EViews: Introductory User Guide

Basic Estimation

Learning support material for the courses:

- ✓ **NMST537 Time Series Analysis**
- ✓ **NEKN432 Econometrics**

Based on official EViews Tutorials & EViews Illustrated.

Regression Analysis in EViews

- EViews has a very powerful and easy-to-use estimation toolkit that allows you to estimate from the simplest to the most complex regression analysis.
- This tutorial explains elementary regression techniques in EViews for single equation regressions using cross-section data.
- The main topics include:
  - ✓ Specifying and estimating a regression equation
  - ✓ Equation Objects (saving, labeling, freezing, printing)
  - ✓ Equation Output: Analyzing and Interpreting results
  - ✓ Multiple Regression Analysis
  - ✓ Estimation with Data Expressions and Functions
  - ✓ Post Estimation: Working with Equations
  - ✓ Hypothesis testing
  - ✓ Estimation Options (robust standard errors, weighted least squares)

EViews: Introductory User Guide

BASIC ESTIMATION: DATA

Description of Data File

- ***Part9.wf1*** has the following data on 9,275 individuals*
  - ✓ Wealth – net total financial wealth (in thousands of dollars)
  - ✓ Income – annual income (in thousands of dollars)
  - ✓ Male – dummy variable, equal to 1 if male, 0 otherwise
  - ✓ Married – dummy variable, equal to 1 if married, 0 otherwise
  - ✓ Age – age in years (minimum age in the dataset is 25 years).
  - ✓ Fsize – family size / number of individuals living in the family

***** This data is from Wooldridge, *Introductory Econometrics* (4th Edition).

EViews: Introductory User Guide

BASIC ESTIMATION: *EQUATION* OBJECT (PART I)

EViews Object: *Equation*

- Single equation regression estimation in EViews is performed using the ***Equation Object.***

There are a number of ways to create a simple OLS Equation Object:

1. From the Main menu, select ***Object*** → ***New Object*** → ***Equation***.
2. From the Main menu, select ***Quick*** → ***Estimate Equation***.
3. On the command window type: `ls`

→ In all cases, the ***Equation Estimation*** box appears.

Equation Estimation Box

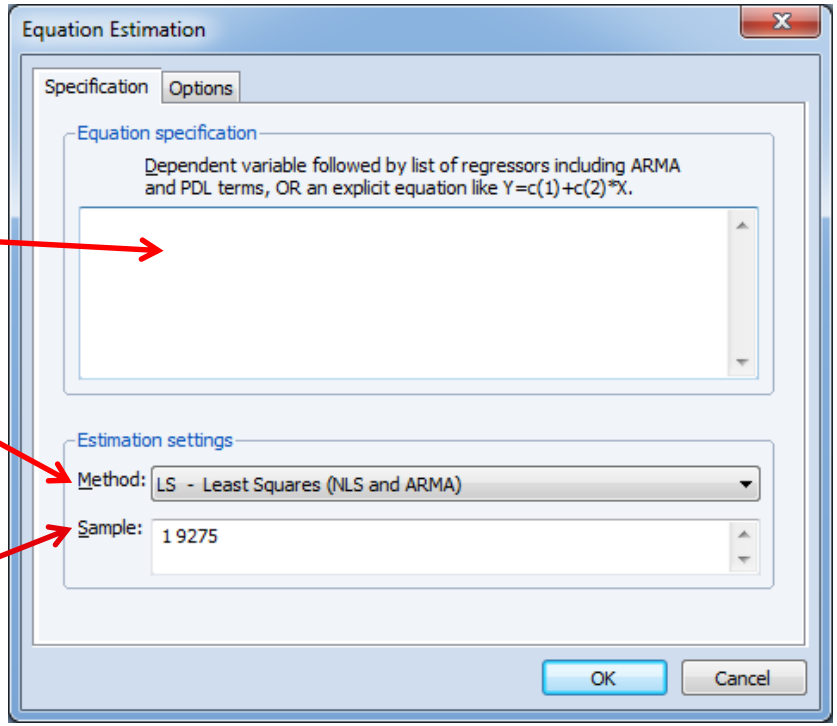
You need to specify three things in this dialogue box:

1. The equation specification.
2. The estimation method.
3. The sample.

Specify your equation either by:
a. List
b. Formula

Specify your estimation method

Specify your sample



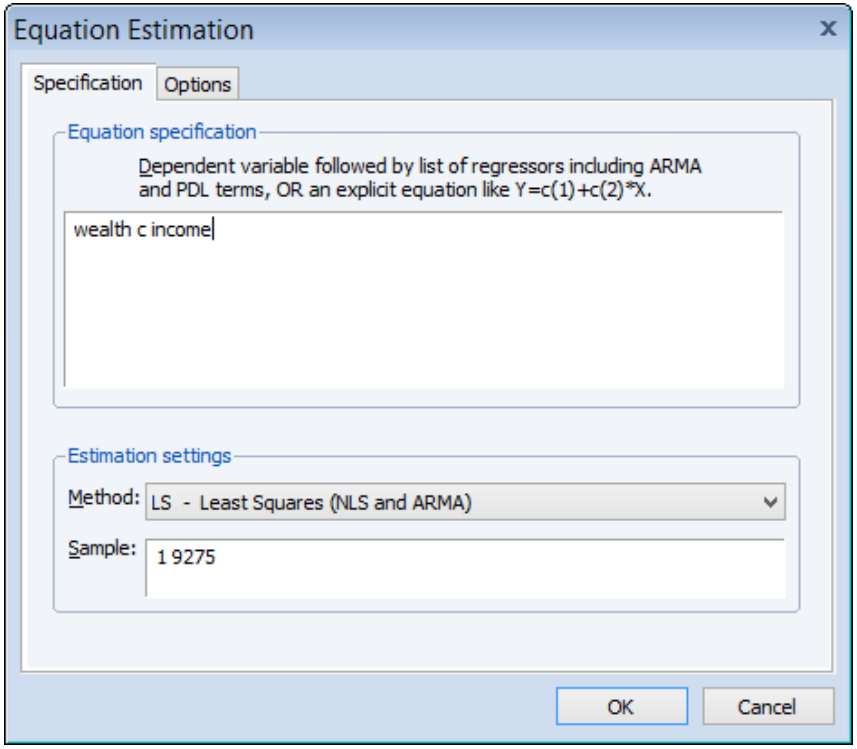
Specifying an Equation by List (Part I)

- The easiest way to specify a linear equation is to provide a list of variables that you wish to use in the equation.

- Suppose that you would like to know how well ***income*** (and an intercept) explains ***financial wealth***.

- To accomplish this, type in the ***Equation Estimation*** box:
  1. The dependent variable (***wealth***).
  2. "c" for constant.
  3. The independent variable (***income***).

  Note that all the entries are all separated by spaces.

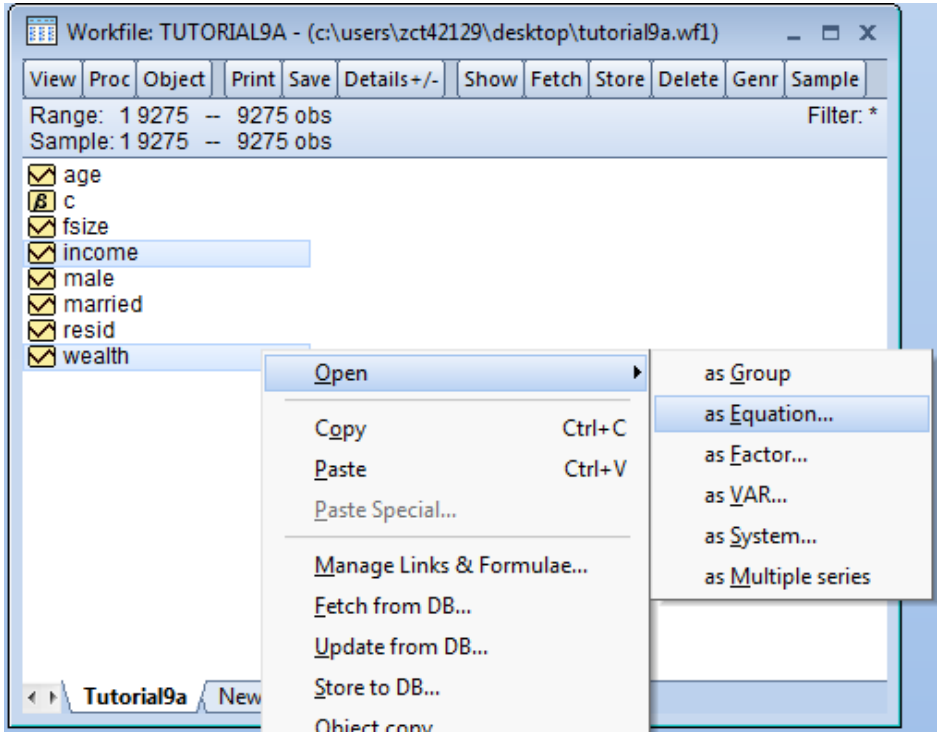


Specifying Equation by List (Part II)

- Alternatively, you can also create an Equation simply by selecting the series and opening them as Equation.

To create an equation:

1. Select ***wealth*** and ***income*** by clicking on these series in the workfile (press **CTRL** to select multiple series). Notice that you need to select the independent variable (***wealth***) first.
2. Right click and select ***Open → as Equation***.
3. The *Equation Estimation* dialog box opens, listing your independent, dependent variables, and the constant.
4. Click **OK** to estimate regression.



Specifying Equation by Formula

- Alternatively, you can also create an Equation simply by the formula.

