

## Phân tích dữ liệu bằng phần mềm EViews

### Giới thiệu

Có rất nhiều phần mềm xử lý dữ liệu cho môn học Phân tích dữ liệu của chúng ta. Các phần mềm thường được người ta sử dụng có các tên gọi như là STATA, SPSS, EXCEL, MINITAB, và EViews . . . Các phần mềm này đều có điểm chung là giúp chúng ta xử lý dữ liệu một cách mau chóng. Tuy nhiên mỗi phần mềm lại có những đặc điểm riêng. STATA có thể tốt các dữ liệu từ các cuộc điều tra lớn, SPSS có ưu điểm xử lý dữ liệu mô tả tốt dưới dạng bảng biểu, EXCEL thì có ở khắp mọi máy tính PC thông thường mà không cần phải cài đặt gì thêm . . . Riêng khoá học phân tích dữ liệu của chúng ta sẽ chủ yếu sử dụng phần mềm Eviews với phiên bản thứ 4. Ưu điểm chính của EViews có thể là cho chúng ta kết quả nhanh chóng về hàm kinh tế lượng cho các dữ liệu chéo, dữ liệu chuỗi thời gian và dữ liệu bảng, ngoài ra phần mềm này lại được chạy trong môi trường Window nên rất ít khi cần nhớ các lệnh cụ thể.

### EViews 4.0 ©

EViews là một sản phẩm đã được đăng ký bản quyền của Quantitative Micro Software, Inc.

### Các nội dung chính trong phần này:

- 1) Khởi động và copy dữ liệu
- 2) Giới thiệu menu chính trong cửa sổ EViews
- 3) Workfile
- 4) Nhập dữ liệu từ phần mềm khác
- 5) Nhập dữ liệu bằng Copy và Paste
- 6) Nhập dữ liệu bằng bàn phím
- 7) Vẽ đồ thị
- 8) Tìm các thống kê mô tả dữ liệu
- 9) Tính hệ số tương quan giữa các biến
- 10) Chọn mẫu nghiên cứu
- 11) Mở rộng Workfile nhập dữ liệu mới
- 12) Tạo các biến mới
- 13) Xây dựng hàm kinh tế lượng
- 14) Kiểm định giả thiết

## Khởi động

### Copy dữ liệu

Sử dụng Windows Explorer, hãy copy thư mục dữ liệu từ thư mục trên mạng  
L:\Econometrics\Data

vào thư mục cá nhân của Anh/Chị trên ổ đĩa Y(tên Anh/Chị):\ . Trong khoá học này, Anh/Chị sẽ luôn lưu giữ kết quả của Anh/Chị trong thư mục cá nhân của mình. **Xin hãy đừng lưu giữ kết quả của mình vào bất kỳ thư mục nào khác. Hãy thận trọng đừng sửa đổi hay viết đè lên các tập tin đã được lưu giữ trên ổ đĩa L: của mạng.**

### Khởi động EViews

Biểu tượng của EViews trên màn hình Windows trông như thế này :



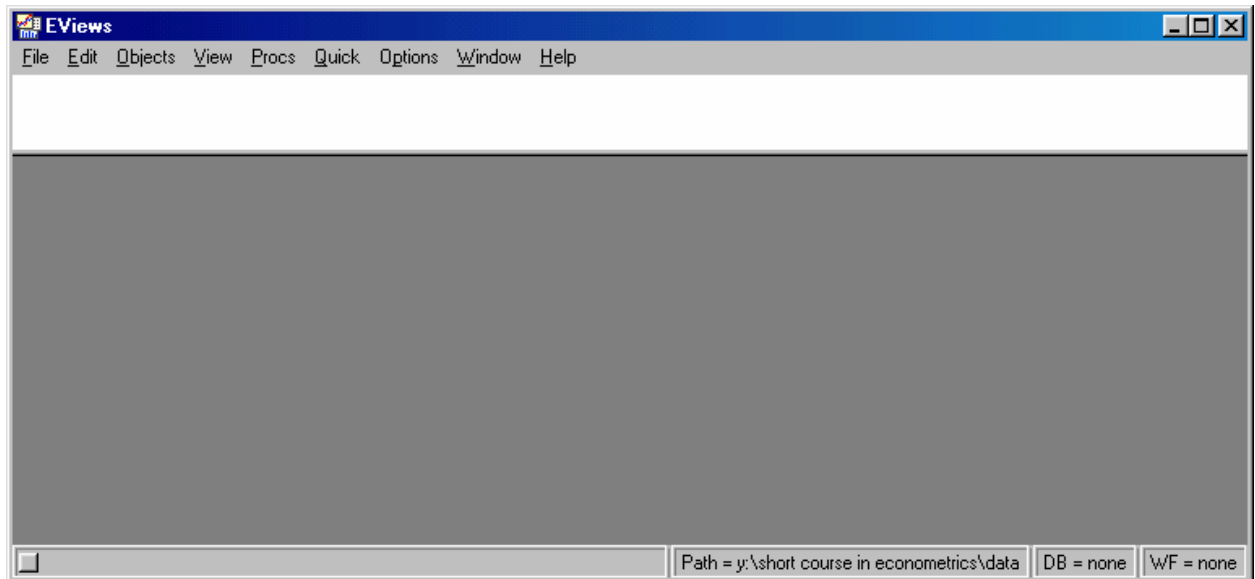
Nhấn đúp vào biểu tượng **EViews** và EViews bắt đầu hoạt động. Hãy ghi nhận menu chính, cửa sổ lệnh, cửa sổ chính, và dòng tình trạng (status line).

**Menu chính** bao gồm những lựa chọn sau:

**File Edit Objects View Procs Quick Options Window Help**

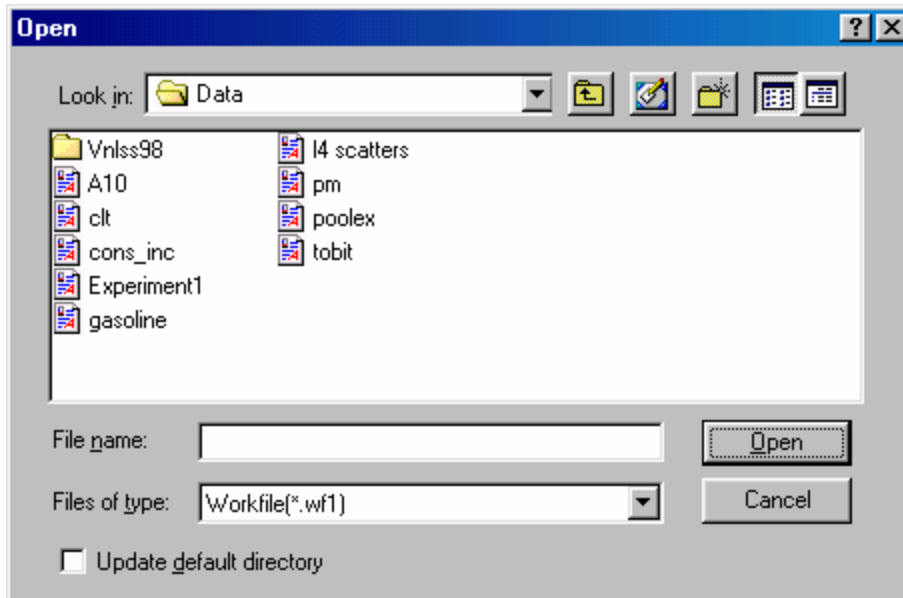
Hãy nhấp vào mỗi một trong những lựa chọn này và kiểm tra các menu phụ xuất hiện ở phía dưới. Anh/Chị có thể muốn tìm hiểu một số chi tiết về các chủ đề trong **Help**. Phương tiện **Help** của **Eviews** rất xuất sắc.

Khi Anh/Chị mở **EViews** lần đầu tiên, cửa sổ chính còn trống vì Anh/Chị còn chưa xác định tập tin làm việc nào (workfile) để sử dụng .



### Mở tập tin làm việc (Workfile) đã lưu giữ từ trước

Nhấp **File/Open/Workfile**; một màn hình mở ra, trên đó liệt kê các tập tin trong thư mục mặc định (default folder). Nếu thư mục mặc định không phải là thư mục mà anh chị mong muốn, thì Anh/Chị có thể tìm trong các thư mục cho tới khi tìm ra thư mục mà Anh/Chị muốn dùng

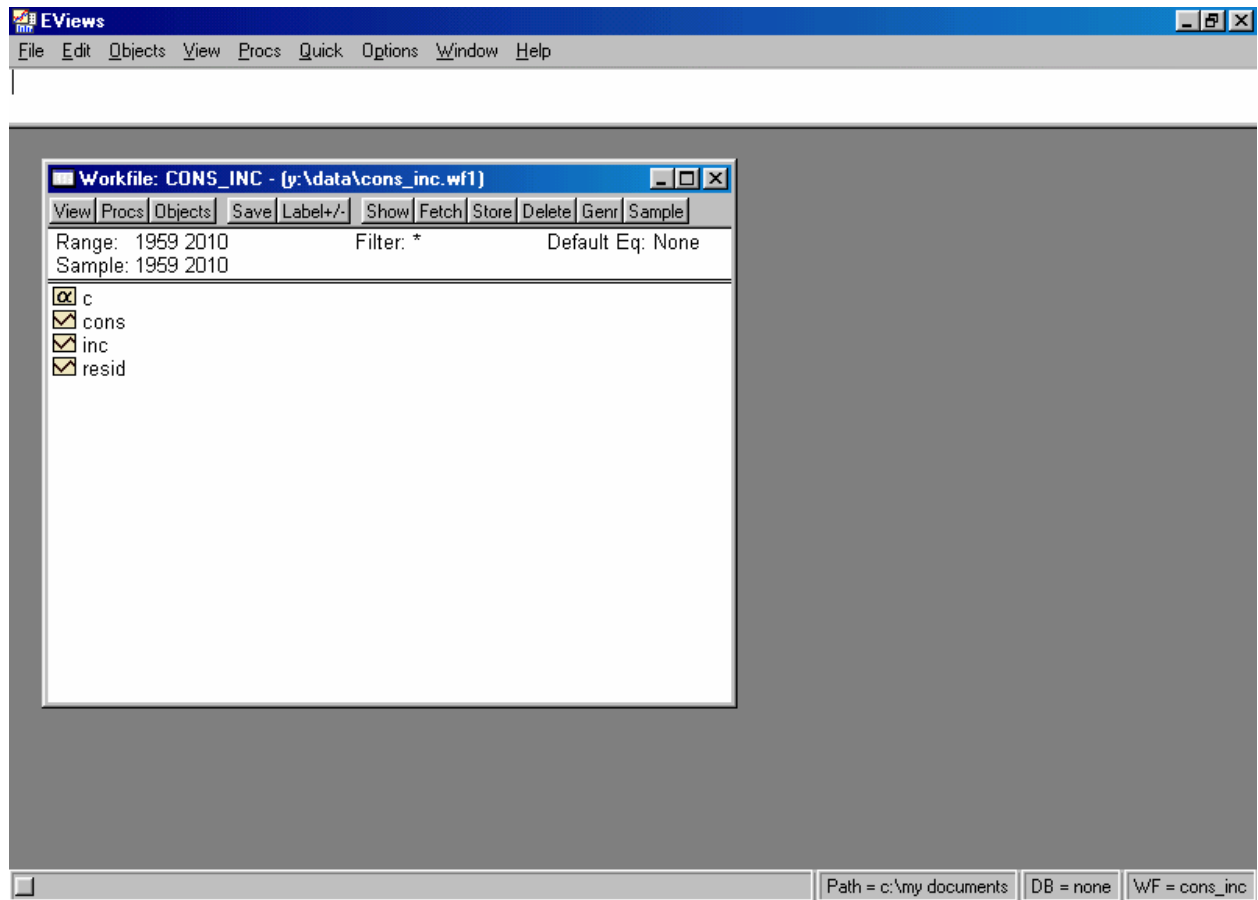


Để chỉ ra các workfile trên **EViews**, dòng "Files of type" cần xác định **Workfile (\*.wf1)**.

Vì Anh/Chị dự định sử dụng thư mục này thường xuyên, nên Anh/Chị nên nhấp chuột để đánh dấu vào ô vuông giúp cập nhật thư mục mặc định ở góc trái-phía dưới của sổ Open như trên đây. Lần sau, khi Anh/Chị khởi động **EViews**, chuỗi **File/Open/Workfile** sẽ tự động chỉ tới thư mục này.

Mở workfile có tên **cons\_inc.wf1** bằng cách nhấp đúp vào nó. Cách khác, Anh/Chị có thể bôi đen nó bằng cách nhấp đơn, sau đó nhấp Open.

Với workfile đang mở, màn hình của Anh/Chị có dạng:



Thanh trên cùng trên menu của Workfile chỉ toàn bộ đường dẫn đối với Workfile này; thanh tình trạng ở dưới đáy của màn hình cũng chỉ đường dẫn này, và nó chỉ workfile đang sử dụng : WF = cons\_inc.

Menu của workfile chứa các nút bấm dành cho

**View Procs Objects Label +/- Show Fetch Store Delete Genr Sample**

Các phím bấm với +/- là các phím chuyển đổi thực hiện các chức năng bật và tắt.

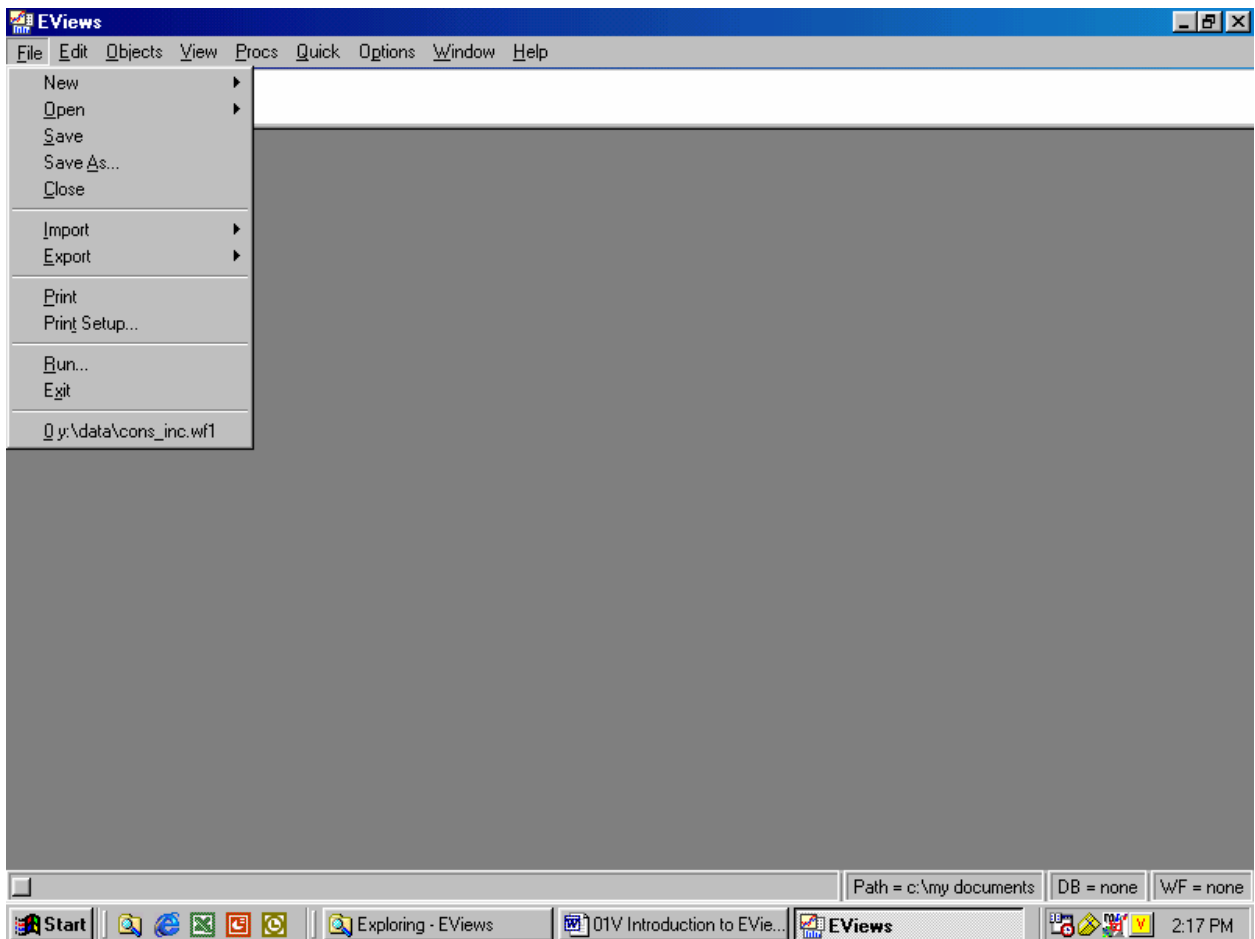
Thông tin về workfile này xuất hiện dưới thanh menu :

<b>Range</b>	<b>Filter</b>	<b>Default Equation</b>
--------------	---------------	-------------------------

Cuối cùng, phía trong cửa sổ của workfile, chúng ta thấy tất cả các đối tượng đang có mặt trong workfile này : các chuỗi (các biến), các nhóm chuỗi, vector hệ số, vector phần dư, và bất cứ phương trình, đồ thị hay bảng nào đã được đặt tên.

