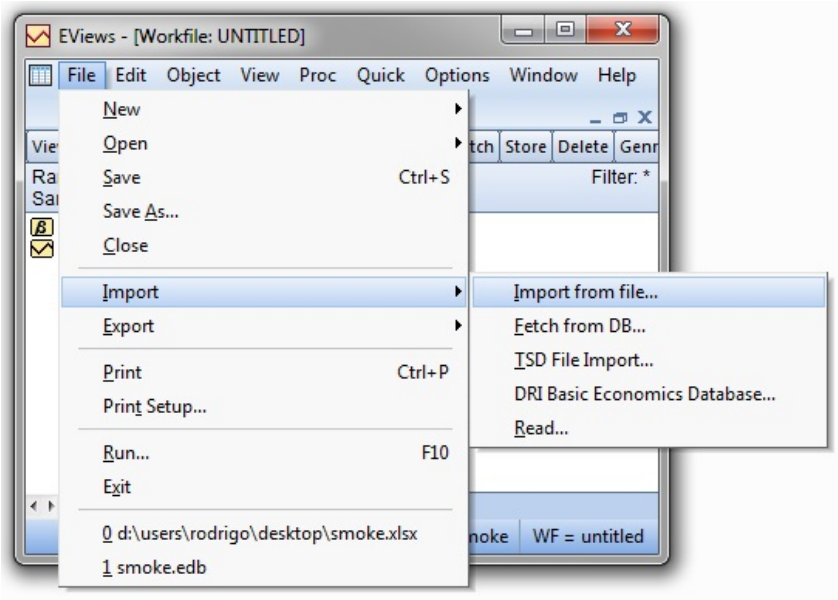
En este apunte se explica como cargar cada serie de una base de datos. Sin embargo, las bases de datos como la CASEN, ELPI o el mismo censo tienen una infinidad de variables que hace engorroso cargar cada base por separado. Acá veremos como cargar una base de datos entera.

Para ello usaremos la base de datos Smoke, que en la primera fila contiene los nombres de las variables:

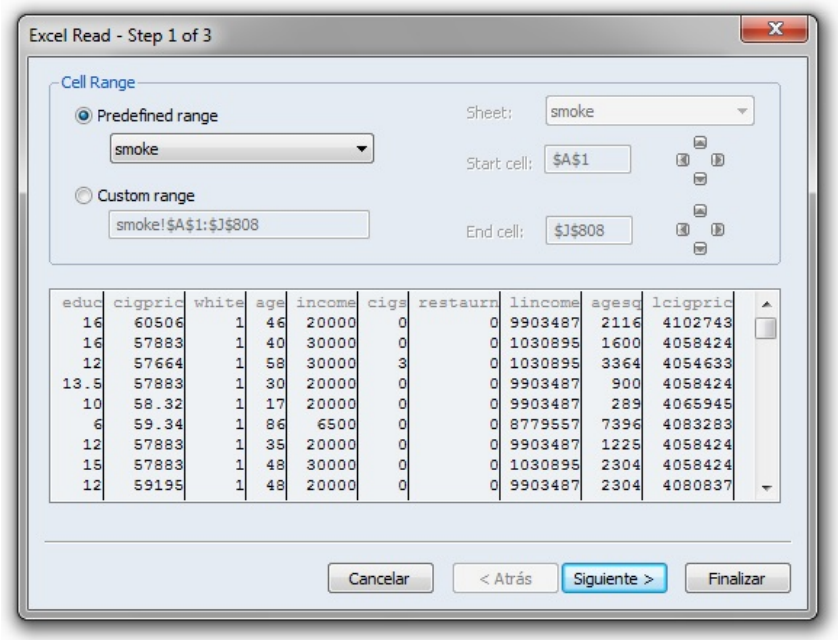


	A	B	C	D	E	F	G
1	educ	cigpric	white	age	income	cigs	restaurr
2	16.00	60506.00	1.00	46.00	20000.00	0.00	
3	16.00	57883.00	1.00	40.00	30000.00	0.00	
4	12.00	57664.00	1.00	58.00	30000.00	3.00	
5	13.50	57883.00	1.00	30.00	20000.00	0.00	
6	10.00	58.32	1.00	17.00	20000.00	0.00	
7	6.00	59.34	1.00	86.00	6500.00	0.00	
8	12.00	57883.00	1.00	35.00	20000.00	0.00	
9	15.00	57883.00	1.00	48.00	30000.00	0.00	
10	12.00	59195.00	1.00	48.00	20000.00	0.00	
11	12.00	58028.00	1.00	31.00	20000.00	0.00	
12	13.50	57883.00	1.00	27.00	30000.00	10.00	
13	12.00	60288.00	1.00	24.00	20000.00	20.00	
14	10.00	57883.00	1.00	71.00	8500.00	30.00	

Ahora, abrimos el software y abrimos un nuevo Workfile, donde definiremos el tamaño de la muestra, más otras cosas si es una base de series de tiempo o de datos de panel. A continuación seleccionamos File-Import-Import from file:

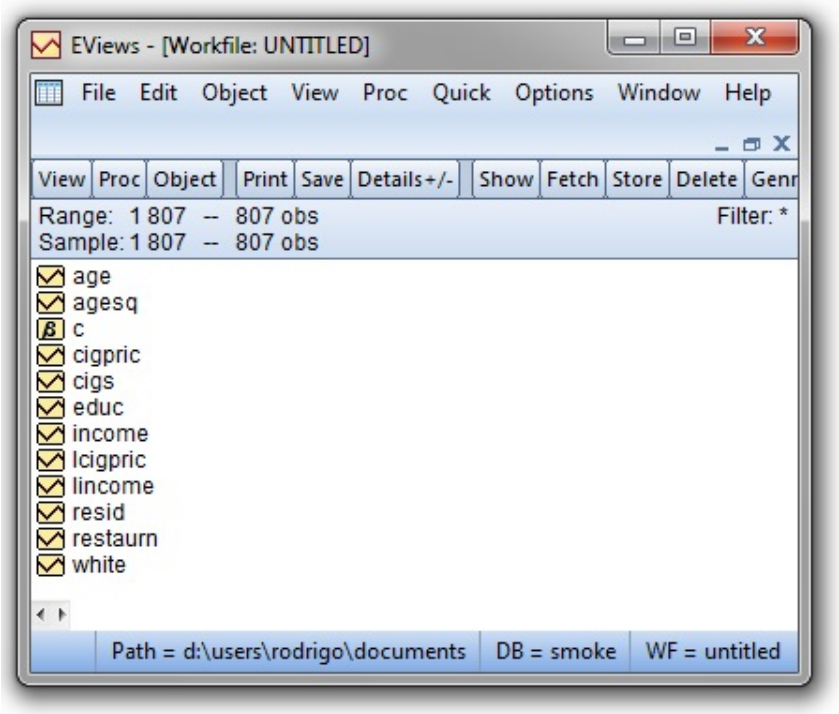


Ahora seleccionamos el archivo Smoke.xlsx que contiene la base de datos. Al seleccionar el archivo aparecerá lo siguiente:



educ	cigpric	white	age	income	cigs	restaurn	lincome	agesq	lcigpric
16	60506	1	46	20000	0	0	9903487	2116	4102743
16	57883	1	40	30000	0	0	1030895	1600	4058424
12	57664	1	58	30000	3	0	1030895	3364	4054633
13.5	57883	1	30	20000	0	0	9903487	900	4058424
10	58.32	1	17	20000	0	0	9903487	289	4065945
6	59.34	1	86	6500	0	0	8779557	7396	4083283
12	57883	1	35	20000	0	0	9903487	1225	4058424
15	57883	1	48	30000	0	0	1030895	2304	4058424
12	59195	1	48	20000	0	0	9903487	2304	4080837

Acá nos indica la forma en como leerá el excel y creara la base de datos. La primera fila, con el nombre de las variables, aparece más clara, lo que significa que lo interpretará como el nombre de las variables. Ahora presione Finalizar y aparecerá cada una de las variables de la base cargadas con su respectivo nombre:



Y listo, ahora podemos empezar a correr las regresiones que queramos.