To specify an equation by the formula:

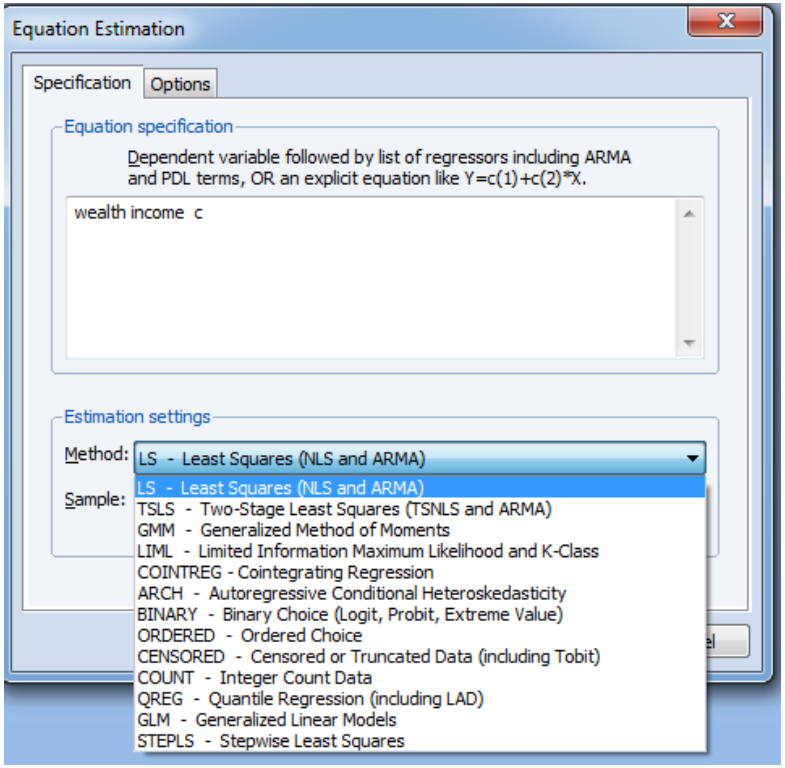
1. Type in the ***Equation Estimation*** box:

   *wealth = c(1) + c(2) * income*,

   where *c(1)* is the first element of the coefficient object *c* (stands for an intercept), *c(2)* is the second element of the coefficient object *c* (stands for the *income* coefficient).
2. Click **OK** to estimate regression.

Estimation Method

- After you specify the variable list, you need to select an estimation method.
- Click on the ***Method*** option, and you see a drop-down menu listing the various estimation method you can perform in EViews.
- Standard, single equation regression, is performed using "***Least Squares***" (***LS***).
- In this tutorial we will use Least Squares and defer discussion of more advanced estimation techniques in *User Guide*.



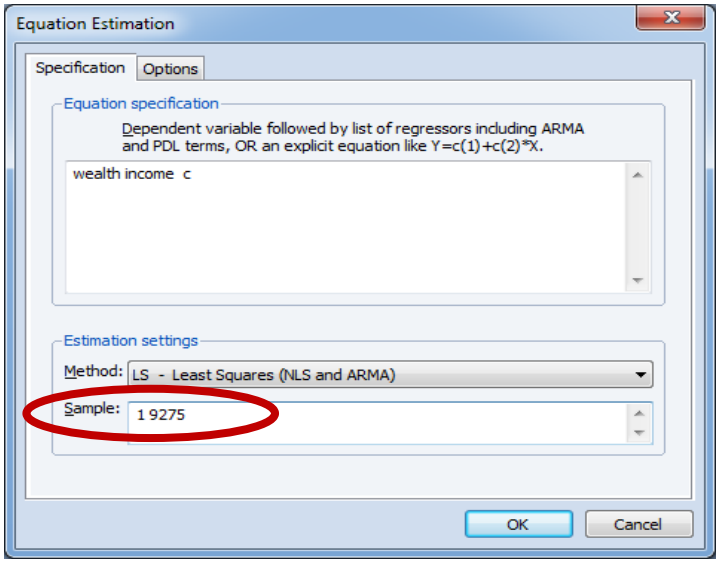
Estimation Sample

- The third item you need to specify in the equation box is the ***Sample***.
- You can specify the sample period in the sample space of the equation box.

For example, to estimate the following regression over the entire sample:

1. You need to include all observations.
2. Click **OK** to estimate the regression.

Alternatively, one can insert any constraints, e.g. the condition "`if married=1 and male=1`".



EViews: Introductory User Guide

BASIC ESTIMATION: *EQUATION* OBJECT (PART II)

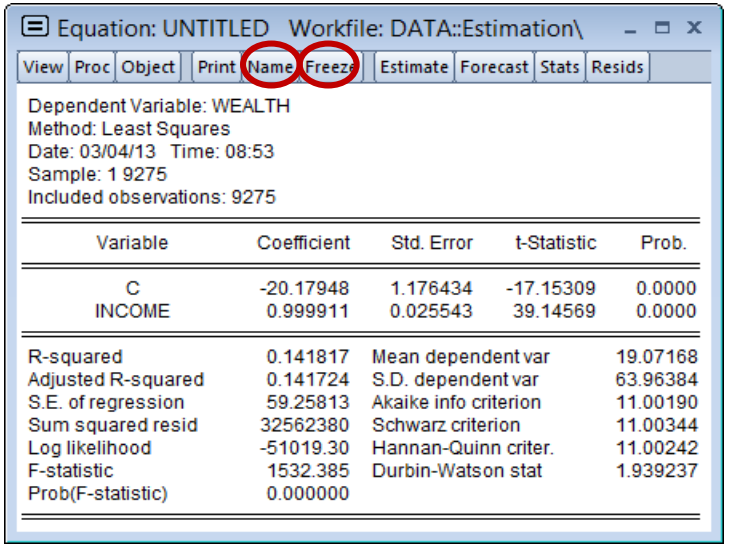
Equation Objects:
Saving, Labeling, Freezing, Printing

- After you estimate an equation, you can **save** the output.

  *[On the Equation box, click the button on the top menu. The **Object Name** dialog box opens. Type the name of the equation and click **OK**.]*

- If you want to **save** the equation output so that **it won't ever change** (even if you re-estimate the regression), you can ***Freeze*** the results.

  [Freezing the equation makes a copy of the current view in the form of a table which is detached from the *Equation Object*. To accomplish this task: Click the ***Freeze*** button on the top menu of the *Equation Box*. You can save this table, by clicking the ***Name*** button in the *Table Object* and name the table in the ***Object Name*** box.]



EViews: Introductory User Guide

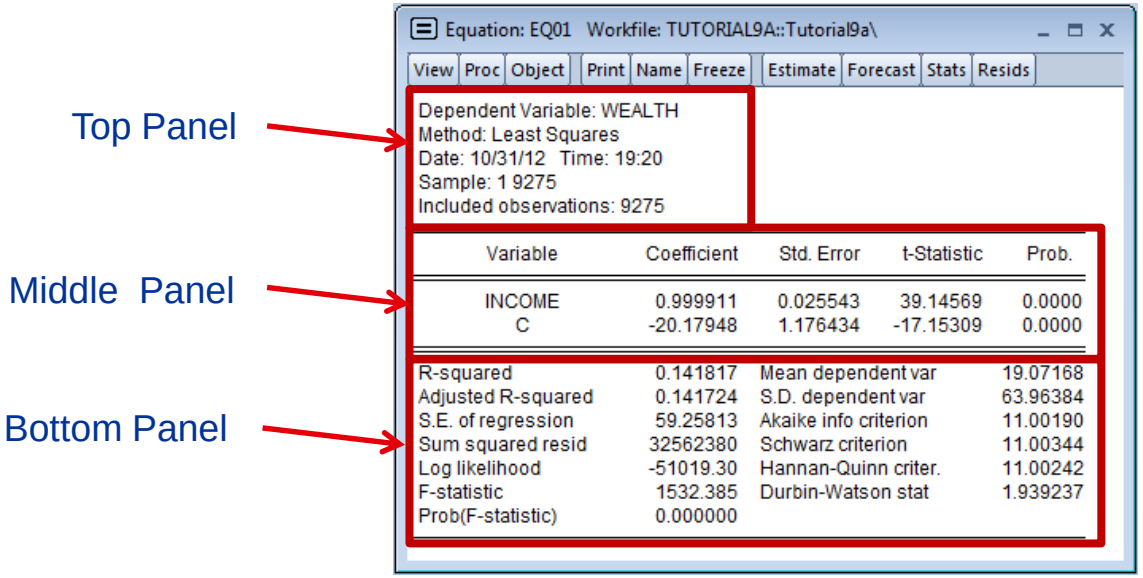
BASIC ESTIMATION: *EQUATION* OUTPUT

Equation Output (Part I)

- Let us analyze the results from our simple estimation, which includes only one explanatory variable (***income***) and an intercept.

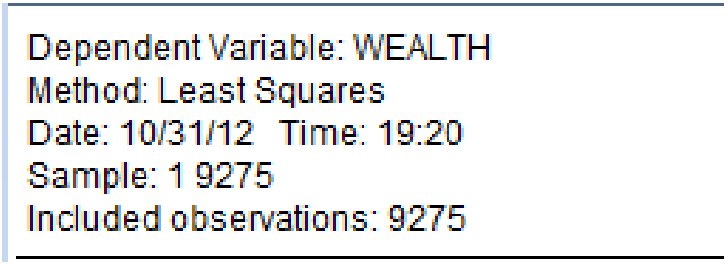
The Equation box has three main parts, which we will discuss in turn:

1. The top panel summarizes the **input** for the regression.
2. The middle panel summarizes information about **regression coefficients**.
3. The bottom panel provides **summary statistics** about the entire regression.



Equation Output (Part II)

- The top panel provides information about **regression inputs**.



Element	Description
Dependent Variable	Denotes the dependent variables.
Method	Denotes the method of estimation (least squares in this example).
Date/Time	Shows the date and time when the regression was carried out.
Sample	Shows the sample period over which the regression is carried out.
Included Observation	Shows the number of observations included in estimation.