Bây giờ hãy đóng **cons\_inc.wf1** bằng cách nhấp vào dấu **X** ở góc phải phía trên của workfile này.

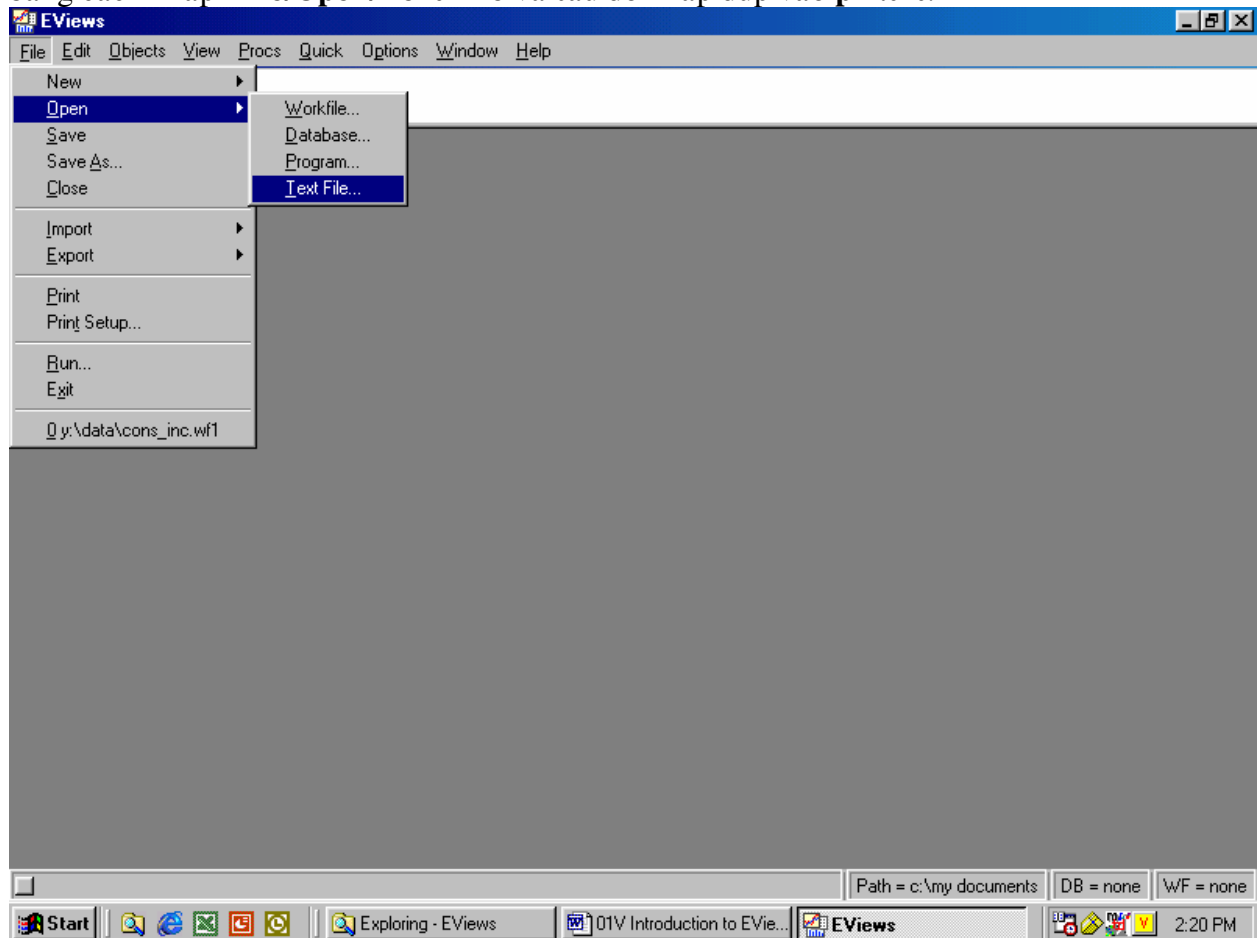
Bây giờ hãy nhấp **File**. Ghi nhớ rằng các workfile mới mở gần nhất sẽ hiện ra ở phía dưới cùng của menu trái ra phía dưới. Nếu Anh/Chị muốn mở lại một trong số đó, chỉ việc nhấp vào tên của nó.

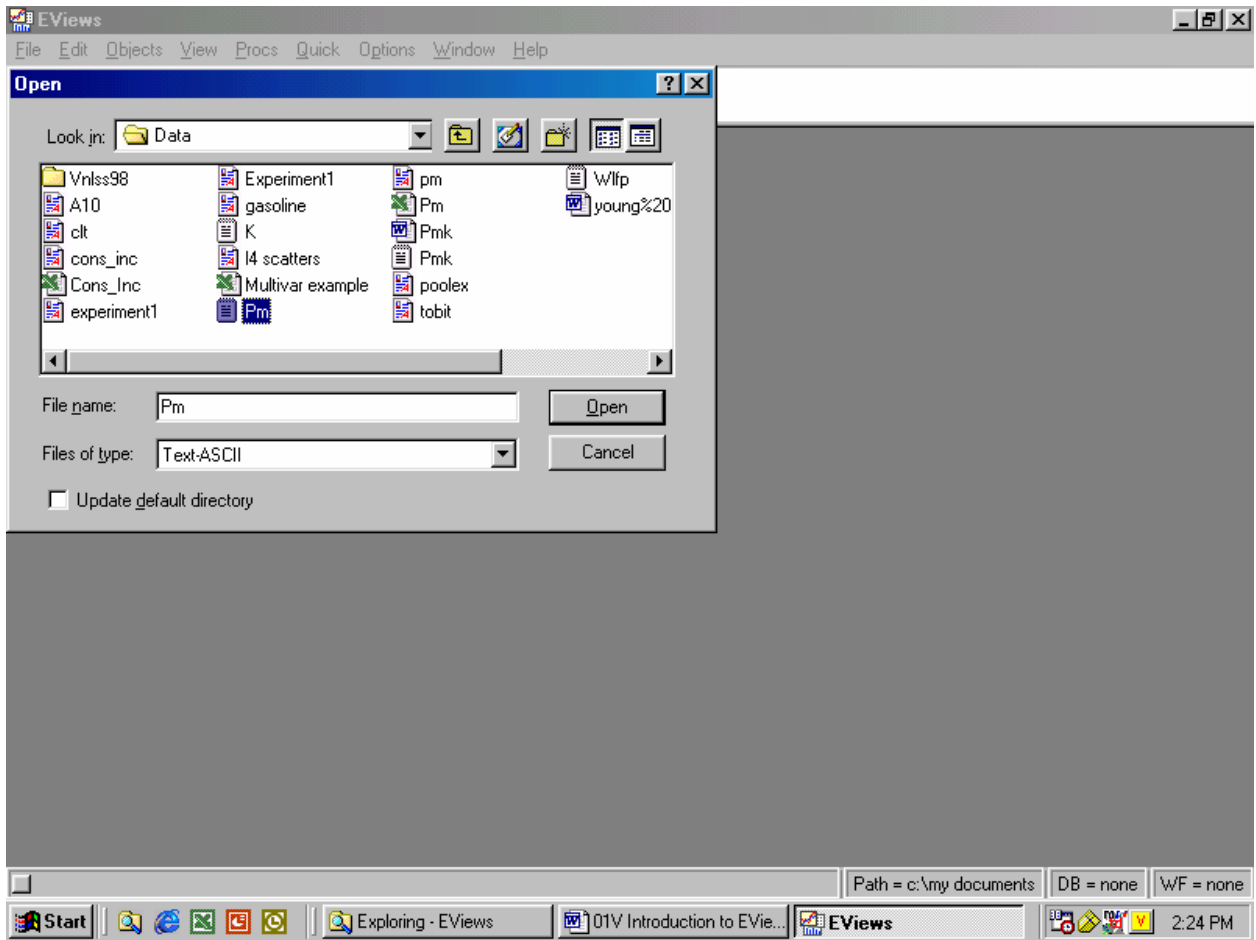


## Nhập dữ liệu từ các phần mềm khác

Dữ liệu có thể được nhập vào từ các tập tin Lotus, Excel, SPSS, MINITAB hoặc ASCII. Trong mỗi trường hợp đều dùng phương pháp như nhau; chúng ta biểu diễn với một tập tin ASCII có chứa dữ liệu sản xuất dành cho SIC 33: Các kim loại sơ cấp (Nguồn : *Phân tích Kinh tế lượng*, in lần thứ ba, của William Greene) SIC là Mã Công nghiệp Chuẩn của Hoa kỳ.

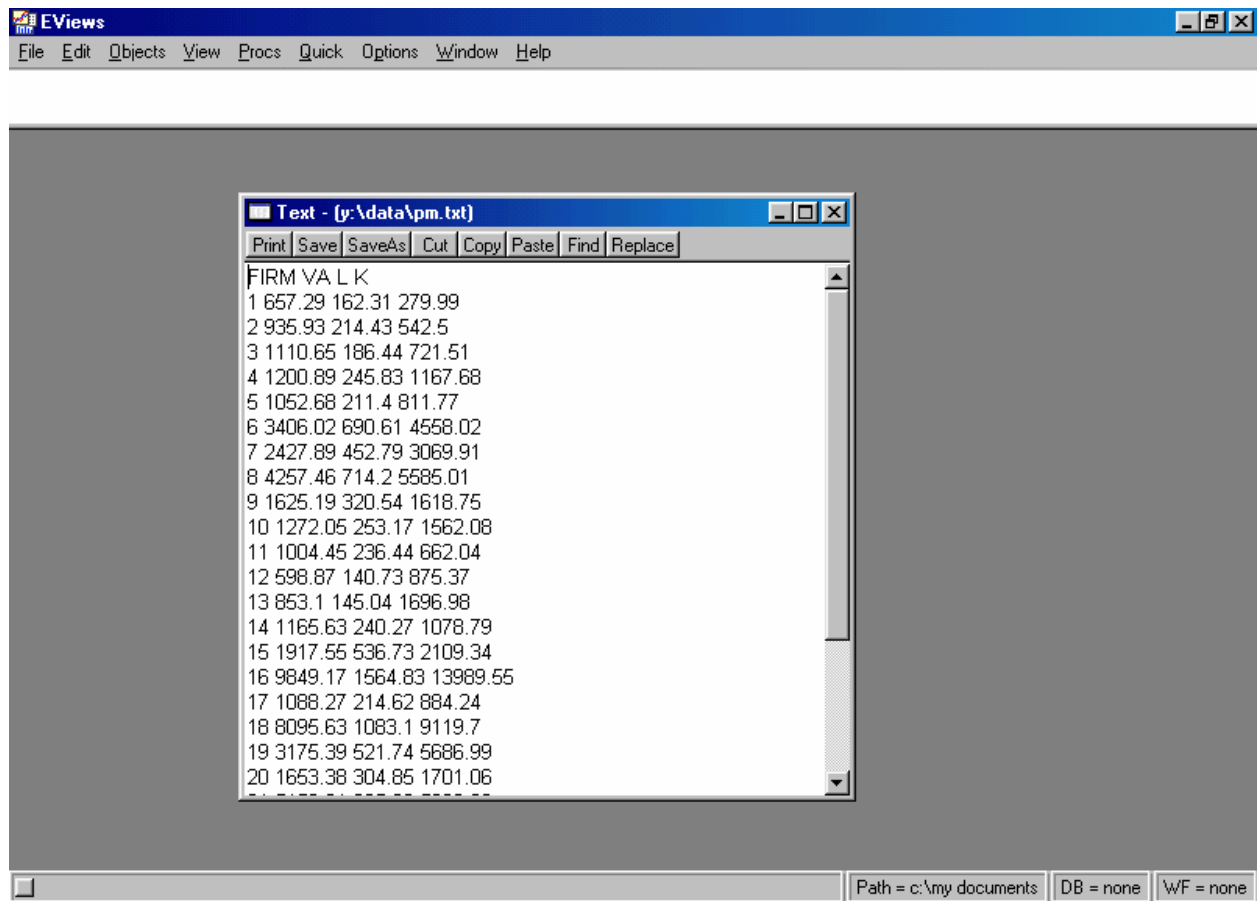
Dữ liệu này nằm trong một tập tin ASCII có tên **pm.txt**. Anh/Chị có thể thấy dữ liệu này bằng cách nhấp **File/Open/Text File** và sau đó nhấp đúp vào **pm.txt**.