Equation Output (Part III)

- The middle panel provides information about the **estimated coefficients**.

Variable	Coefficient	Std. Error	t-Statistic	Prob.
INCOME	0.999911	0.025543	39.14569	0.0000
C	-20.17948	1.176434	-17.15309	0.0000

Element	Description
Coefficient Values	• *Income* coefficient measures the marginal contribution of income to wealth. • C is the estimated constant (or intercept) of the regression.
Standard Errors	• Reports the standard errors of the coefficient estimates. • The larger the standard errors, the more noisy the estimates.
t-Statistic	• Reports the t-statistics, computed by dividing coefficient estimates by their standard errors. • Is used to test whether the coefficient in that row equals zero.
Prob. (p-value)	• Reports probability of drawing a t-statistic as extreme as the one actually estimated. • Is used to test whether the coefficient is equal to zero (against a two-sided alternative).

Equation Output (Part IV)

- The bottom panel provides information regarding the **summary statistics** for the entire regression.

R-squared	0.141817	Mean dependent var	19.07168
Adjusted R-squared	0.141724	S.D. dependent var	63.96384
S.E. of regression	59.25813	Akaike info criterion	11.00190
Sum squared resid	32562380	Schwarz criterion	11.00344
Log likelihood	-51019.30	Hannan-Quinn criter.	11.00242
F-statistic	1532.385	Durbin-Watson stat	1.939237
Prob(F-statistic)	0.000000		

Statistic	Description
R-squared	Measures the success of the regression in predicting the values of depended variable.
Adjusted R-squared	Adjusts for the number of independent regressors by penalizing R-squared for additional regressors.
S.E. of regression	Is a summary measure based on estimated variance of the residuals.
Sum squared resid	Reports the sum of squared residuals. The same as $(S.E.\ of\ regression)^2 * (T-k-1)$, where T is the number of observations (9,275 here), k is the number of independent variables (k=1 here).
Log-likelihood	Reports the log likelihood function evaluated at coefficient estimates assuming normally distributed errors.
F-statistic	Tests whether all slope coefficients (excluding the constant) are zero.
Prob(F-Static)	Reports the probability of drawing an F-statistics as the one estimated.

Equation Output (Part V)

R-squared	0.141817	Mean dependent var	19.07168
Adjusted R-squared	0.141724	S.D. dependent var	63.96384
S.E. of regression	59.25813	Akaike info criterion	11.00190
Sum squared resid	32562380	Schwarz criterion	11.00344
Log likelihood	-51019.30	Hannan-Quinn criter.	11.00242
F-statistic	1532.385	Durbin-Watson stat	1.939237
Prob(F-statistic)	0.000000		

Statistic	Description
Mean dependent var	Shows the mean of the dependent variable (in this case, *wealth*).
S.D. dependent var	Shows the standard deviation of the dependent variable (i.e, *wealth*).
Akaike info criterion	Used in model selection; smaller values are preferred.
Schwarz criterion	An alternative to Akaike information (AIC) used also for model selection. Imposes a larger penalty for including additional explanatory variables
Hannan-Quinn criter.	An alternative to AIC and Schwarz criteria used for model selection. It employs a slightly different penalty function than the other two.
Durbin-Watson stat	Measures serial correlation in the residuals. As a rule of thumb, a DW statistic less than 2 is an indication of positive serial correlation.

EViews: Introductory User Guide

BASIC ESTIMATION: MULTIPLE REGRESSION

Multiple Regression Analysis: Estimation

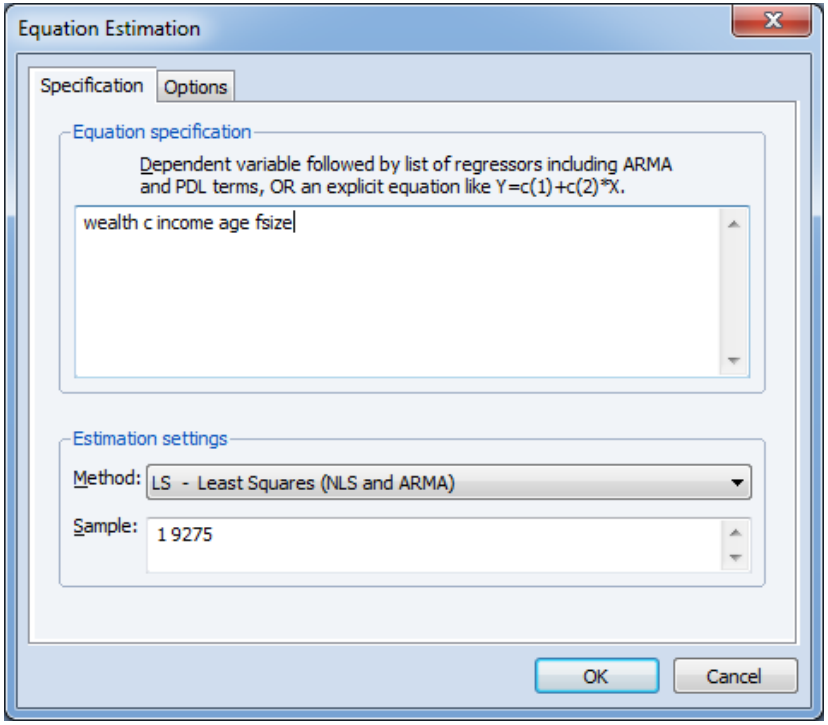
- Let us assume a broader model to explain wealth would be one that includes the age of the individuals as well as the family size, in addition to their income.

For this, augment the previous model with the new independent variables by typing in the equation box:

1. **Wealth** – dependent variable
2. **c** – for constant
3. **Income** – the 1st independent variable
4. **Age** – 2nd independent variable
5. **Fsize** – 3rd independent variable

Alternatively, one might equivalently write:

*wealth = c(1) + c(2)*income + c(3)*age + c(4)*fsize*



Multiple Regression Analysis: Interpreting Results

- The estimation output now includes the ***Coefficient***, ***Standard Error***, ***t-Statistics*** and ***associated p-values*** for each of the regressors in the multiple regression.

Equation: UNTITLED Workfile: DATA::Estimation\

View | Proc | Object | Print | Name | Freeze | Estimate | Forecast | Stats | Resids

Dependent Variable: WEALTH
Method: Least Squares
Date: 03/04/13 Time: 09:07
Sample: 1 9275
Included observations: 9275

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-52.67379	2.829309	-18.61719	0.0000
INCOME	0.973735	0.025377	38.37042	0.0000
AGE	1.013050	0.059022	17.16389	0.0000
FSIZE	-2.805637	0.398602	-7.038697	0.0000

R-squared	0.173474	Mean dependent var	19.07168
Adjusted R-squared	0.173207	S.D. dependent var	63.96384
S.E. of regression	58.16114	Akaike info criterion	10.96474
Sum squared resid	31361180	Schwarz criterion	10.96782
Log likelihood	-50845.00	Hannan-Quinn criter.	10.96579
F-statistic	648.6115	Durbin-Watson stat	1.929874
Prob(F-statistic)	0.000000		

EViews: Introductory User Guide

BASIC ESTIMATION: ESTIMATION VIA EXPRESSIONS

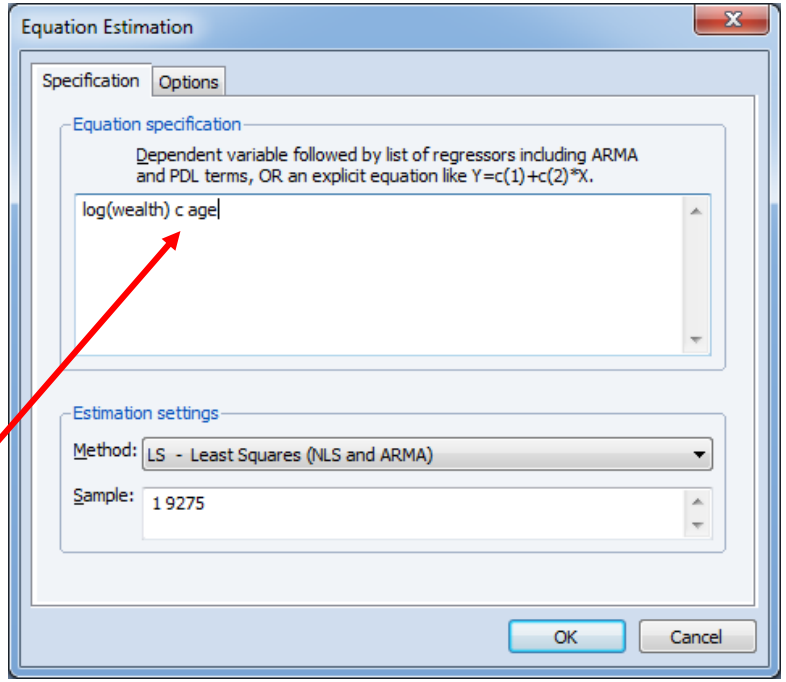
Estimation with Data Expressions: Example 1 (Part I)

- You can use data expressions directly in the equation box to estimate a regression without having to first create these series.
- Often, log or quadratic functions are used to capture nonlinearities in data. These can be specified directly in the equation box.