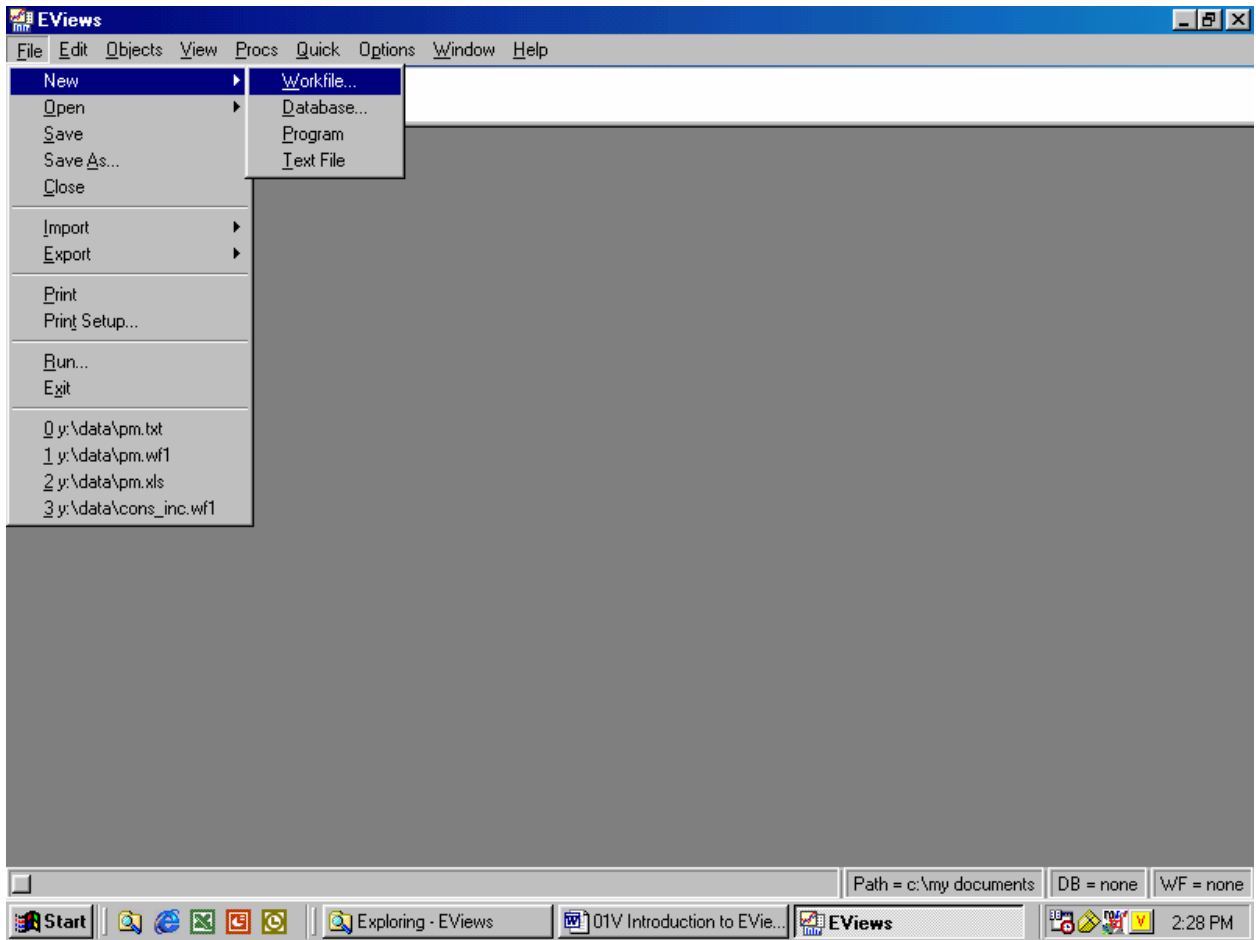


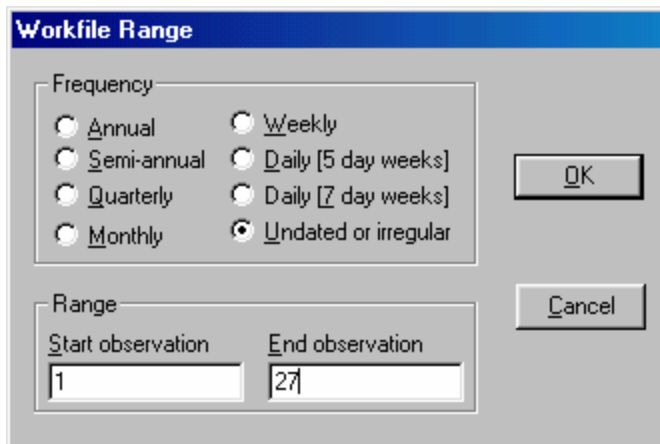


Anh/Chị cần thẩm tra tập tin này để xác định (1) có bao nhiêu biến, (2) có phải các tên của biến xuất hiện ở phía trên của tập tin không, (3) (các) ranh giới giữa các cột dữ liệu, và (4) có bao nhiêu quan sát. Hãy xem rồi sau đó đóng tập tin này lại.



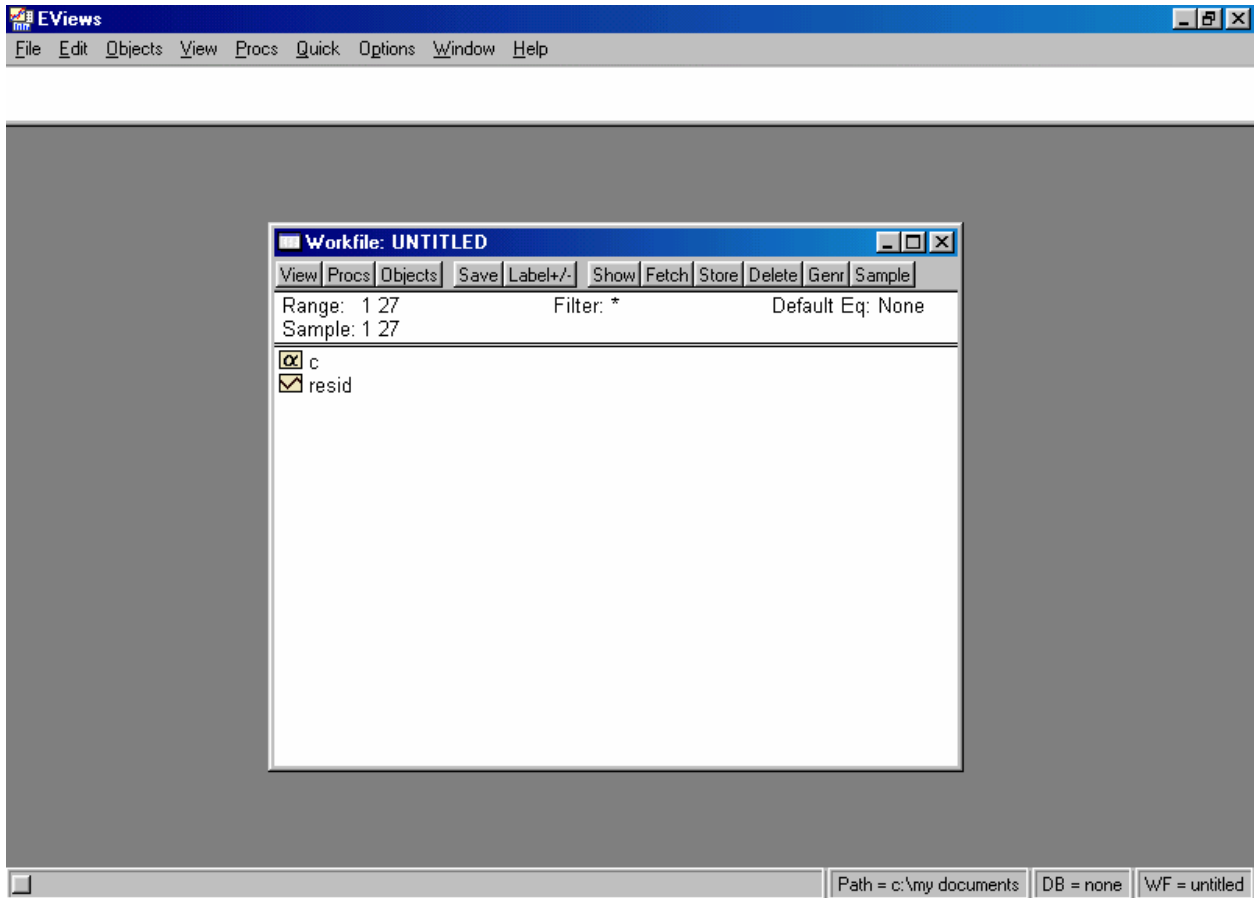
Trở lại với menu chính : nhấn **File/New/Workfile**. Bằng việc nhấp **New** trong chuỗi này chúng ta báo cho **EViews** rằng chúng ta có ý định phát triển một workfile mới. Việc này làm cho **EViews** mở ra một hộp thoại để chúng ta xác định các thuộc tính nhất định của dữ liệu. Trong trường hợp này, dữ liệu là dữ liệu chéo, nên chúng ta nhấp vào ô dành cho "**undated**," và chúng ta đánh máy **27** vào "**last observation**."

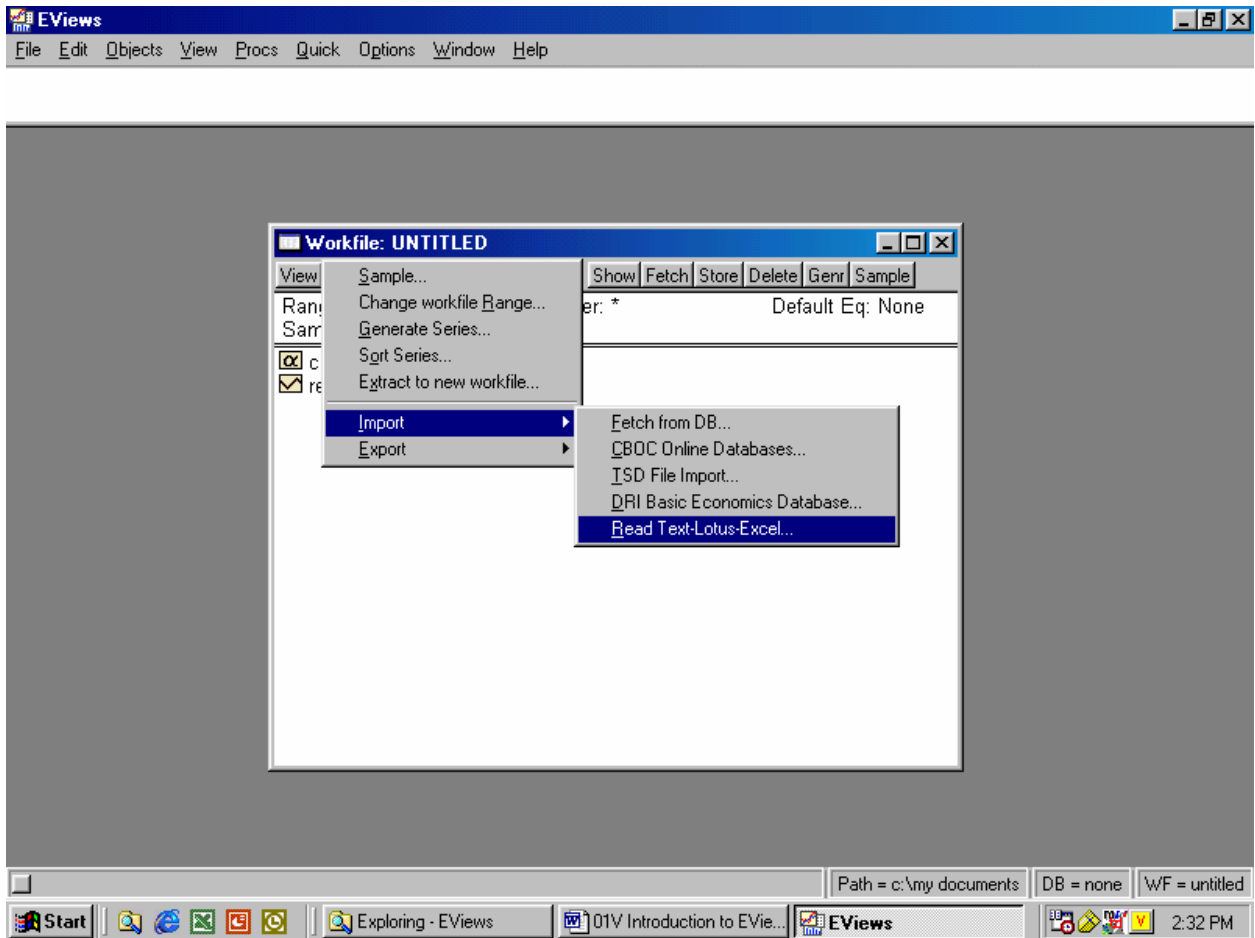




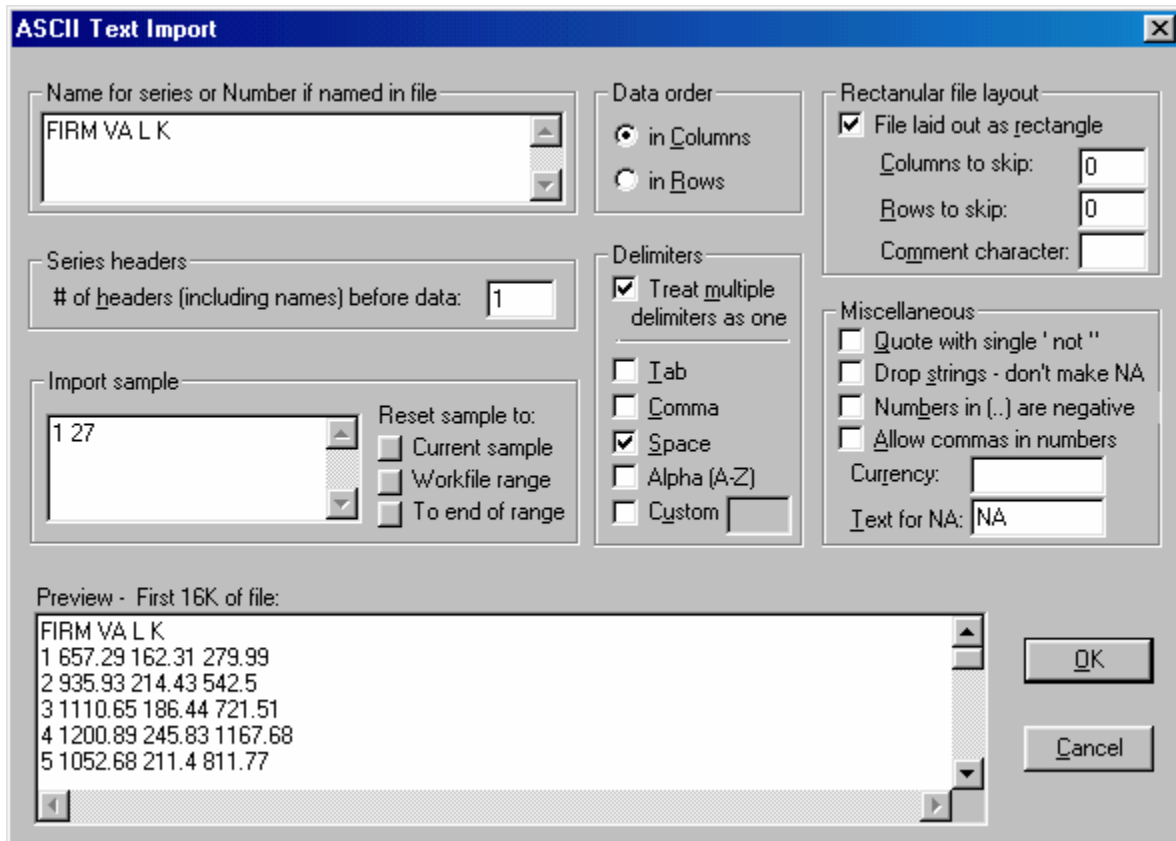
Bây giờ hãy nhấp **OK**. Một workfile mới xuất hiện.

Nhấn **Procs/Import/Read Text-Lotus-Excel...** và một hộp thoại xuất hiện để mở các tập tin. Hãy đảm bảo là dòng files of type chỉ **Text-ASCII**. Excel và Lotus cũng là những lựa chọn có sẵn. Hãy đưa về **pm.txt** và nhấp để mở nó.