- For example, a linear regression of ***wealth*** on ***age***, assumes that each additional year increases wealth by a *constant amount*, whether this is the 25th year of your life or the 50th (remember that the minimum age in the data is 25).
- A better characterization would be that each year (beyond the 25th) increases ***wealth*** by a *constant percentage*.
- Define the following regression model:

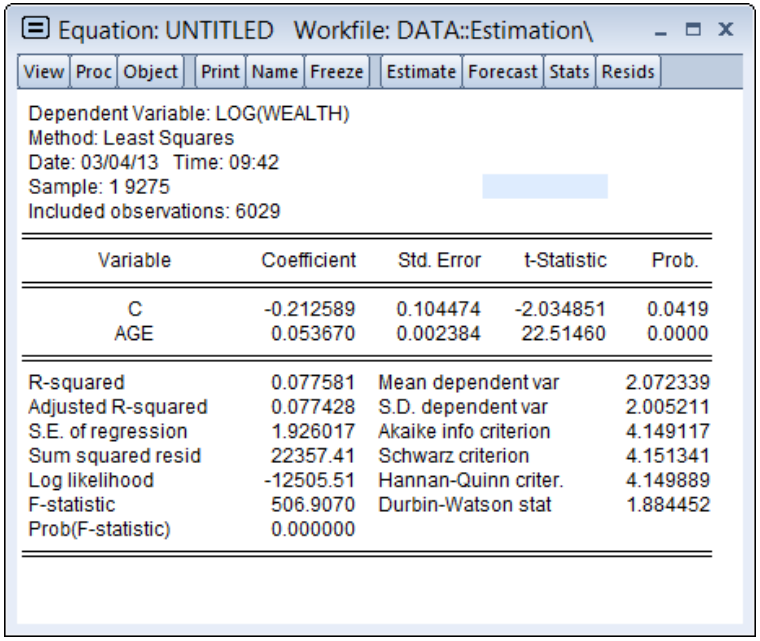
$$\log(wealth_i) = \beta_0 + \beta_1 age_i + \varepsilon_i, i = 1, \dots, 9275,$$
$$i: wealth_i > 0.$$

Note that by default, EViews will automatically convert any observations which cannot be evaluated into NAs, and remove them from the regression.



Estimation with Data Expressions: Example 1 (Part II)

- The output of this estimation is shown here. Note how the **Dependent Variable** label as the top of the output has changed to show we are now estimating with log(wealth) as our dependent variable.
- The **Included observations**: label shows that only 6029 observations were used in the regression. The remaining 3246 were removed due to negative wealth values.



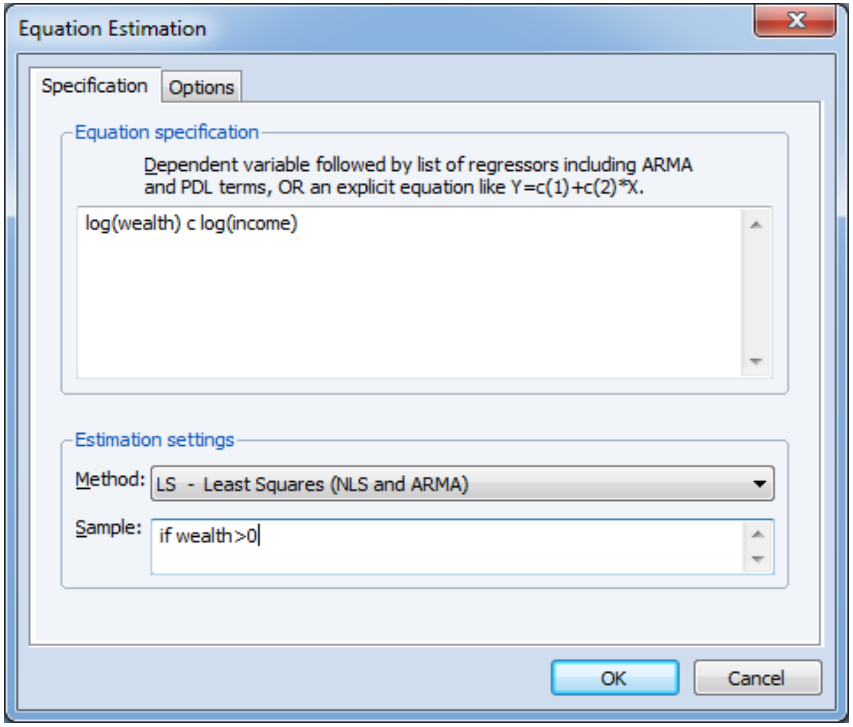
Estimation with Data Expressions: Example 2 (Part I)

- You can use logs for both dependent and independent variables.
- You can estimate a constant-elasticity model relating wealth to income:

$$\log(wealth_i) = \beta_0 + \beta_1 \log(income_i) + \varepsilon_i, i = 1, \ldots, 9275,$$
$$i: wealth_i > 0 \text{ and } income_i > 0.$$

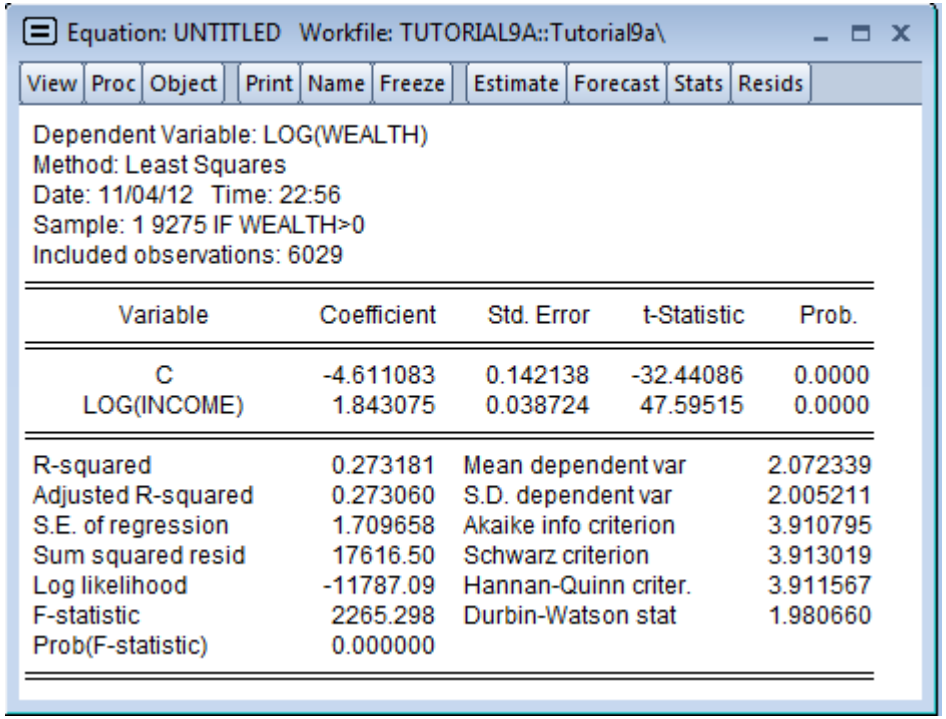
Specify the constant-elasticity model:

1. **log(wealth) –** for dependent variable
2. **c –** for constant
3. **log(income)** – for independent variable
4. This time rather than letting EViews automatically remove non-valid observations, we restrict the sample so that wealth is a positive number. Enter in Sample: **if wealth>0.**



Estimation with Data Expressions: Example 2 (Part II)

- The output is presented here.



Equation: UNTITLED Workfile: TUTORIAL9A::Tutorial9a\

View | Proc | Object | Print | Name | Freeze | Estimate | Forecast | Stats | Resids

Dependent Variable: LOG(WEALTH)
Method: Least Squares
Date: 11/04/12 Time: 22:56
Sample: 1 9275 IF WEALTH>0
Included observations: 6029

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-4.611083	0.142138	-32.44086	0.0000
LOG(INCOME)	1.843075	0.038724	47.59515	0.0000

R-squared	0.273181	Mean dependent var	2.072339
Adjusted R-squared	0.273060	S.D. dependent var	2.005211
S.E. of regression	1.709658	Akaike info criterion	3.910795
Sum squared resid	17616.50	Schwarz criterion	3.913019
Log likelihood	-11787.09	Hannan-Quinn criter.	3.911567
F-statistic	2265.298	Durbin-Watson stat	1.980660
Prob(F-statistic)	0.000000		

Estimation with Dummy Variables (Part I)

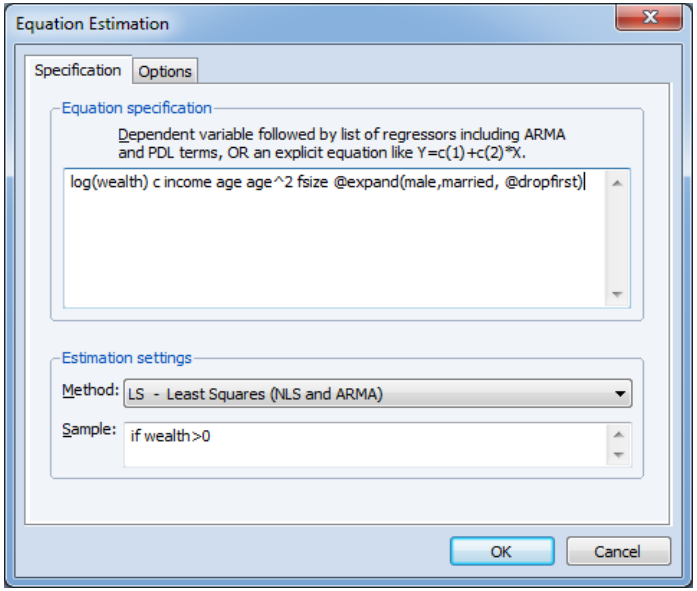
- You can also use ***@expand*** function in a regression to estimate the impact of categorical dummy variables. Let us assume the following example.
- In our dataset: *Male* = 1 if male, 0 if female, *Married* =1 if married, 0 if single
- Suppose you want to find out how wealth depends on the gender of the individual and his/her marital status in addition to income and age, i.e.:

$$\log(wealth_i) = \beta_0 + \beta_1 income_i + \beta_2 age_i + \beta_3 age_i^2 + \beta_4 fsize_i + \alpha_1 1_{[Male=0,Married=1]} + \alpha_2 1_{[Male=1,Married=0]} + \alpha_3 1_{[Male=1,Married=1]} + \varepsilon_i, i = 1, \dots, 9275, i: wealth_i > 0.$$

To determine this, type in the equation box:

1. **wealth** – for dependent variable
2. **c** – for constant
3. **income** – 1st independent variable
4. **age** – 2nd independent variable
5. **@expand(male, married, @dropfirst)**
6. *Sample*: if wealth>0

Note ***@dropfirst*** *explicitly excludes one of the dummy variables.*



Estimation with Dummy Variables (Part II)

- As can be seen, one of the dummy variables corresponding to single females (*male=0* and *married=0*) has dropped out.
- The base group therefore is single females; the rest of the dummy coefficients will be interpreted against this base group.

Equation: UNTITLED Workfile: DATA::Estimation\

View | Proc | Object | Print | Name | Freeze | Estimate | Forecast | Stats | Resids

Dependent Variable: LOG(WEALTH)
Method: Least Squares
Date: 03/04/13 Time: 11:22
Sample: 1 9275 IF WEALTH>0
Included observations: 6029

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-1.788760	0.378241	-4.729154	0.0000
INCOME	0.036971	0.000918	40.25965	0.0000
AGE	0.073948	0.018257	4.050434	0.0001
AGE^2	-0.000318	0.000207	-1.537785	0.1242
FSIZE	-0.151355	0.018949	-7.987434	0.0000
MALE=0,MARRIED=1	0.178583	0.066133	2.700382	0.0069
MALE=1,MARRIED=0	0.048592	0.074845	0.649239	0.5162
MALE=1,MARRIED=1	0.068904	0.108994	0.632180	0.5273

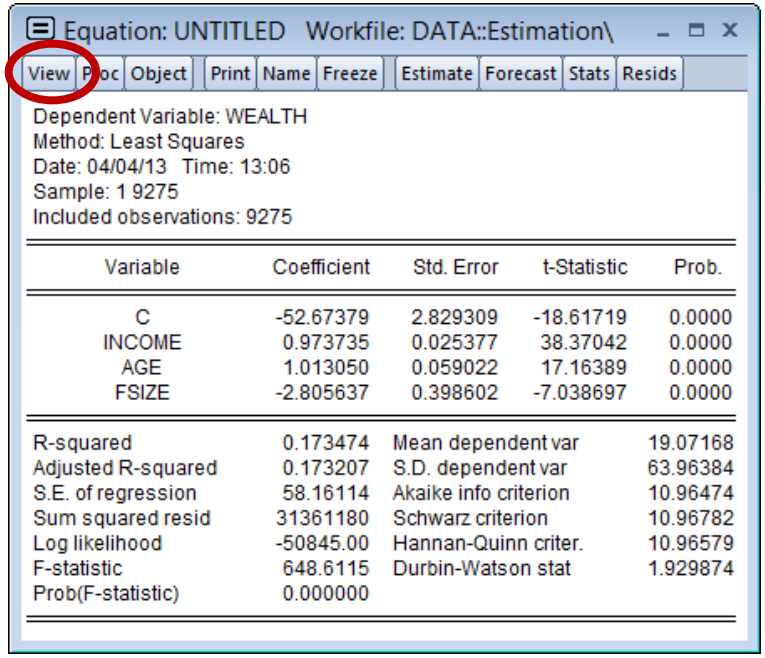
R-squared	0.310209	Mean dependent var	2.072339
Adjusted R-squared	0.309407	S.D. dependent var	2.005211
S.E. of regression	1.666369	Akaike info criterion	3.860497
Sum squared resid	16719.03	Schwarz criterion	3.869393
Log likelihood	-11629.47	Hannan-Quinn criter.	3.863586
F-statistic	386.8181	Durbin-Watson stat	1.972344
Prob(F-statistic)	0.000000		

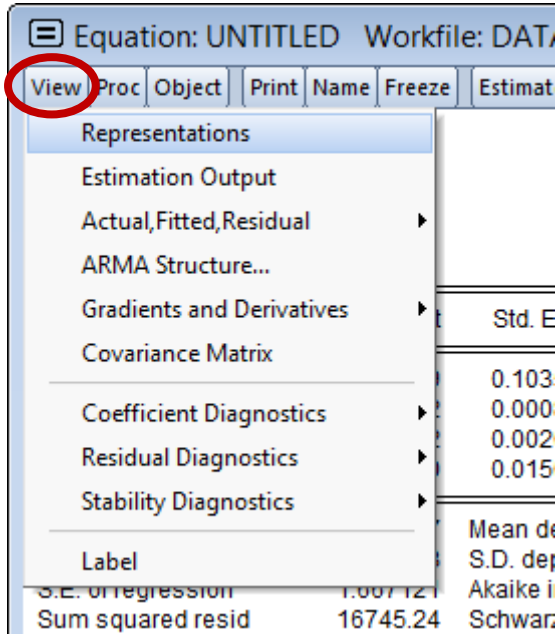
EViews: Introductory User Guide

BASIC ESTIMATION: WORKING WITH OUTPUTS

Post Estimation: *View* Menu

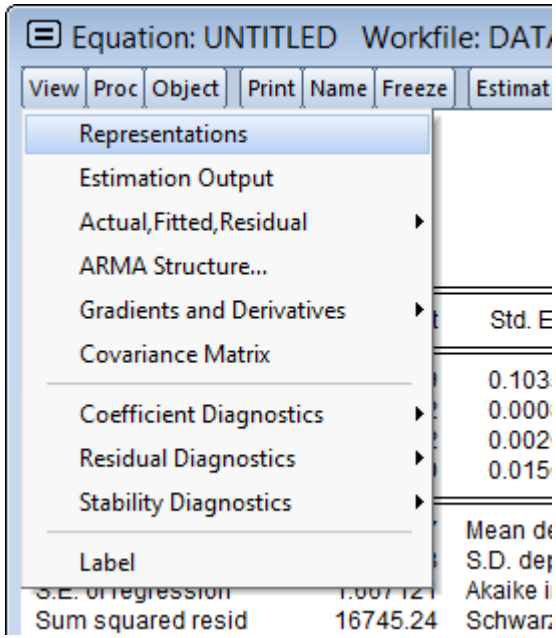
- Once the equation object has been estimated you can perform a number of post-estimation tests, diagnostics and other actions from the ***View*** and ***Proc*** menus.
- Let's first discuss a few main options in the ***View*** menu (others will be discussed in subsequent tutorials):
  - ✓Representation
  - ✓Estimation Output
  - ✓Actual, Fitted, Residual

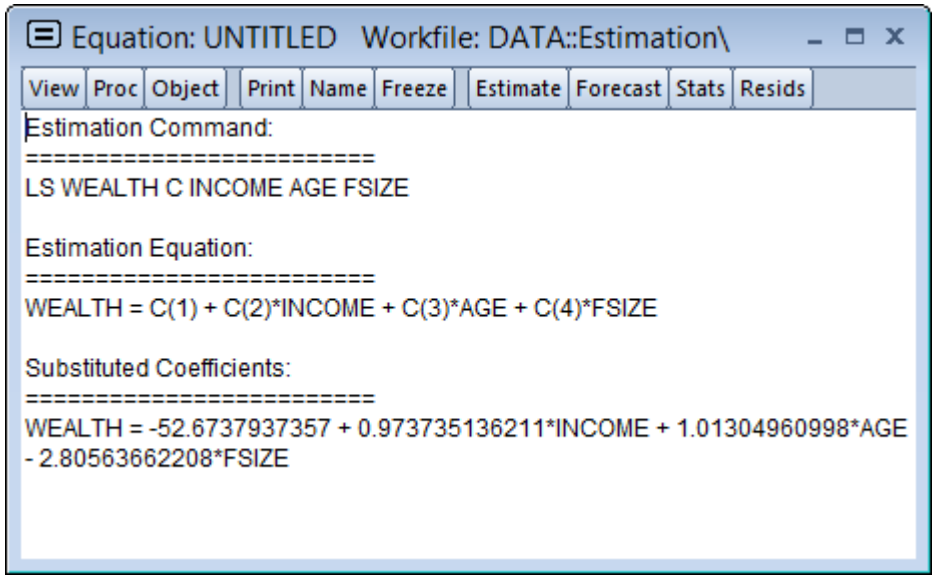




Post Estimation: *View* Menu - Representations

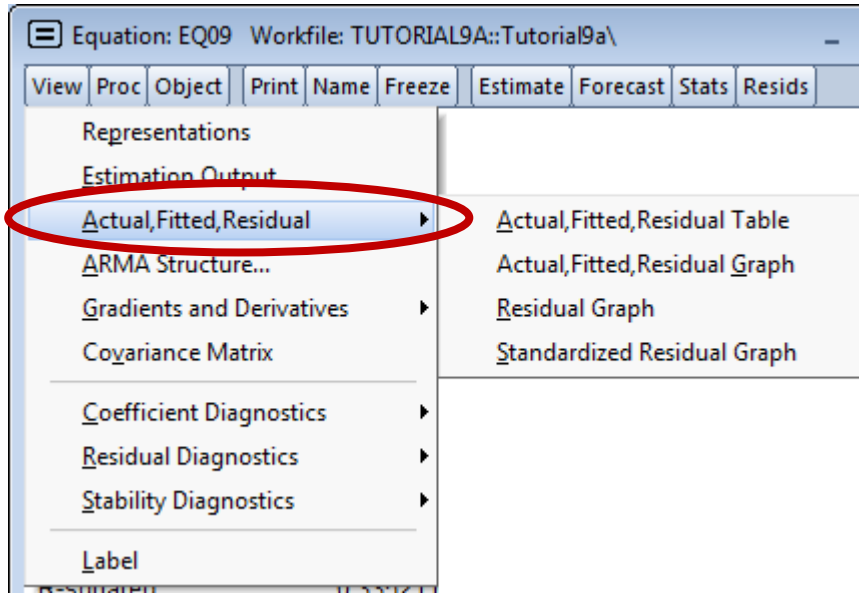
- If you click on ***View*** → ***Representation***, the equation display changes.
- This option displays the equation in three forms:
  - ✓ EViews command form
  - ✓ Algebraic equation with symbolic coefficients
  - ✓ Equation with estimated coefficients