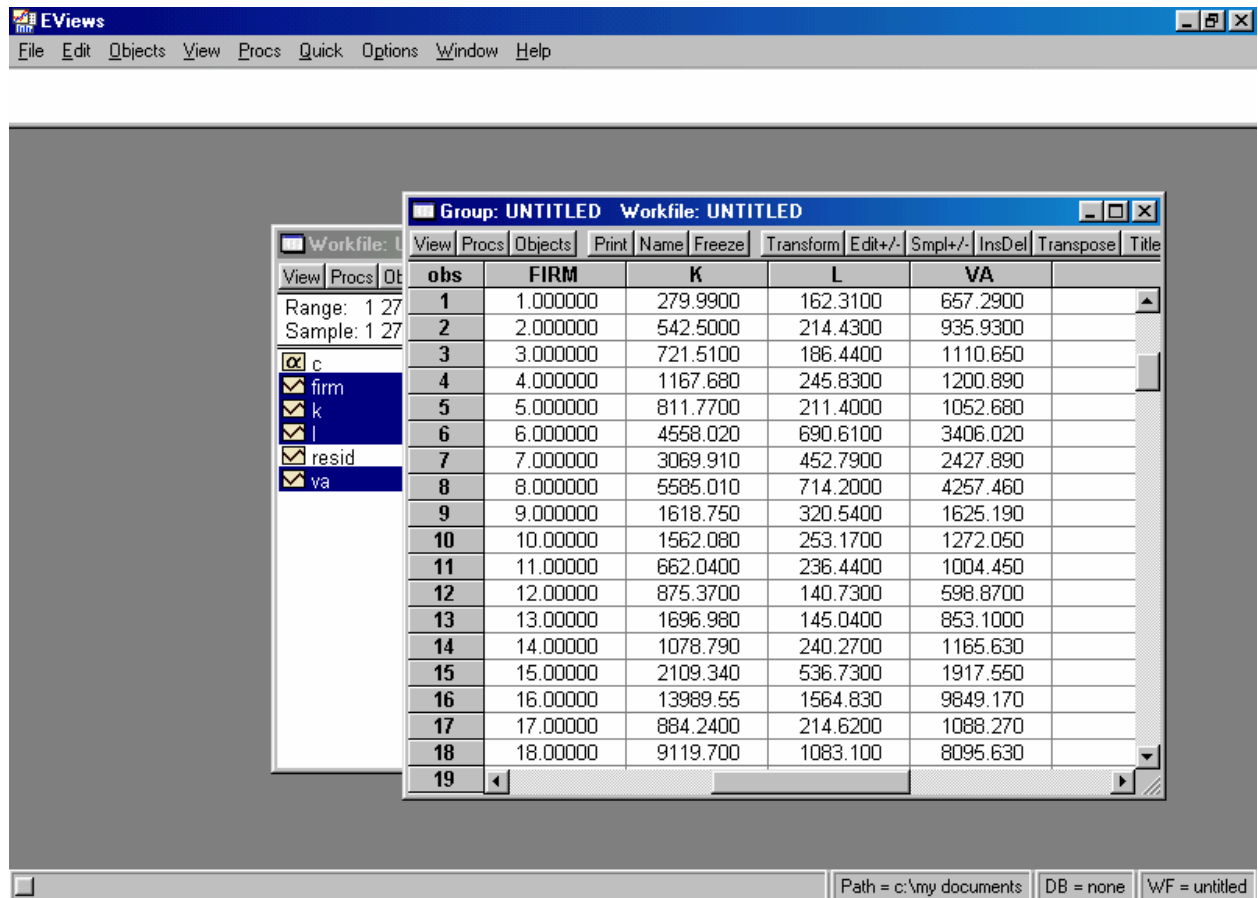


Một hộp thoại khác — phức tạp hơn — xuất hiện. Đưa danh sách các biến vào theo thứ tự **Firm V A L K**. Chúng là : Mã số hãng, Giá trị gia tăng, Lao động, và Vốn. Hãy chỉ ra rằng những biến này ở trong các cột. Kiểm tra lại các lựa chọn còn lại (bạn không cần bất cứ lựa chọn nào trong số này đối với bài tập này, nhưng hãy xem qua).



Nhấn **OK** và dữ liệu đã được nhập vào.

Xem và kiểm tra dữ liệu: sử dụng phím Ctrl và nhấp chuột vào các biến FIRM VA L K sau đó nhấp đúp vào vùng đánh dấu bảng dữ liệu sẽ được hiện ra như sau:



The screenshot shows the EViews software interface. The main window displays a workfile with the following data:

obs	FIRM	K	L	VA
1	1.000000	279.9900	162.3100	657.2900
2	2.000000	542.5000	214.4300	935.9300
3	3.000000	721.5100	186.4400	1110.650
4	4.000000	1167.680	245.8300	1200.890
5	5.000000	811.7700	211.4000	1052.680
6	6.000000	4558.020	690.6100	3406.020
7	7.000000	3069.910	452.7900	2427.890
8	8.000000	5585.010	714.2000	4257.460
9	9.000000	1618.750	320.5400	1625.190
10	10.00000	1562.080	253.1700	1272.050
11	11.00000	662.0400	236.4400	1004.450
12	12.00000	875.3700	140.7300	598.8700
13	13.00000	1696.980	145.0400	853.1000
14	14.00000	1078.790	240.2700	1165.630
15	15.00000	2109.340	536.7300	1917.550
16	16.00000	13989.55	1564.830	9849.170
17	17.00000	884.2400	214.6200	1088.270
18	18.00000	9119.700	1083.100	8095.630
19				

Lưu giữ EViews workfile mới này là một ý hay. Nhấp **File/Save As** và lưu giữ tập tin này dưới tên **pm**. EViews tự động bổ sung thêm phần đuôi **.wf1** của workfile.

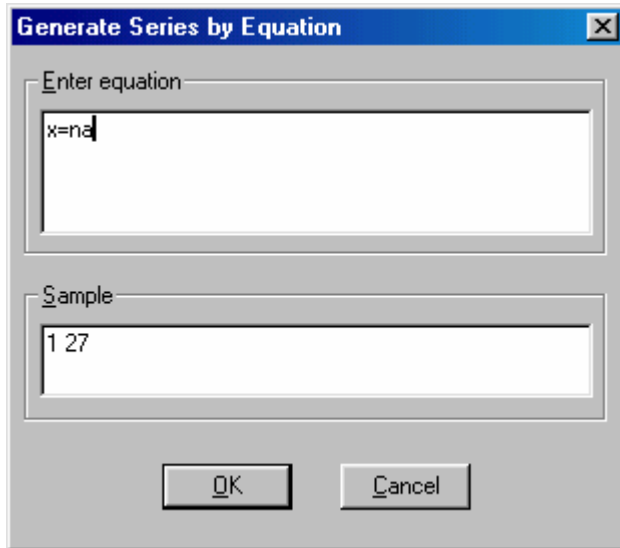
### Nhập dữ liệu bằng cách copy và dán (Copy và Paste)

Sau khi các biến đã được xác định, các giá trị của chúng có thể được thu vào bằng cách sử dụng lệnh copy, sau đó dán vào hình của bảng tính EViews.

Anh/Chị có thể copy và dán từ nhiều nguồn, bao gồm cả thư điện tử. Đôi khi Anh/Chị phải lưu giữ dữ liệu này trước hết như là một tập tin ngôn ngữ (text file), sau đó copy và dán nó.

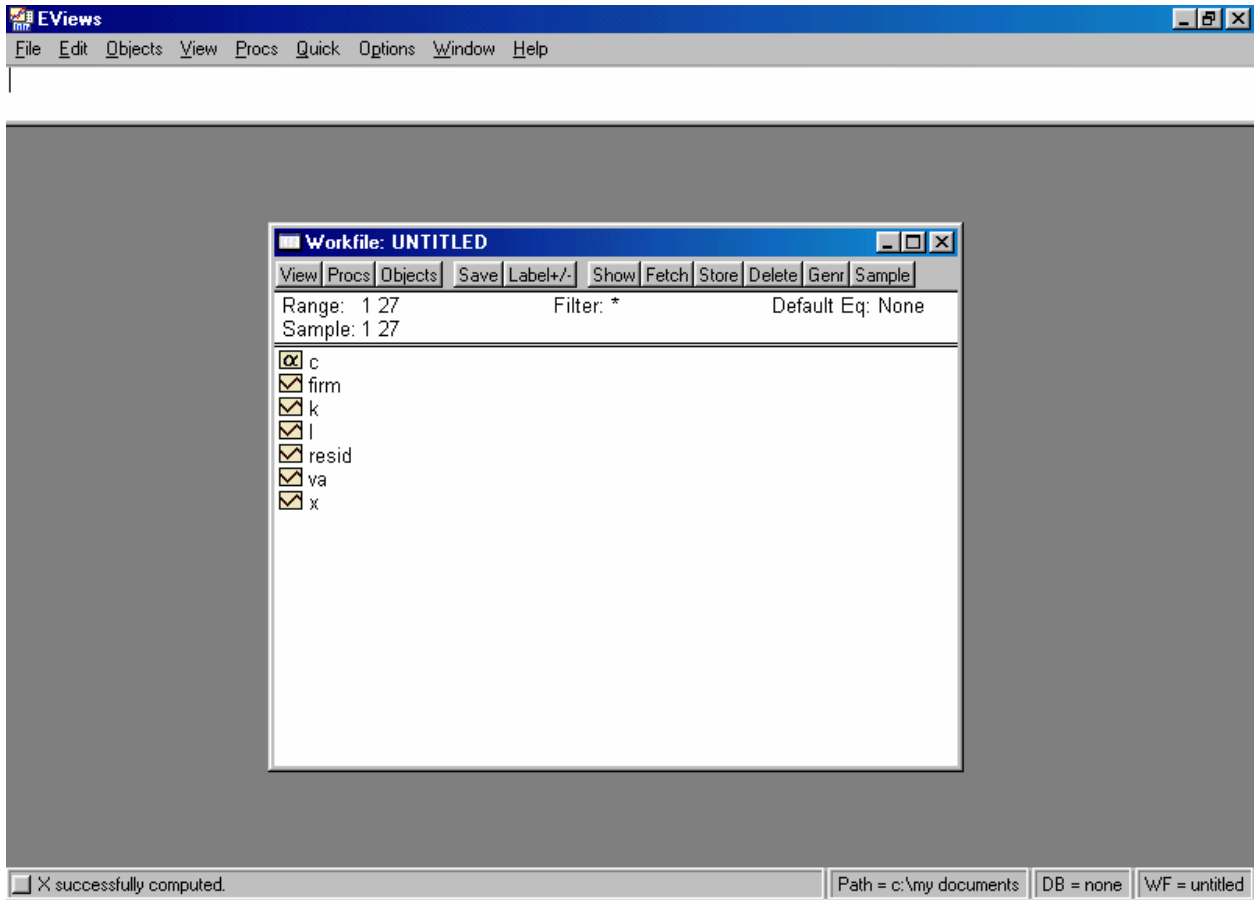
Trước khi bắt đầu thao tác copy/paste này, Anh/Chị phải tạo ra các biến trong workfile mục tiêu. Hãy tưởng tượng rằng Anh/Chị muốn copy và dán một biến có tên **X** vào workfile của mình.

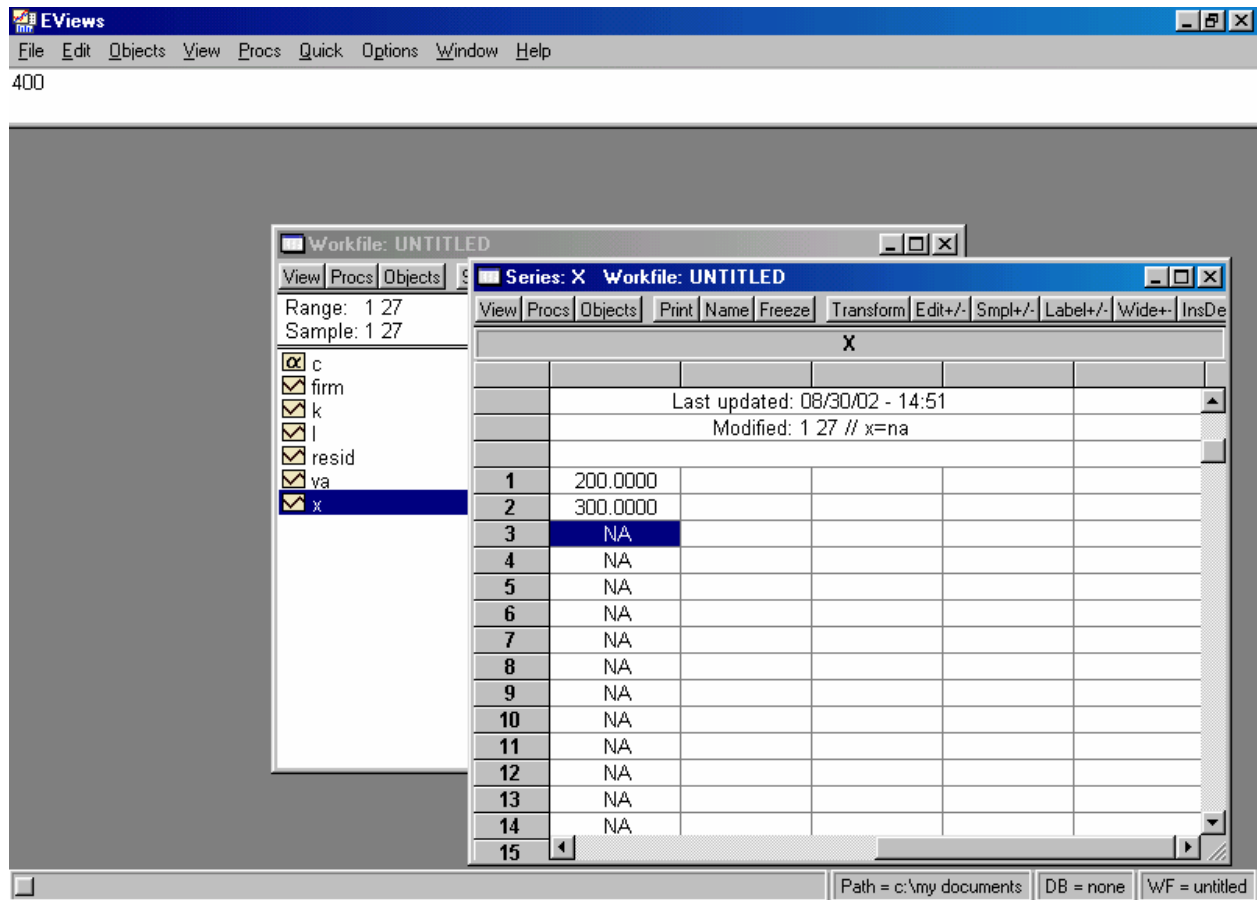
Đầu tiên, hãy mở workfile này, sau đó nhấp **Genr** và đánh máy **X = NA**; động tác này tạo ra một biến có tên **X** nhưng tất cả các giá trị của nó đều không có .



Mở chuỗi **X** . Bây giờ nhấp **Edit+/-** . Lúc này, Anh/Chị hãy bôi đen và copy dữ liệu cho **X** từ nguồn của nó, sau đó nhấp vào ô đầu tiên của chuỗi **X** trên **EViews** và nhấp **Edit/Paste** ở phần menu chính. Dữ liệu này rơi vào vị trí và Anh/Chị hoàn tất thao tác này bằng cách tắt edit (nhấp **Edit+/-**).







## Nhập dữ liệu qua bàn phím

Phương pháp này đơn giản nhưng phiền phức; nếu việc nạp dữ liệu cần được thực hiện bằng tay thì người ta thích làm bằng cách sử dụng một chương trình bảng tính (như là Excel), sau đó nhập vào EViews.

## Làm việc trong một Workfile

Làm việc trong một Workfile là thực hiện các mục tiêu như thống kê mô tả, vẽ đồ thị, xây dựng hàm kinh tế lượng . . .