Post Estimation: *View* Menu – Actual, Fitted, Residual (Part I)

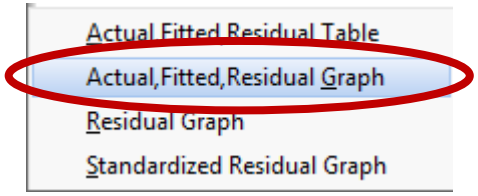
- The *View* Menu, ***Actual, Fitted, Residual*** option, provides several different ways at looking at the residuals and the fitted values of an equation.
- If you click on ***View*** → ***Actual, Fitted, Residual*** a number of options appear:
  - ✓ Actual, Fitted, Residual Table
  - ✓ Actual, Fitted, Residual Graph
  - ✓ Residual Graph
  - ✓ Standardized Residual Graph



Post Estimation: *View* Menu – Actual, Fitted, Residual (Part II)

- For a first look, perhaps it's best to select:

  View → Actual, Fitted, Residual → Actual, Fitted, Residual Graph

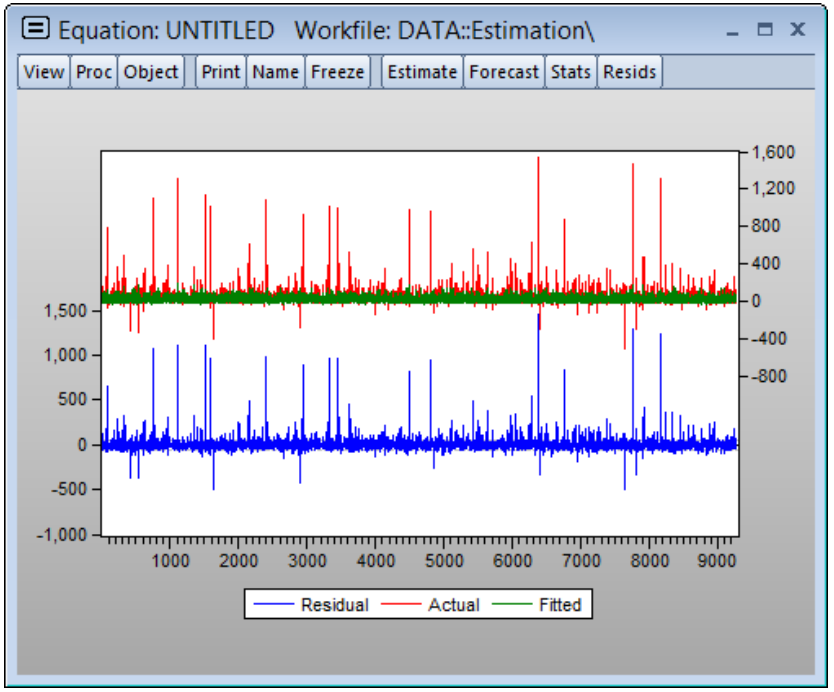


- This displays three series:
  - ✓ The actual series (dependent variable *wealth*) – (in red) plotted against the right vertical axis.
  - ✓ The fitted values ($\widehat{wealth}$) from the regression – (in green) plotted against the right vertical axis.
  - ✓ Residuals – (in blue) plotted against the left vertical axis.

Post Estimation: *View* Menu – Actual, Fitted, Residual (Part III)

- In this case, the fitted values do not approximate the actual values as well as one would hope.
- Similarly, the residuals of the equation are relatively large.

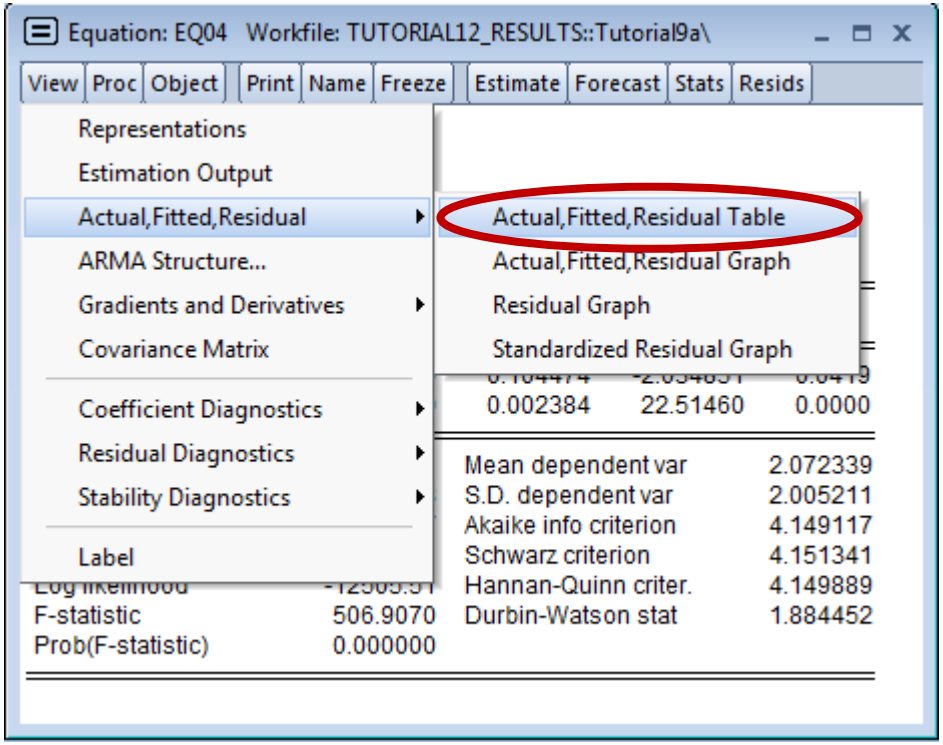
*<u>Note:</u> You can get exactly the same view by clicking the **Resids** button on the top menu of the Equation Object.*



Post Estimation: *View* Menu – Actual, Fitted, Residual (Part IV)

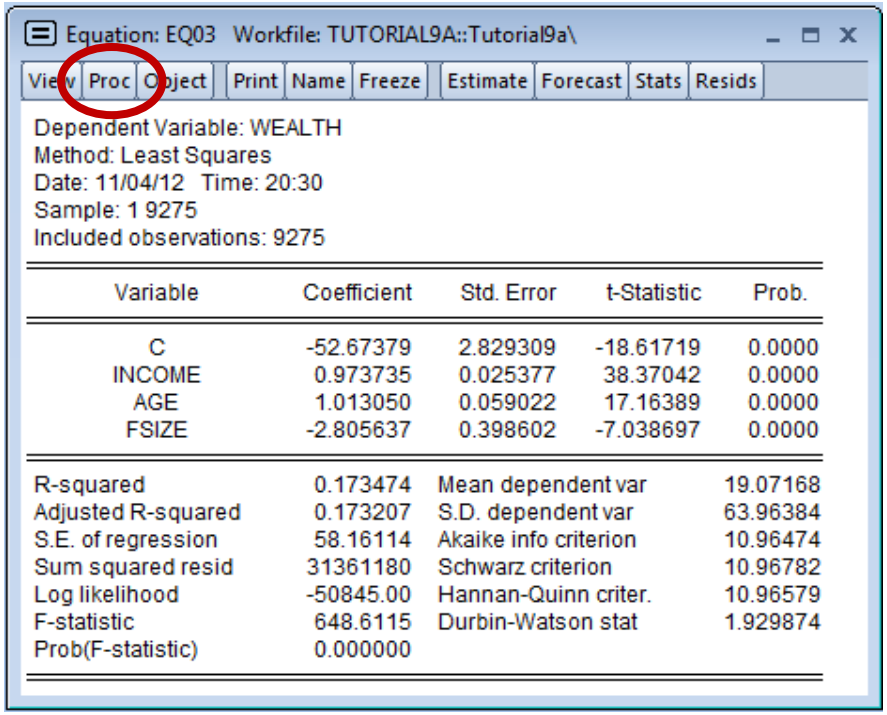
- If you would like to view specific numbers from those graphs, select:

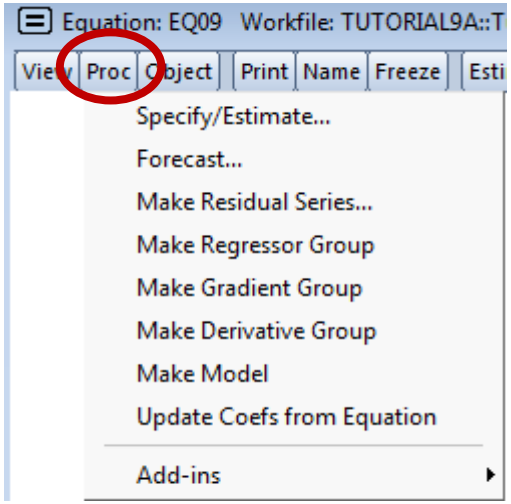
View → Actual, Fitted, Residual → Actual, Fitted, Residual Table