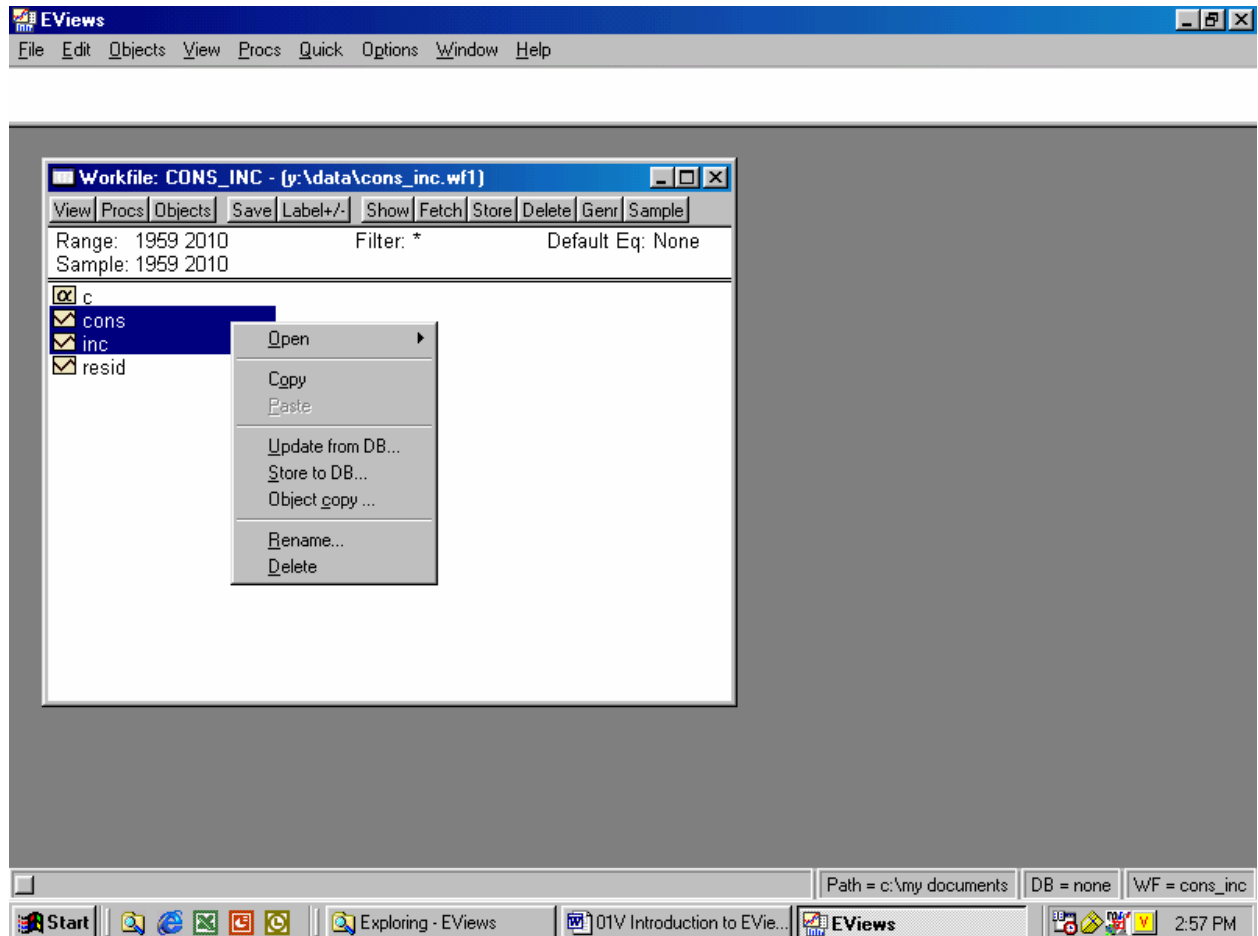
## Lựa chọn các đối tượng từ một Workfile

Nhấp **File/Open/Workfile** sau đó nhấp đúp vào **cons\_inc.wf1**.

Với mục đích lựa chọn một đối tượng (hoặc nhiều đối tượng) để làm việc, Anh/Chị hãy bôi

đen chúng. Nếu Anh/Chị muốn bôi đen nhiều đối tượng, thì hãy giữ phím **Ctrl** và nhấp vào mỗi mục, hoặc kéo con trỏ qua các mục mà bạn muốn bôi đen. Hãy bôi đen **cons** và **inc**.

Nhấp đúp vào một trong các mục đã bôi đen (**hoặc nhấp chuột phải**), chọn **Open Group**, và sau đó cả hai mục đều hiện trên bảng tính. Nhấp **name** và tên đặt tên cho nhóm **cons\_inc**; đóng bảng tính lại.

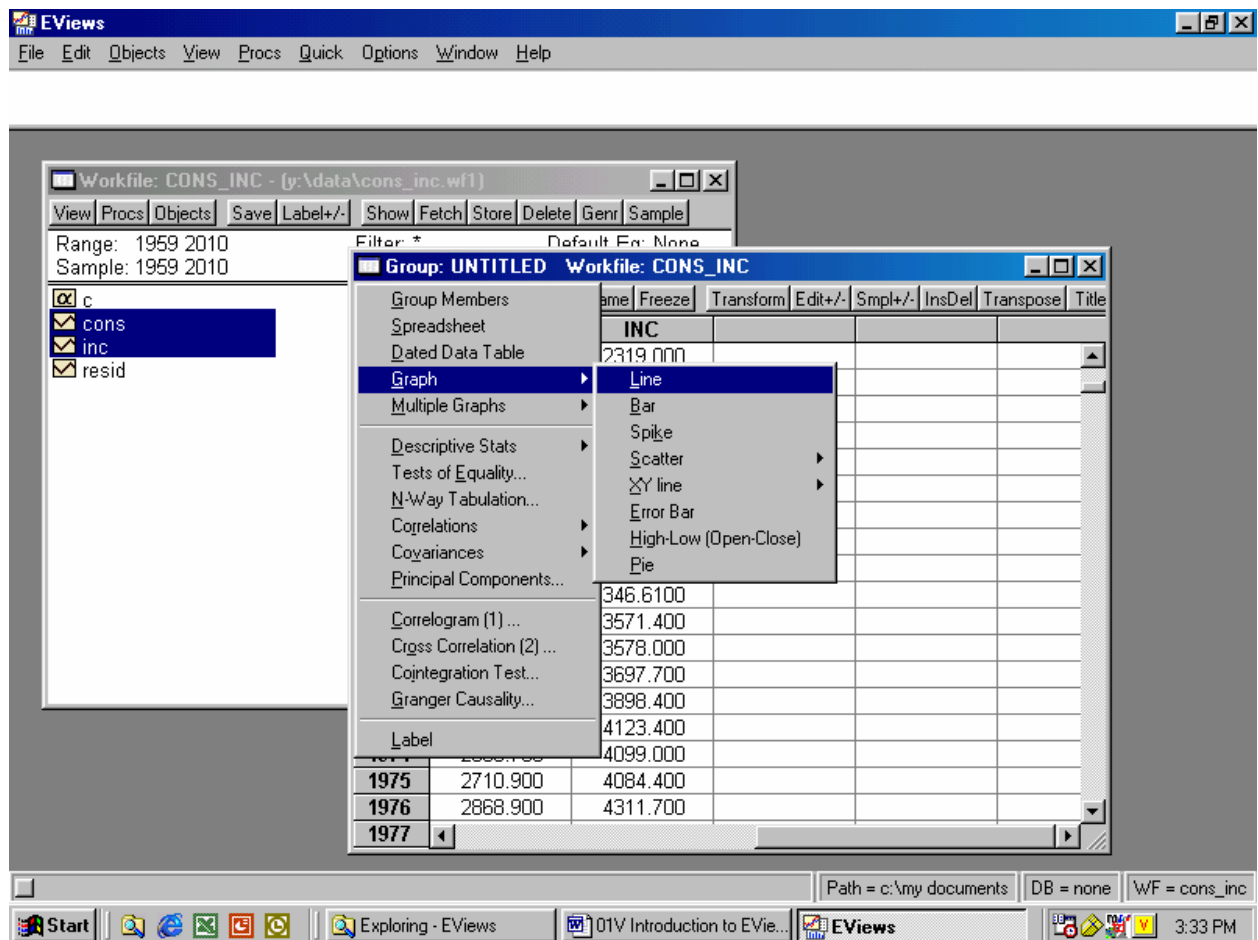


Nhấp vào nhóm **cons\_inc**; bây giờ nó đã được bôi đen. Nhấn đúp vào nó, và nó mở ra thành dạng bảng tính.

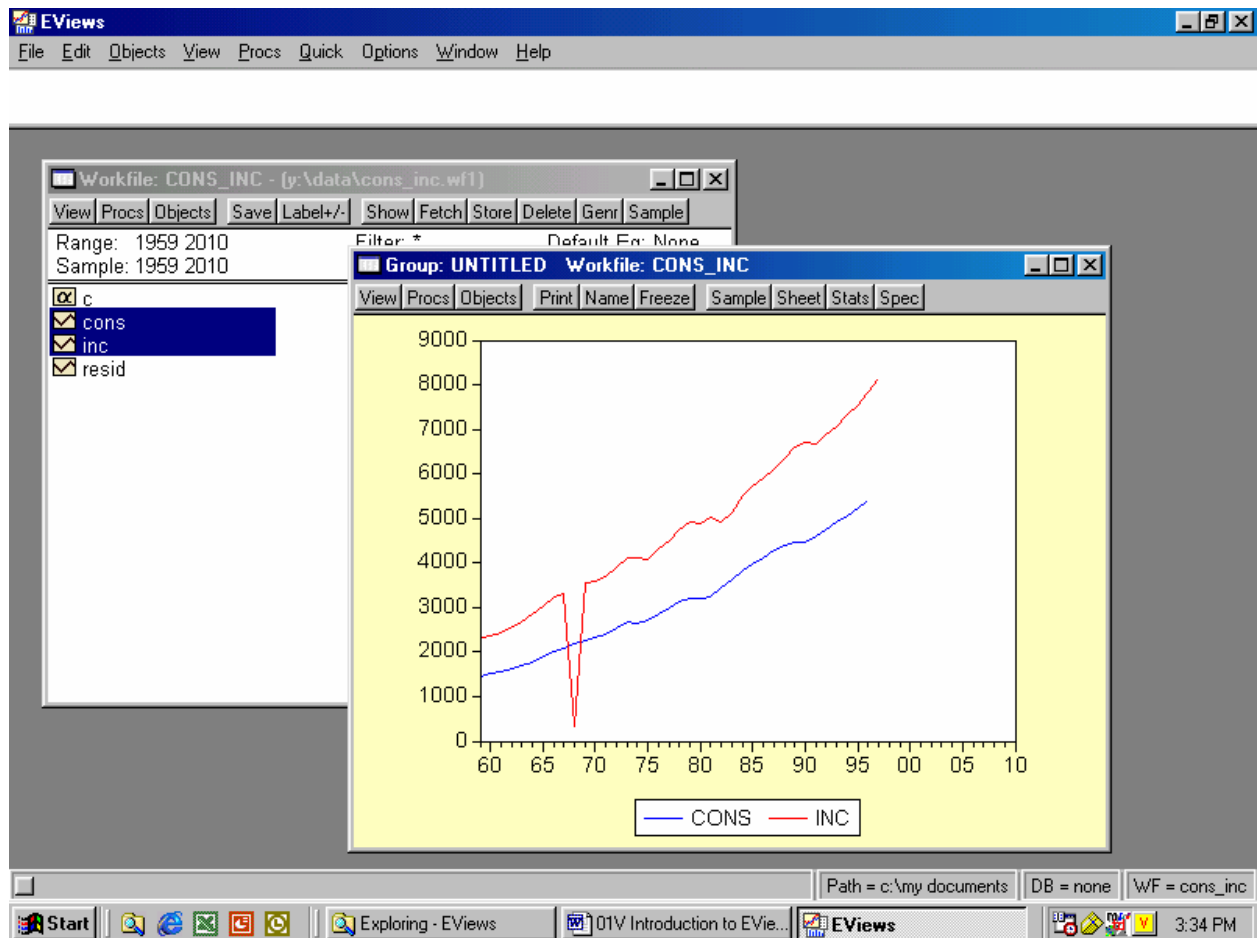
### Xem và biên tập các hạng mục vào Workfile

Kiểm tra chuỗi **cons** và **inc**. Hãy sử dụng các phím  $\leftarrow \uparrow \rightarrow \downarrow$  để di chuyển trên bảng tính.

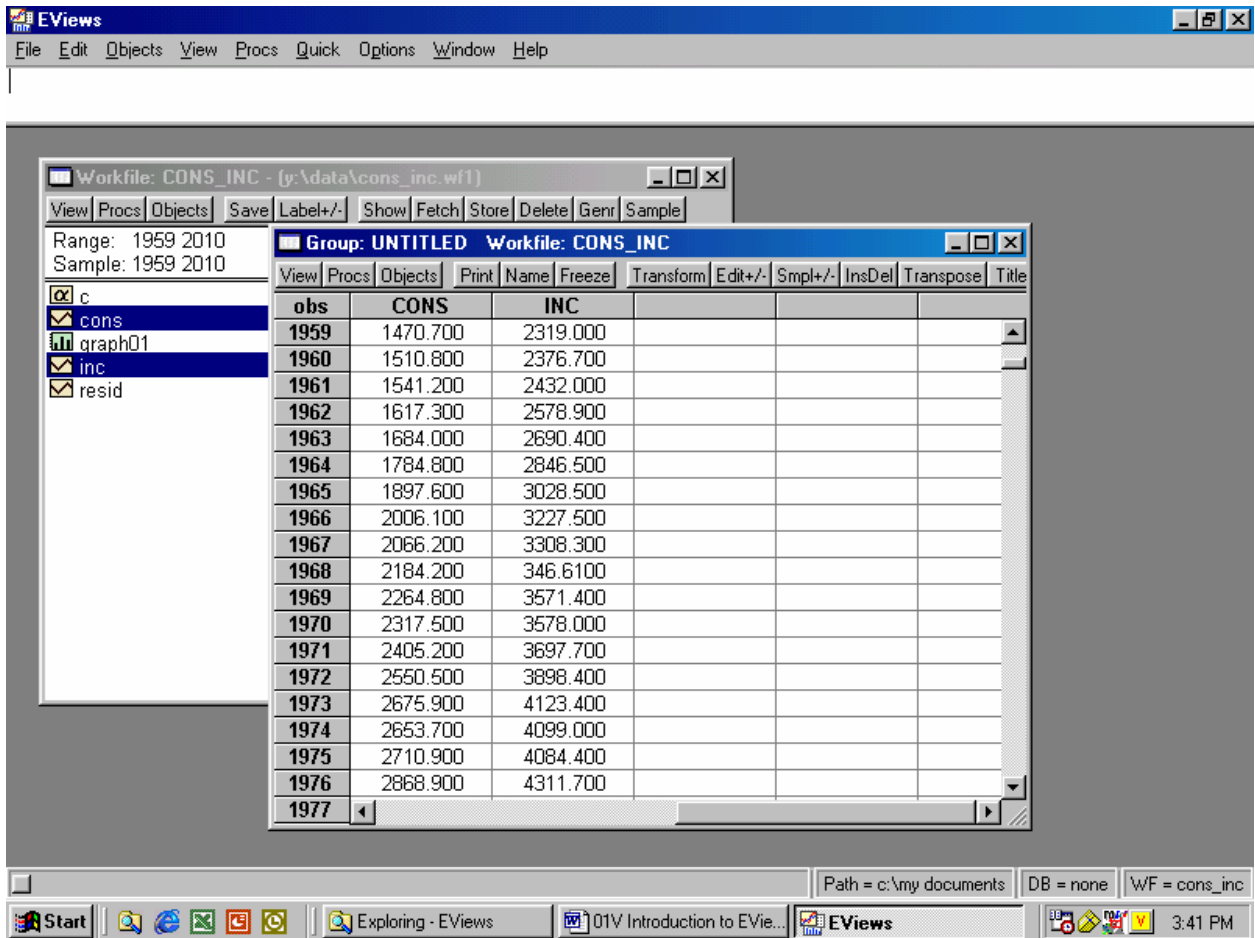
## Vẽ đồ thị trong EViews

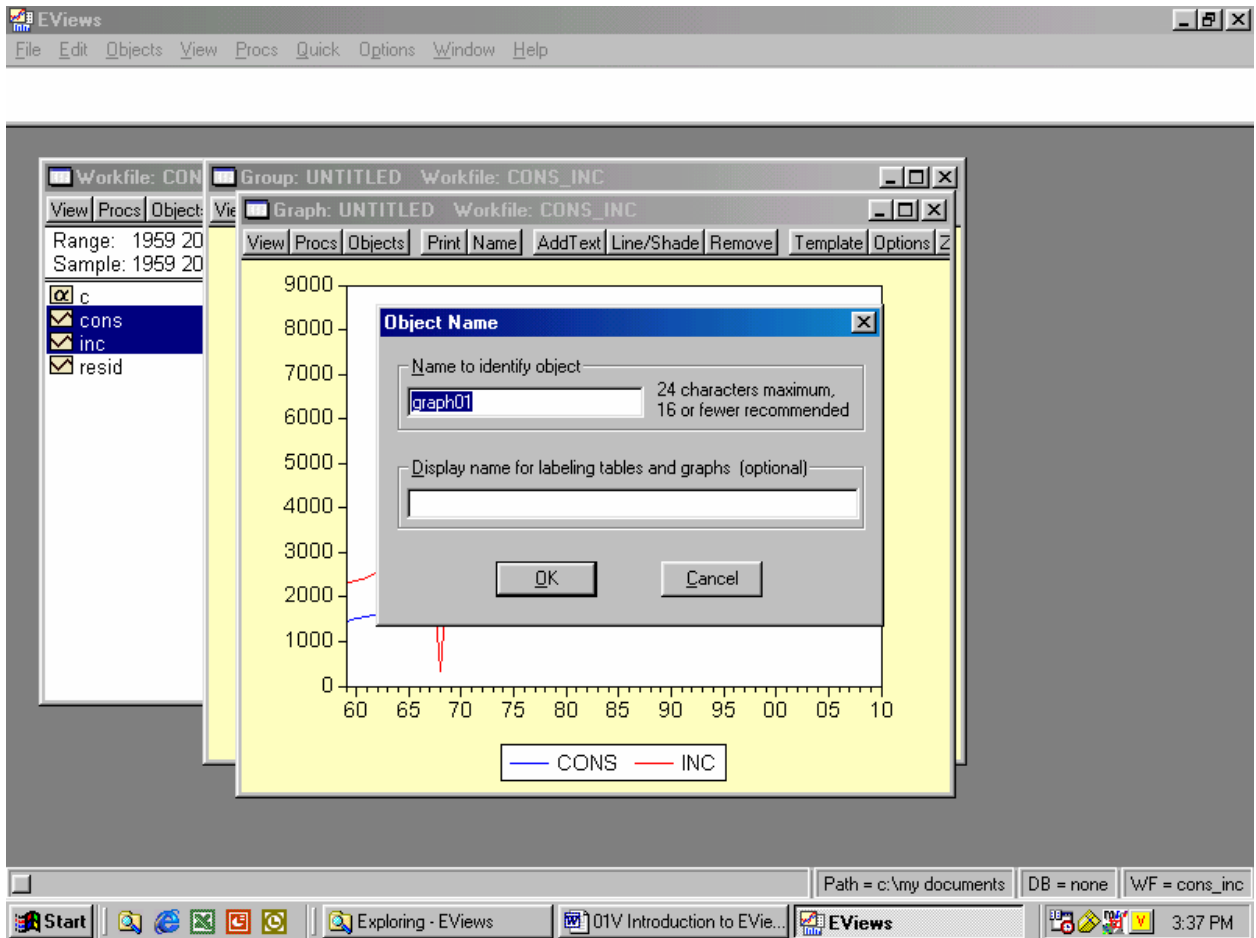


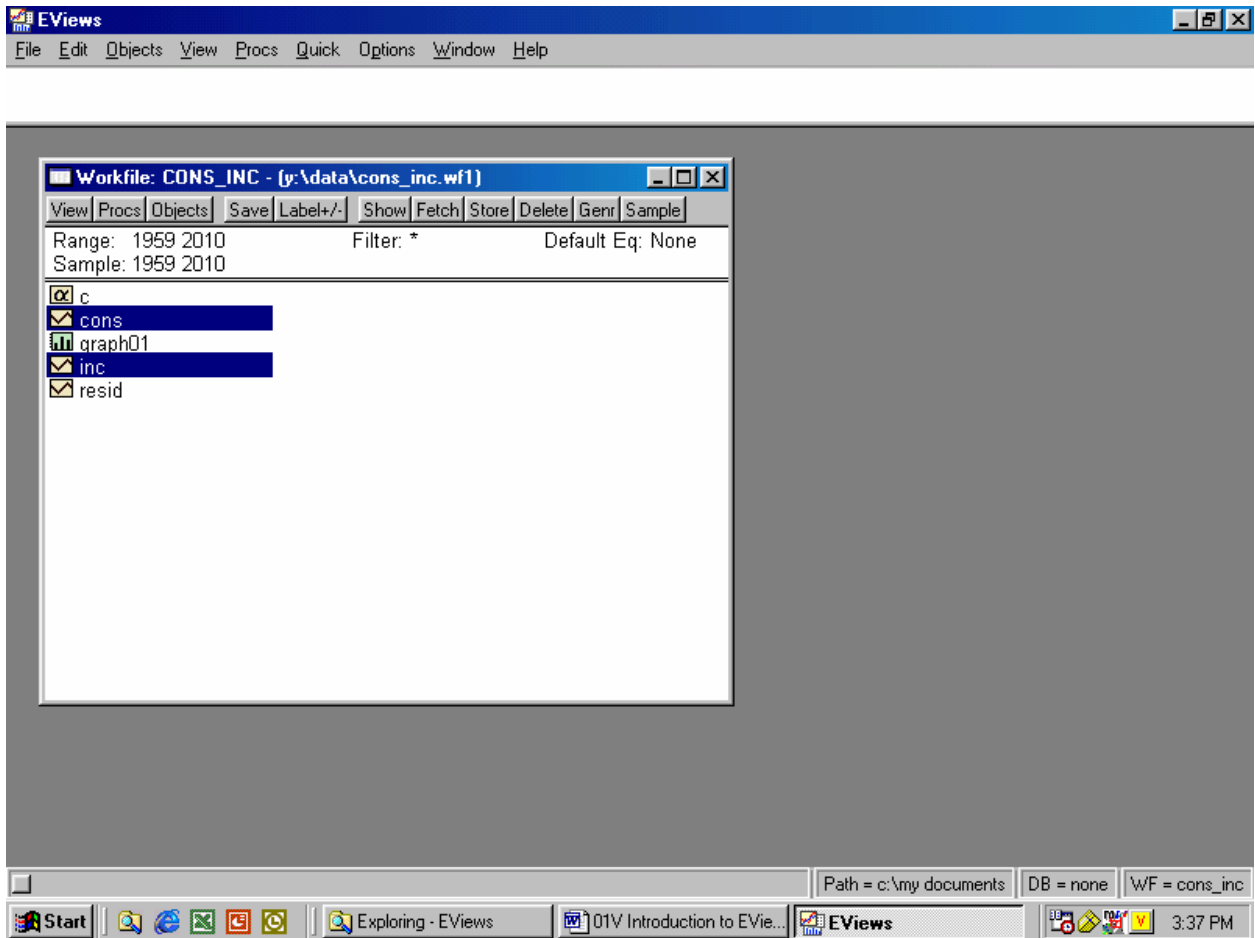
Nhấn **View/Graph/Line** trên menu của nhóm. **Cons** và **inc** có các xu hướng tương tự nhau, nhưng **inc** có một điểm dị biệt vào năm 1968. Kiểm tra dữ liệu gốc cho thấy rằng số liệu đưa vào cho **inc** trong năm 1968 cần phải là 3466.1 thay cho 346.61



Chúng ta sẽ thay đổi dữ liệu của năm 1968 cho chuỗi **inc** này. EViews cập nhật năng động các đồ thị và bảng tính cho phù hợp với các thay đổi trong dữ liệu. Nếu chúng ta muốn bảo lưu đồ thị này như một ghi nhận cho dữ liệu gốc, thì chúng ta có thể nhấn vào phím **Freeze**, sau đó đặt cho đồ thị này tên gọi : **Graph\_1** (hay bất cứ tên nào bạn muốn). Lưu ý rằng khi chúng ta Freeze (đóng băng) đồ thị này, các lựa chọn trong menu mới sẽ sẵn sàng. Chúng ta sẽ trở lại với chúng sau. Đóng đồ thị có tên **Graph\_1**. Khi Anh/Chị **Freeze** một bảng tính, nó được lưu giữ như là một **bảng**.





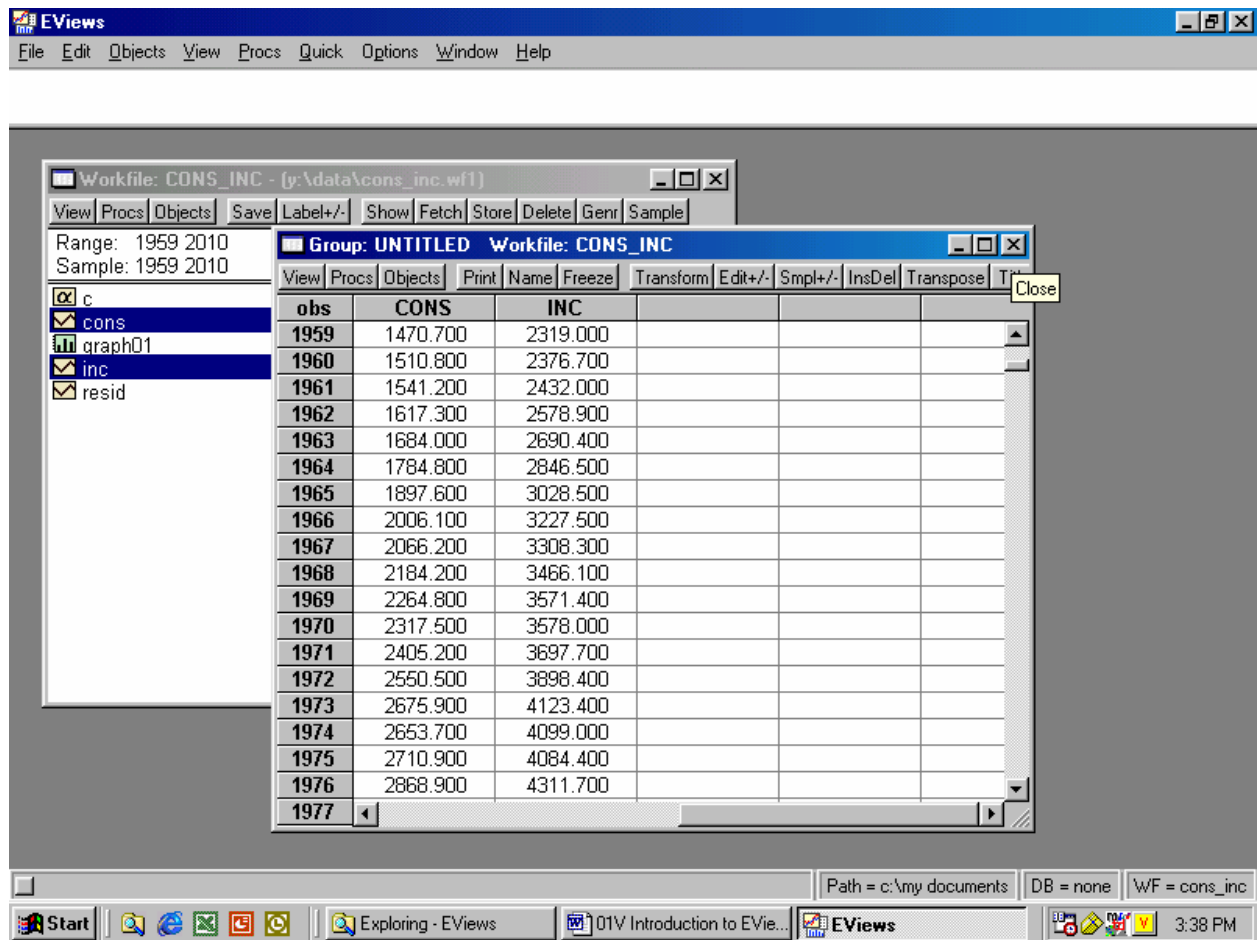


Bây giờ, trên đồ thị chưa có tên hãy nhấp **View/Spreadsheet**. Nhấn **Edit+/-** để bật chức năng hiệu chỉnh đồ thị.

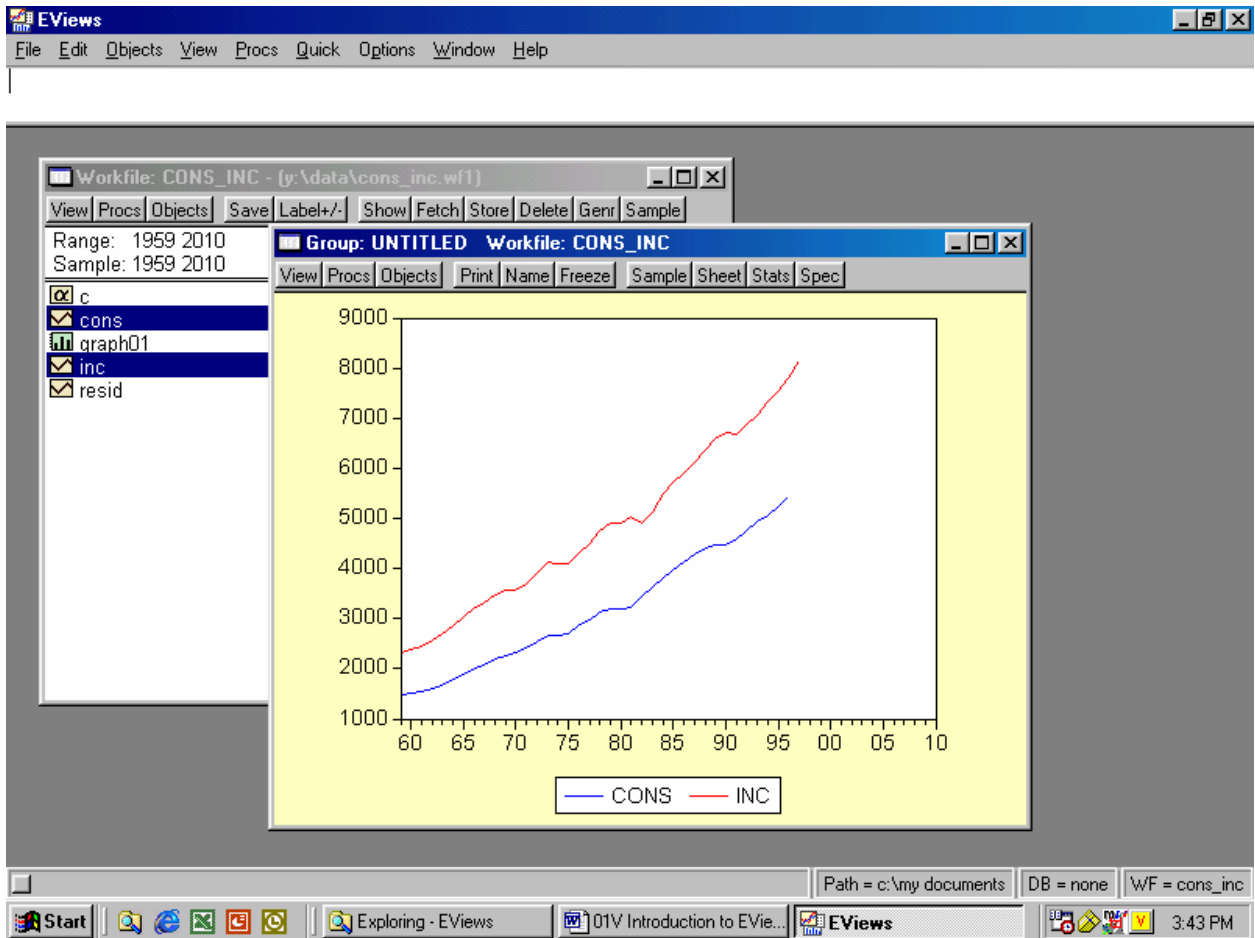
Nhấp vào 346.61 và đánh máy giá trị chính xác : 3466.1.