Post Estimation: *Proc* Menu

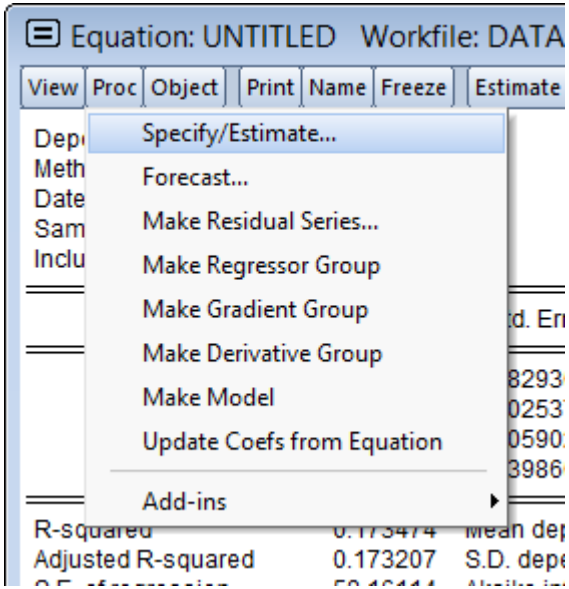
- The ***Proc*** menu, also offers a number of procedures, after estimation is carried out.
- Let's discuss a few main options in the ***Proc*** menu:
  - ✓ Specify/Estimate
  - ✓ Make Regressor Group
  - ✓ Make Residual Series

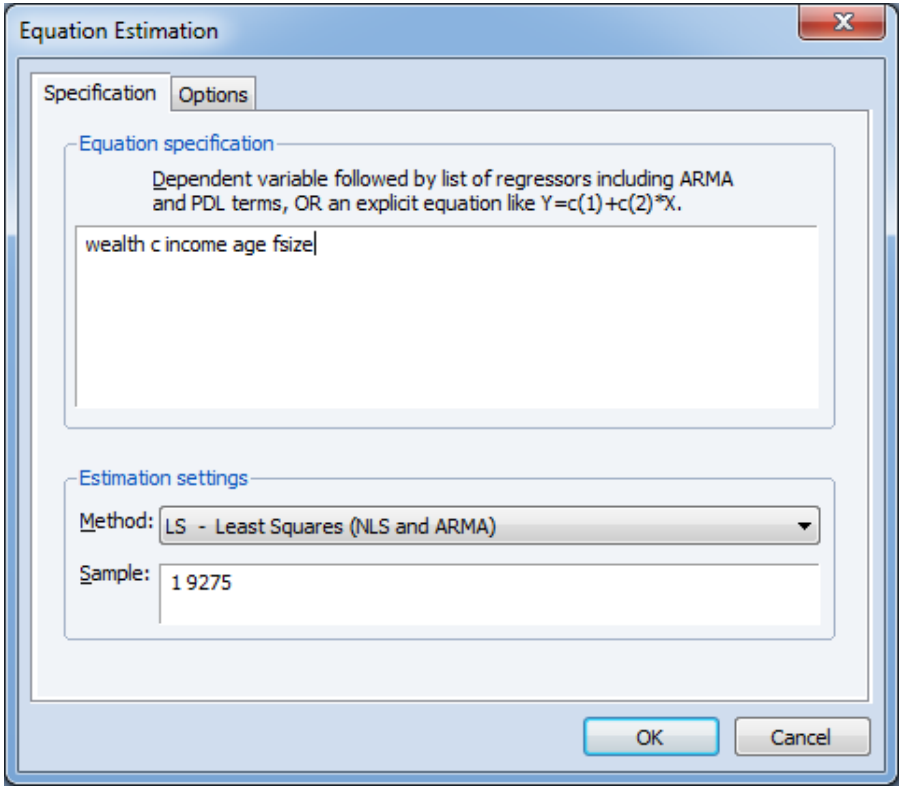




Post Estimation: *Proc* Menu – Specify/Estimate

- If you click on ***Proc*** → ***Specify/Estimate,*** the ***Equation Specification*** dialog box opens up. Here, one can modify your specification using this dialog box (edit equation, change estimation method, change sample, etc.).





Post Estimation: *Proc* Menu – Make Regressor Group

- You can also create a group consisting of all the variables included in the equation (except the constant). To accomplish this, click on ***Proc*** → ***Make Regressor Group***.

Post Estimation: *Proc* Menu – Make Residual Series

- You may also want to store residuals so you can recall them later.
- Every time you estimate an equation, EViews automatically places the residuals of the just-estimated equation in the  ***resid*** series.
- The problem is that this series cannot be used in an estimation command, because the act of estimation itself changes the values stored in ***resid***. New residuals are stored in ***resid*** for every new round of estimation.

<u>If one wants to save these residuals for later use, one can do so by following these steps:</u>

1. Click on **Proc → Make Residual Series.**
2. The **Make Residuals** dialog box opens up. Depending on the estimation you may choose from three types of residuals: ordinary, standardized, and generalized. For ordinary least squares, only the ordinary residuals may be saved.
3. Name the residuals.

EViews: Introductory User Guide

BASIC ESTIMATION: HYPOTHESIS TESTING

Hypothesis Testing

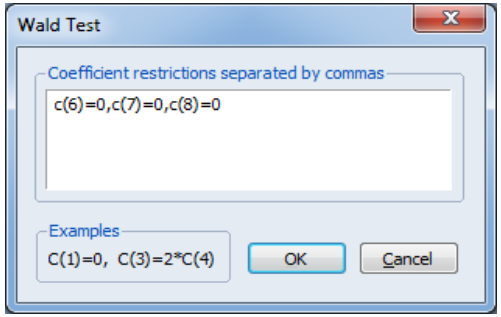
- EViews allows you to test more complex hypothesis just as easy.

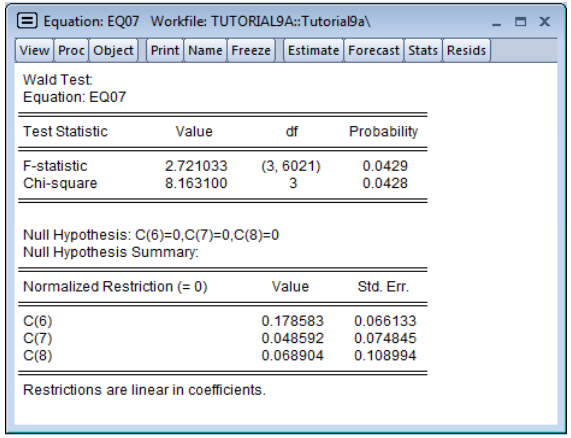
To accomplish this, follow these steps:

1. Click ***View*** → ***Coefficient Diagnostics*** → ***Wald Test – Coefficient Restrictions***
2. The Wald Test dialog box opens up. You will notice:
   - ✓ EViews names coefficients c(1), c(2), c(3), etc., numbering them in the order they appear in the regression (including the constant).
   - ✓ Specify the hypothesis as an equation in the Wald Test box. For example, to test ***H0: c(6)=0, c(7)=0, c(8)=0*** in the previously discussed model:

$$\log(wealth_i) = \beta_0 + \beta_1 income_i + \beta_2 age_i + \beta_3 age_i^2 + \beta_4 fsize_i + \alpha_1 1_{[Male=0,Married=1]} + \alpha_2 1_{[Male=1,Married=0]} + \alpha_3 1_{[Male=1,Married=1]} + \varepsilon_i, i = 1, \dots, 9275, i: wealth_i > 0,$$

   type ***c(6)=0, c(7)=0, c(8)=0***. The F-test is used to test multiple restriction hypothesis.





EViews: Introductory User Guide

BASIC ESTIMATION: HETEROSKEDASTICITY AND ROBUST STANDARD ERRORS

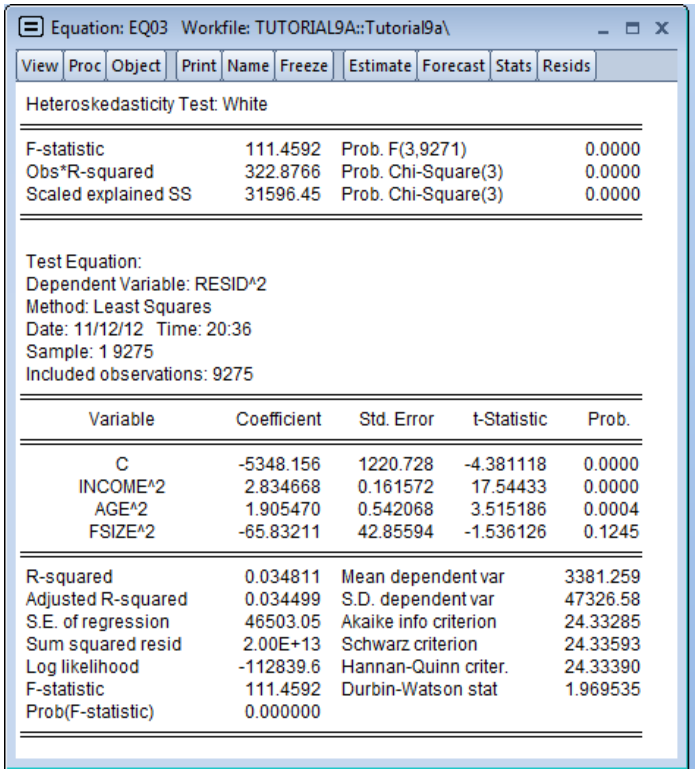
Testing for Heteroskedasticity

- EViews allows you to employ a number of different heteroskedasticity tests.
- We will demonstrate the White heteroskedasticity tests.
  - ✓ The White test is a test of the null hypothesis of no heteroskedasticiy, against heteroskedasticity of unknown, general form. It essentially tests whether the independent variable (and/or their cross terms, x_1^2, x_2^2, x_1*x_2, etc.) help explain the squared residuals.

<u>To perform the White test:</u>

1. On the equation box, select ***View*** → ***Residual Diagnostics*** → ***Heteroskedasticity Tests***.
2. The ***Heteroskedasticity Tests*** window opens up. Select ***White*** under the drop-down menu.
3. You may chose to include or exclude the cross terms. If you do not wish to include the cross term, uncheck the box "*Include White cross terms*". The test will simply be carried out with only the squared terms.

- ❑ The top panel shows the results of the White test, while the bottom panel shows the auxiliary regression used to compute the test statistics.
- ❑ The ***Obs*R-squared*** statistic (in the top panel) is the White statistic with a "x^2" (Chi-Square) distribution.
- ❑ From the Prob. Chi-Square value, we decidedly reject the null of homoskedascity. This means that the error term is heteroskedastic and we should adjust standard errors accordingly.



Equation: EQ03 Workfile: TUTORIAL9A::Tutorial9a\

View | Proc | Object | Print | Name | Freeze | Estimate | Forecast | Stats | Resids

Heteroskedasticity Test: White

F-statistic	111.4592	Prob. F(3,9271)	0.0000
Obs*R-squared	322.8766	Prob. Chi-Square(3)	0.0000
Scaled explained SS	31596.45	Prob. Chi-Square(3)	0.0000

Test Equation:
Dependent Variable: RESID^2
Method: Least Squares
Date: 11/12/12 Time: 20:36
Sample: 1 9275
Included observations: 9275

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-5348.156	1220.728	-4.381118	0.0000
INCOME^2	2.834668	0.161572	17.54433	0.0000
AGE^2	1.905470	0.542068	3.515186	0.0004
FSIZE^2	-65.83211	42.85594	-1.536126	0.1245

R-squared	0.034811	Mean dependent var	3381.259
Adjusted R-squared	0.034499	S.D. dependent var	47326.58
S.E. of regression	46503.05	Akaike info criterion	24.33285
Sum squared resid	2.00E+13	Schwarz criterion	24.33593
Log likelihood	-112839.6	Hannan-Quinn criter.	24.33390
F-statistic	111.4592	Durbin-Watson stat	1.969535
Prob(F-statistic)	0.000000		

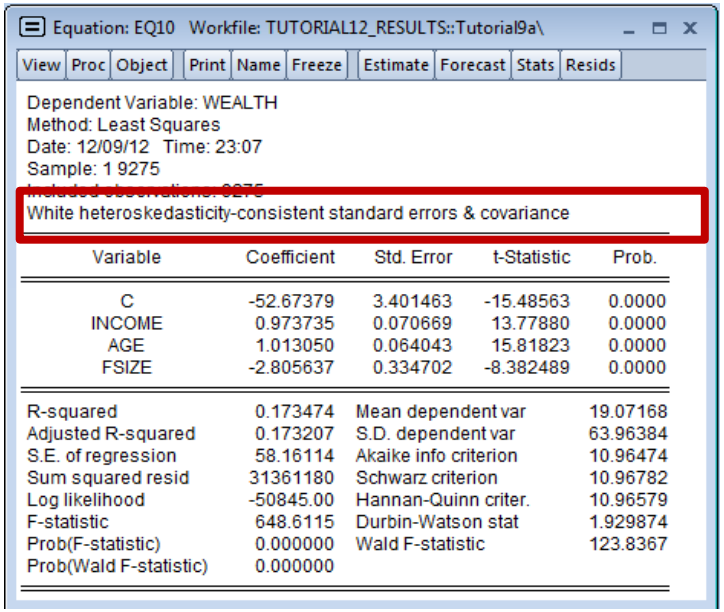
Addressing Heteroskedasticity: Robust Standard Errors

- One approach to dealing with heteroskedasticity is to correct the standard errors to account for heteroskedasticity.
- EViews provides built-in tools that allows you to adjust standard errors for heteroskedasticity of unknown form.

To derive the White-heteroskedasticity consistent standard errors:

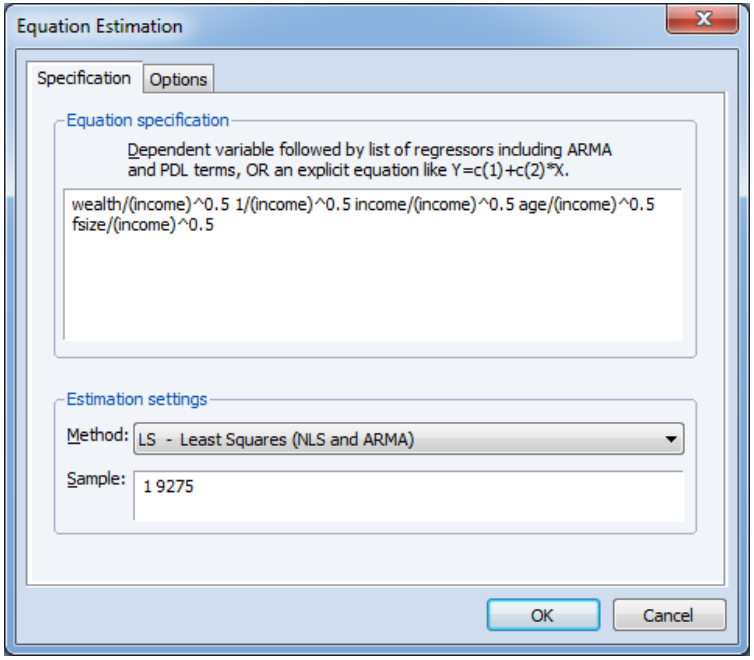
1. Click ***Estimate*** on the ***Equation Box***.
2. The ***Equation Estimation*** box opens up. Click ***Options*.**
3. Under the *Coefficient Covariance matrix* drop-down menu, choose **White**.

❑ EViews re-estimates the equation, this time adjusting the standard errors for heteroskedasticity.



Weighted Least Squares (Part I)

- Suppose that you know the exact nature of the heteroskedasticity.
- For example, suppose you suspect that heteroskedastity is present in the financial wealth regression because the variance of the unobserved factors affecting financial wealth increases with income.
- To express this as an equation: $var(\varepsilon_i|income_i) = income_i * \sigma^2, \sigma^2 > 0$.
- You could transform the model by dividing by $\sqrt{income_i}$ as shown here.

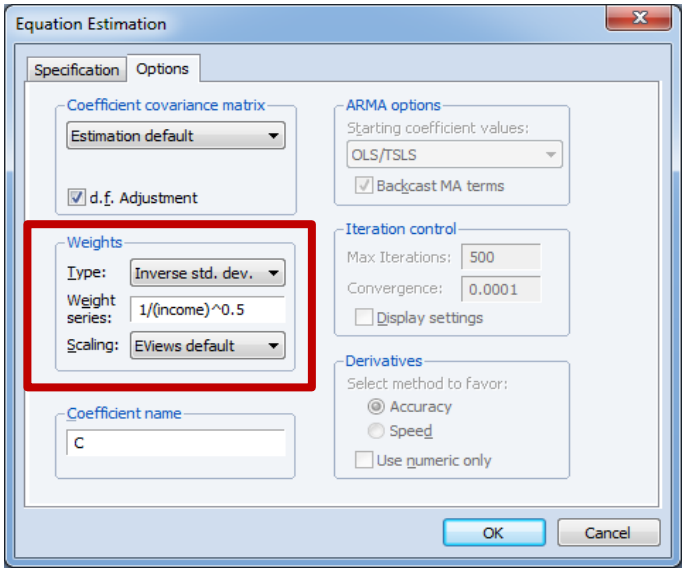


Weighted Least Squares (Part II)

- This approach to define the model isn't ideal – it cumbersome and complicated.
- EViews has a built-in method that allows us to perform weighted least squares (WLS).

To implement WLS in EViews:

1. Click ***Estimate*** on the ***Equation Box***.
2. The ***Equation Estimation*** box opens up. Click ***Options*.**
3. Under ***Weights*** → ***Type*** choose ***Inverse std. dev.***
4. Under ***Weights*** → ***Weight series***, specify the type of weights you will use to transform your data (in this case $\frac{1}{\sqrt{income_i}}$).
5. Click ***OK***.



Weighted Least Squares (Part III)

- Reading results:
  - ✓ The top panel displays the estimation setting showing the weights.
  - ✓ The middle panel shows the estimated coefficients, standard errors, and t-stats.
  - ✓ The bottom panel shows two types of statistics:
    - a. ***Weighted Statistics –*** corresponding to the actual estimated equation.
    - b. ***Unweighted Statistics –*** computed using the unweighted data and the WLS (weighted least square coefficients).

Equation: EQ12 Workfile: TUTORIAL12_RESULTS::Tutorial9a\

View | Proc | Object | Print | Name | Freeze | Estimate | Forecast | Stats | Resids

Dependent Variable: WEALTH
Method: Least Squares
Date: 11/12/12 Time: 22:44
Sample: 1 9275
Included observations: 9275
Weighting series: 1/(INCOME)^0.5
Weight type: Inverse standard deviation (EViews default scaling)

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-34.01599	1.871159	-18.17910	0.0000
INCOME	0.790422	0.023702	33.34809	0.0000
AGE	0.669814	0.039154	17.10695	0.0000
FSIZE	-1.891179	0.262116	-7.215055	0.0000

Weighted Statistics

R-squared	0.140570	Mean dependent var	14.03565
Adjusted R-squared	0.140292	S.D. dependent var	43.79404
S.E. of regression	41.47023	Akaike info criterion	10.28826
Sum squared resid	15944081	Schwarz criterion	10.29134
Log likelihood	-47707.81	Hannan-Quinn criter.	10.28931
F-statistic	505.4606	Durbin-Watson stat	1.957885
Prob(F-statistic)	0.000000	Weighted mean dep.	10.28726

Unweighted Statistics

R-squared	0.164630	Mean dependent var	19.07168
Adjusted R-squared	0.164359	S.D. dependent var	63.96384
S.E. of regression	58.47151	Sum squared resid	31696780
Durbin-Watson stat	1.926647		

Radek Hendrych
Charles University
Faculty of Mathematics and Physics
Department of Probability and Mathematical Statistics
Sokolovská 83, 186 75 Prague, Czech Republic

E-mail: hendrych@karlin.mff.cuni.cz