Nhấn **Edit+/-** để tắt chức năng biên tập.





Bây giờ hãy biểu diễn đồ thị này một lần nữa bằng cách nhấn **View/Graph/Line** trên menu của nhóm; hãy so sánh nó với đồ thị **Graph\_1**. Nhìn vào đồ thị là một cách tốt để phát hiện các sai sót.



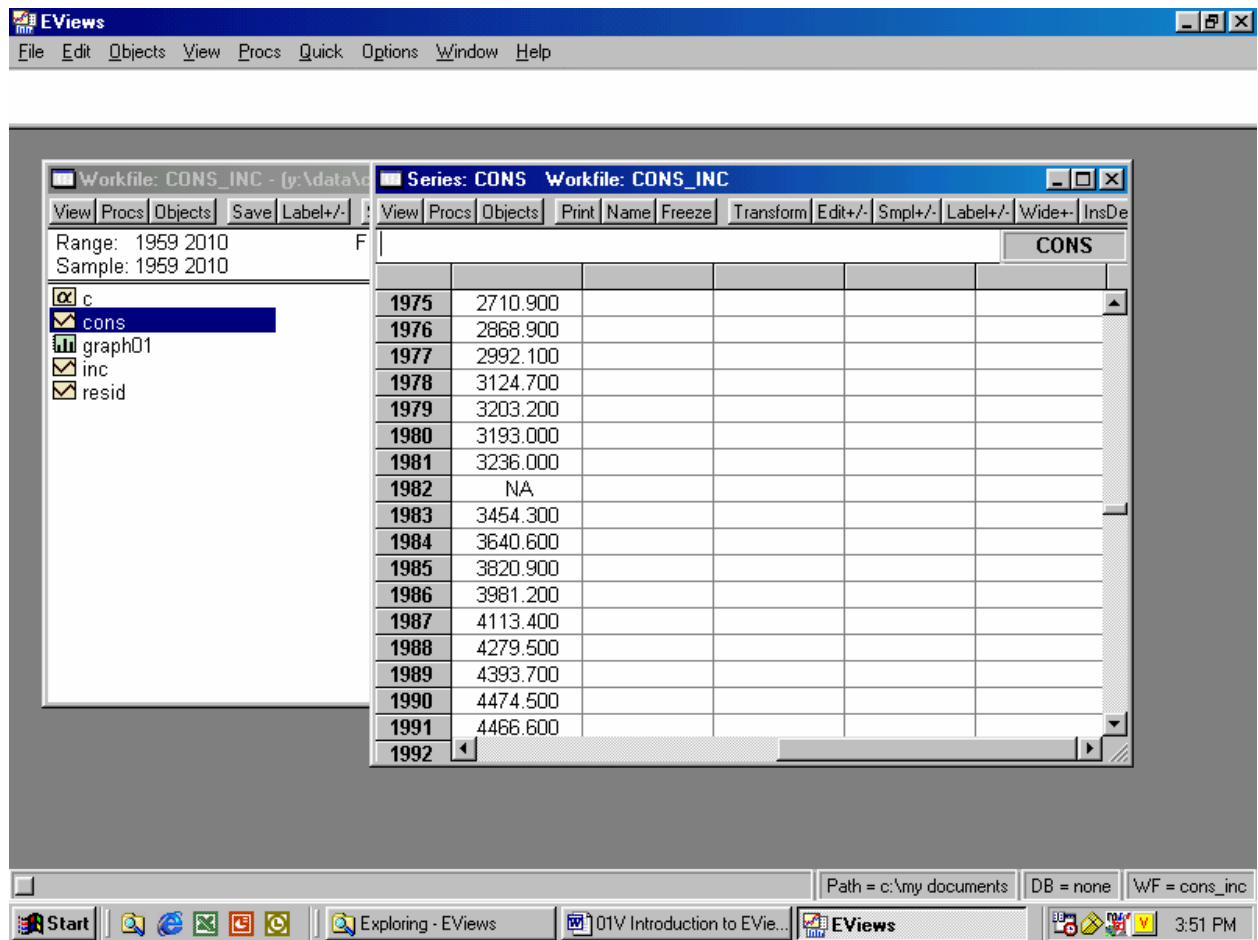
Anh/Chị có thể nhấn đúp vào đồ thị này để biến đổi các lựa chọn khi vẽ đồ thị; chúng ta sẽ sử dụng đặc điểm này sau .

The screenshot shows the EViews software interface. A window titled 'Group: UNTITLED Workfile: CONS\_INC' is open, displaying a data table. The table has columns for 'obs', 'CONS', and 'INC'. The rows represent years from 1980 to 1998. The 'CONS' column contains values for each year, except for 1997, which is 'NA'. The 'INC' column contains values for each year. The status bar at the bottom indicates 'Path = c:\my documents', 'DB = none', and 'WF = cons\_inc'.

obs	CONS	INC
1980	3193.000	4900.900
1981	3236.000	5021.000
1982	3454.300	4919.300
1983	3640.600	5132.300
1984	3820.900	5505.200
1985	3981.200	5717.100
1986	4113.400	5912.400
1987	4279.500	6113.300
1988	4393.700	6368.400
1989	4474.500	6591.800
1990	4466.600	6707.900
1991	4594.500	6676.400
1992	4748.900	6880.000
1993	4928.100	7062.600
1994	5075.600	7347.700
1995	5237.500	7543.800
1996	5423.900	7813.200
1997	NA	8159.500
1998		

Anh/Chị cũng có thể đã nhận thấy rằng không có số liệu cho **cons** trong năm 1997. Lý do là dữ liệu gốc cho thấy số liệu năm 1982 đã bị bỏ qua và điều này làm thiếu dữ liệu năm 1997. Dữ liệu gốc cho biết cons có giá trị 3275.5. Giải pháp của chúng ta là:

Đóng nhóm này lại, sau đó chỉ mở **cons** theo dạng bảng. Nhấp **InsDel**, thay đổi nó thành 1982, sau đó nhấn **OK**. EViews đã định vị **N/A** trong ô dành cho năm 1982 và đã đẩy tất cả mọi quan sát còn lại xuống dưới.



Bây giờ nhấn **Edit+/-** để chuyển sang chức năng biên tập.

Nhấn vào **N/A**

Đánh máy 3275.5

Nhấn **Edit+/-** để chuyển khỏi chức năng biên tập.

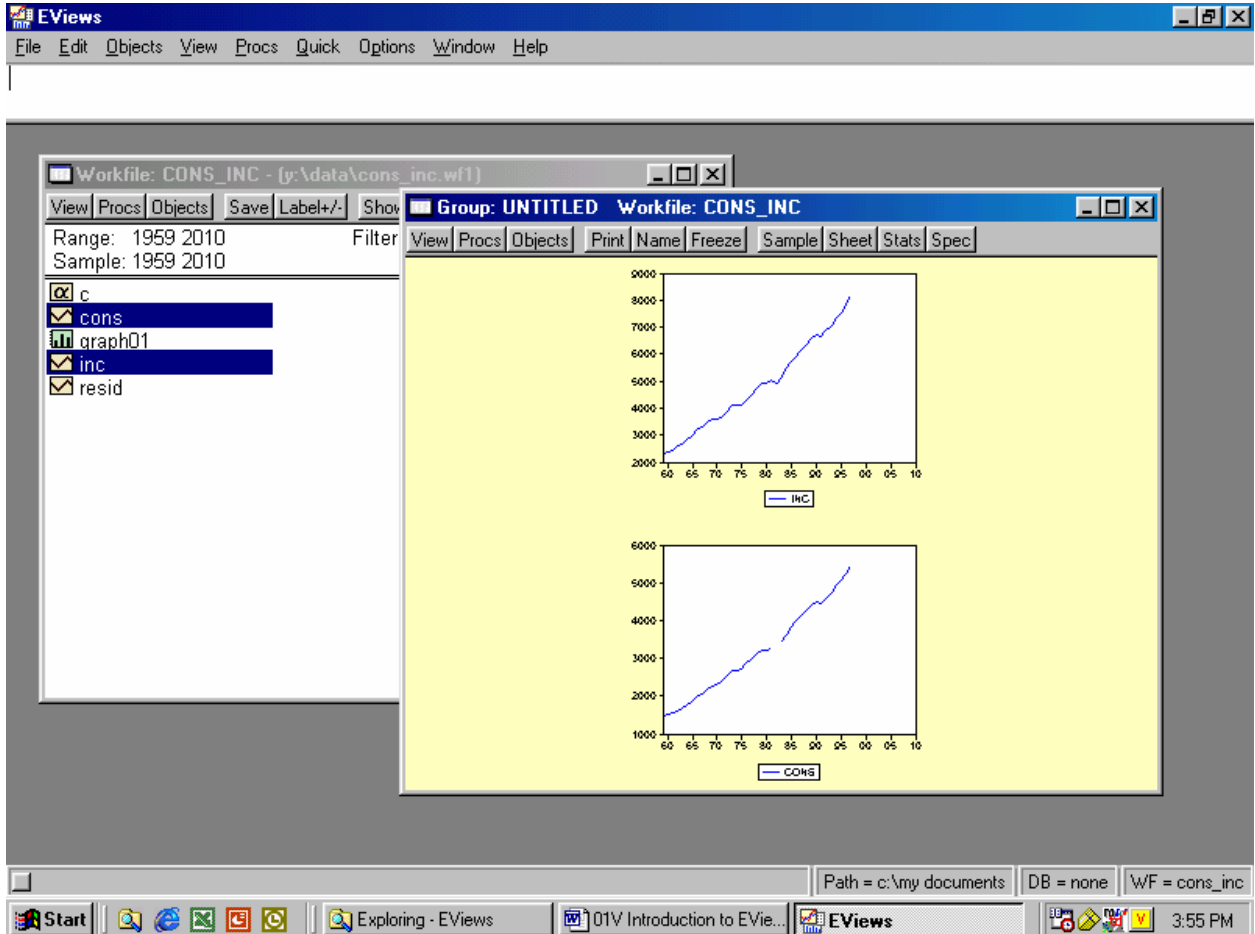
Đóng nhóm này lại .

Bây giờ hãy nhấn đúp vào **cons\_inc** và kiểm tra bảng tính. Nó cũng đã được sửa lại.

Bây giờ hãy kiểm tra một số dạng khác:

## Nhấp View/Multiple Graphs/Line

Có công dụng vẽ đồ thị từng chuỗi theo thời gian

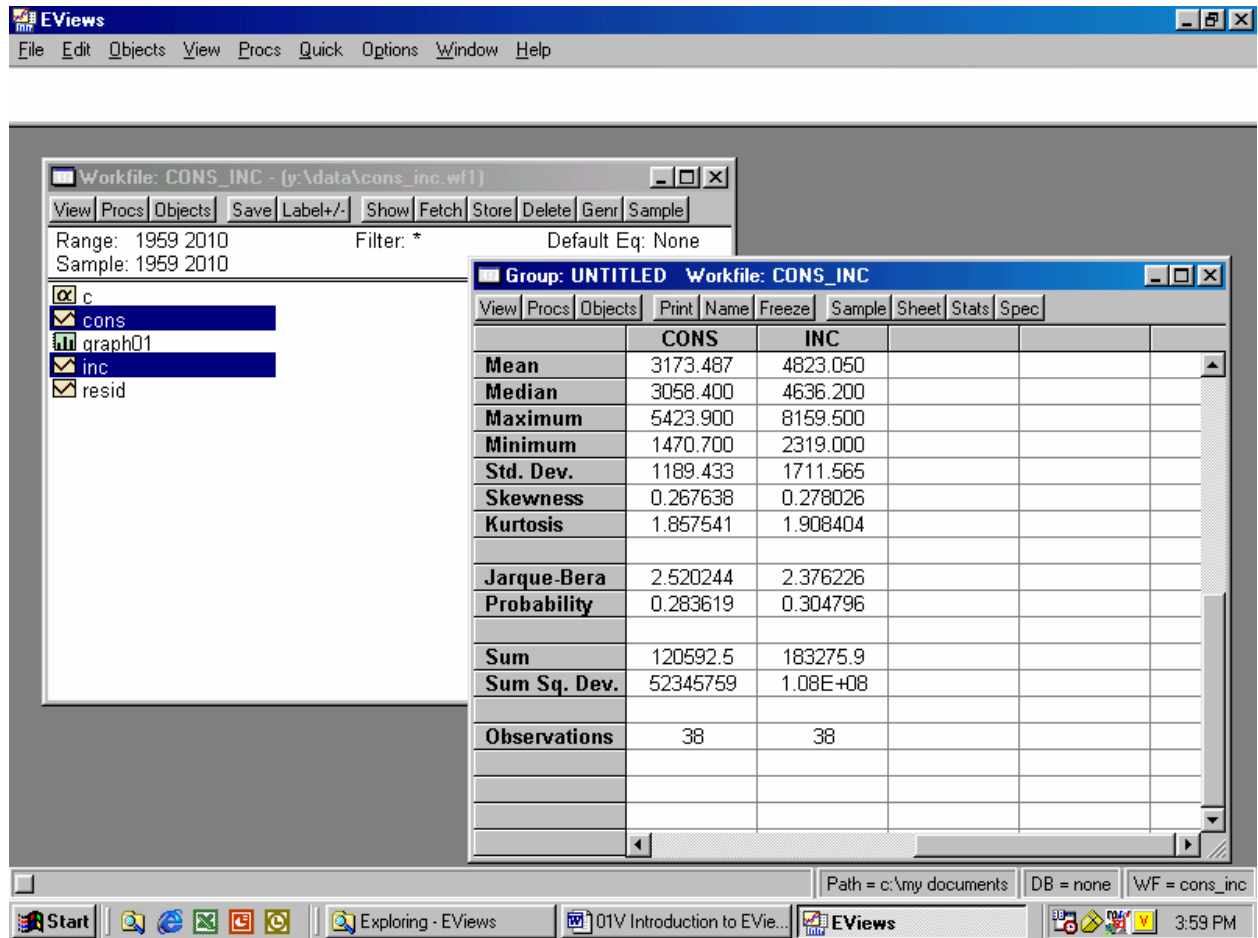


## Tìm các tiêu thức thống kê của mẫu

### Nhập View/Descriptive Statistics/Common Sample

The screenshot shows the EViews software interface. The main window is titled 'Workfile: CONS\_INC - (y:\data\cons\_inc.wf1)'. It displays a list of objects on the left: 'c', 'cons', 'graph01', 'inc', and 'resid'. The 'cons' object is selected. The 'View' menu is open, showing 'Descriptive Stats' with a sub-menu containing 'Common Sample' and 'Individual Samples'. The 'Common Sample' option is highlighted. The background shows a data table with columns 'INC' and 'Label'. The table contains data for the years 1975, 1976, and 1977.

Year	INC	Label
1975	2710.900	4084.400
1976	2868.900	4311.700
1977		

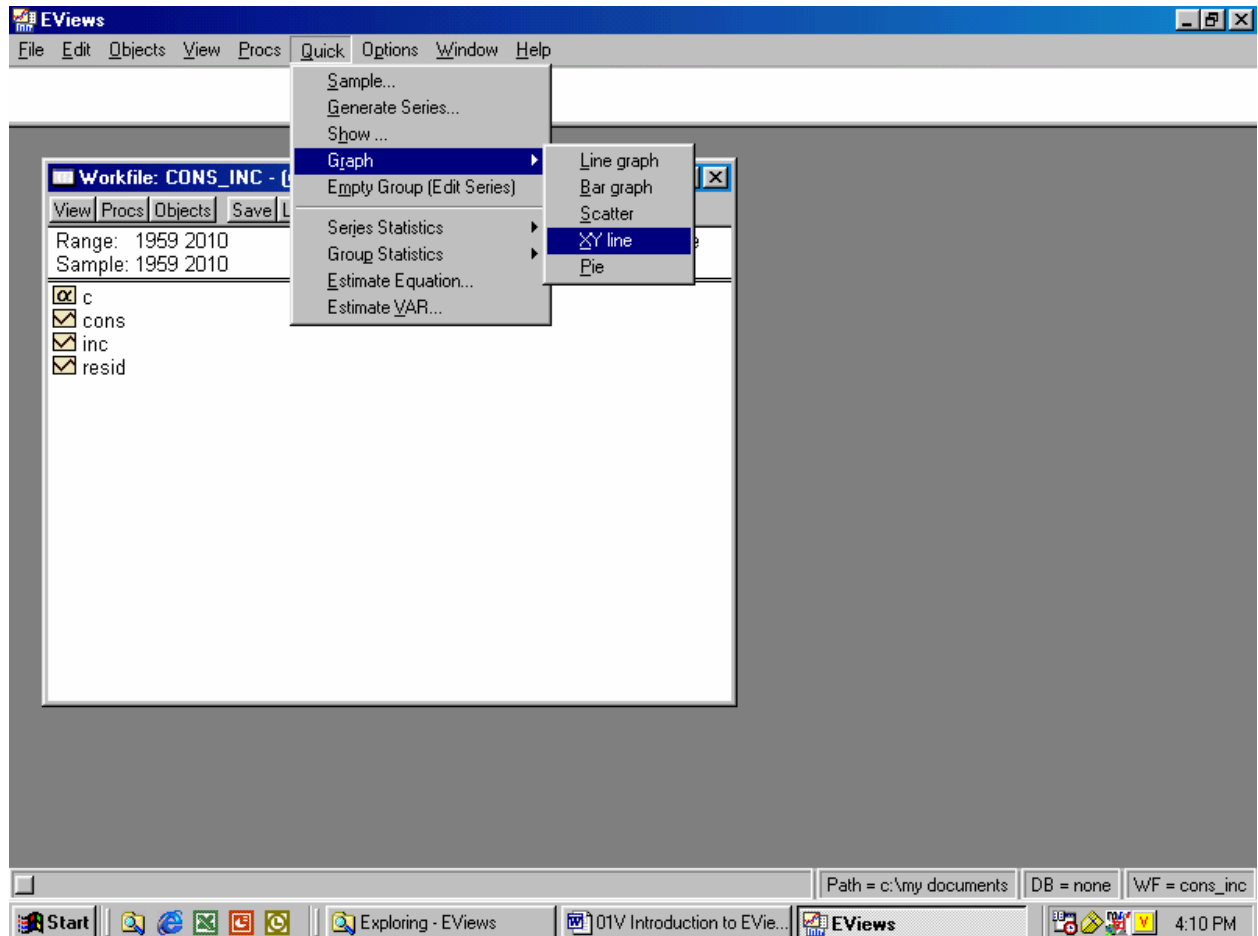




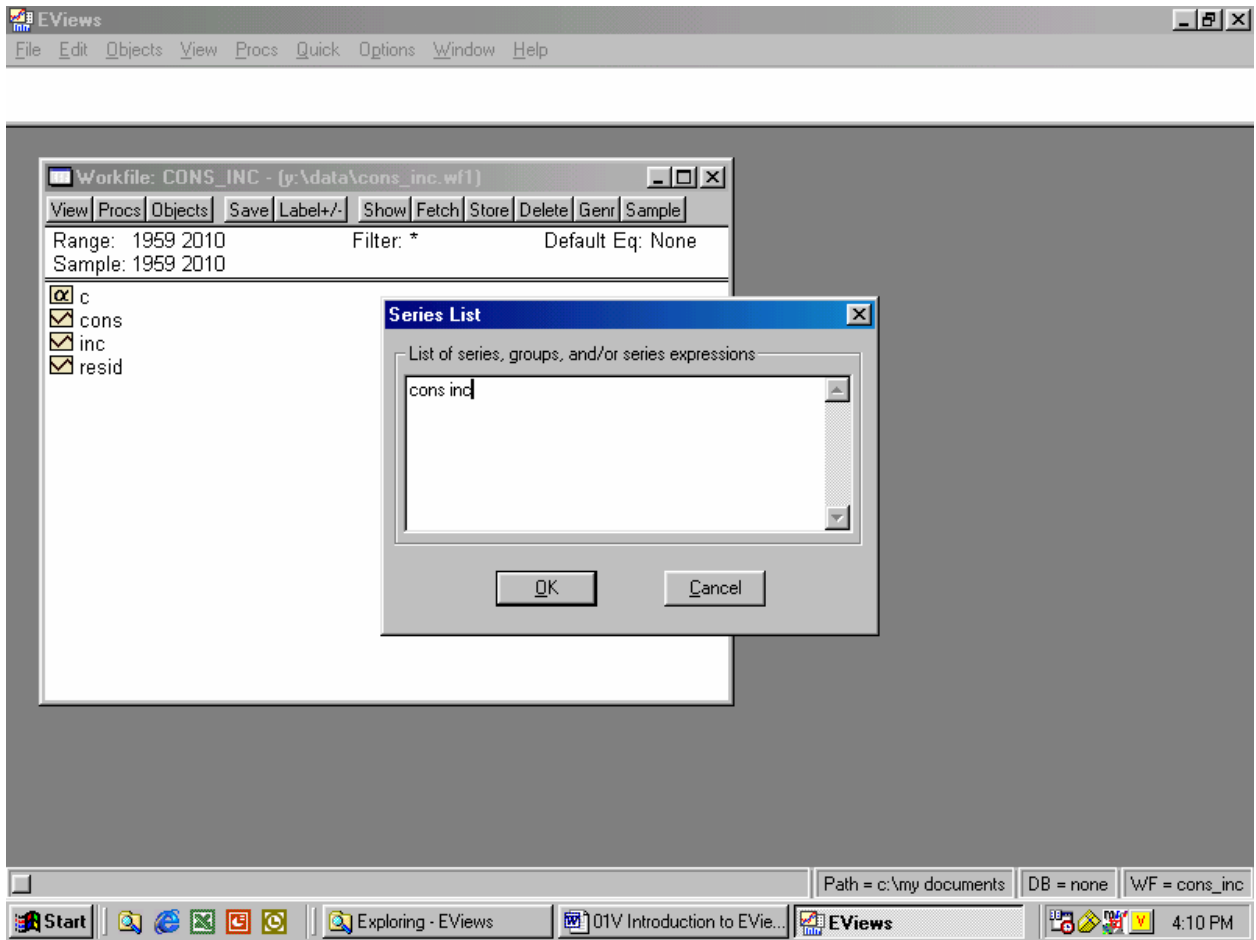


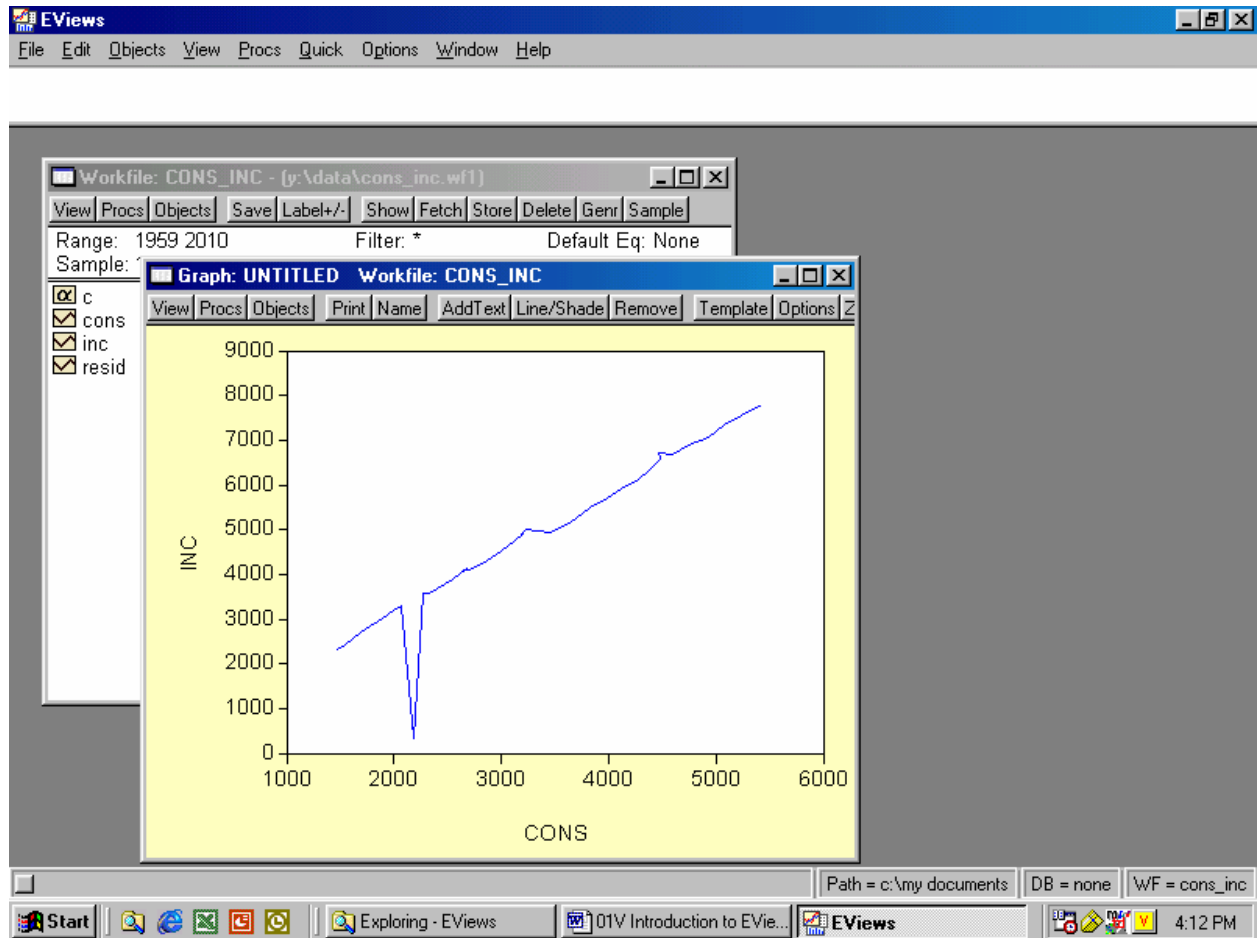
## Vẽ đồ thị không cần mở chuỗi dưới dạng bảng

Bây giờ hãy trở lại workfile. Hãy vào menu chính và nhấp vào **Quick/Graph** dùng để vẽ đồ thị linh động hơn. Cách vẽ này không cần phải mở chuỗi dữ liệu như cách đã vẽ ở trên.



Một cửa sổ mở ra yêu cầu Anh/Chị xác định chuỗi hoặc các nhóm; nó chỉ ra **cons\_inc**; Anh/Chị cần đánh máy **cons inc** và nhấn **OK**. Một cửa sổ mở ra cho phép Anh/Chị lựa chọn nhiều phương án vẽ đồ thị khác nhau. Nhấn vào phím ? phía dưới **Graph Type**, chọn **Scatter Diagram**, và nhấn **OK**. Biện mà Anh/Chị liệt kê đầu tiên được vẽ trên trục hoành. Hãy lưu ý tới sự đa dạng của các lựa chọn trong menu.

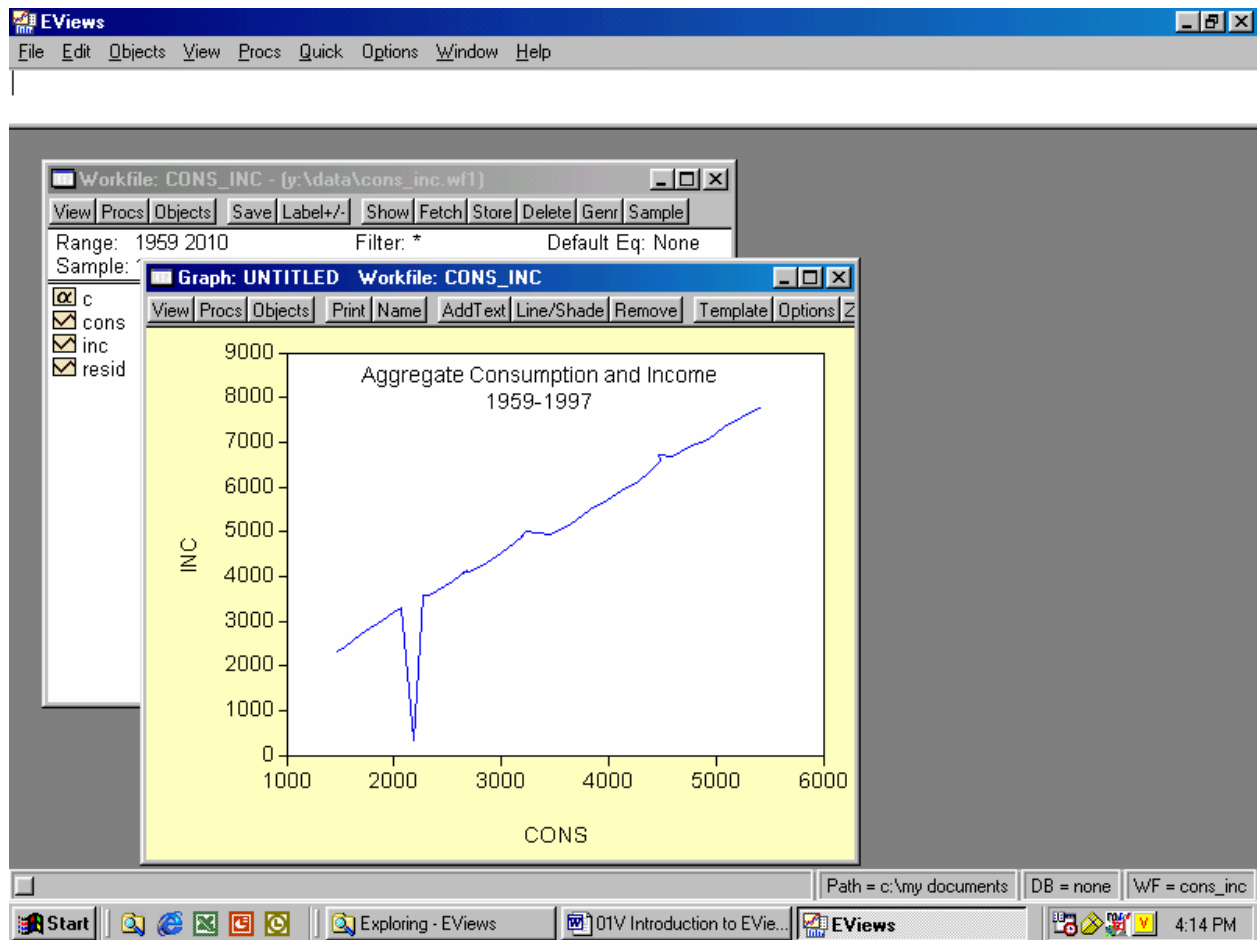




Nhấp **Add Text** và đánh máy

### **Aggregate Consumption and Income 1959 - 1997**

vào khoảng trống đã cho. Nhấn **OK**. Bây giờ, nhấp **Name** và nhấp **OK** dành cho GRAPH1; GRAPH1 được bổ sung vào workfile của Anh/Chị.



Nhấp **Objects/Freeze Output** trong menu đồ thị và đặt cho đồ thị tên **Graph\_2**. Bây giờ workfile của Anh/Chị chứa hai đồ thị lâu dài: **Graph\_1** (phản ánh dữ liệu gốc) và **Graph\_2** phản ánh các hiệu chỉnh.

Bây giờ nhấn **File/Save** và đóng workfile này.

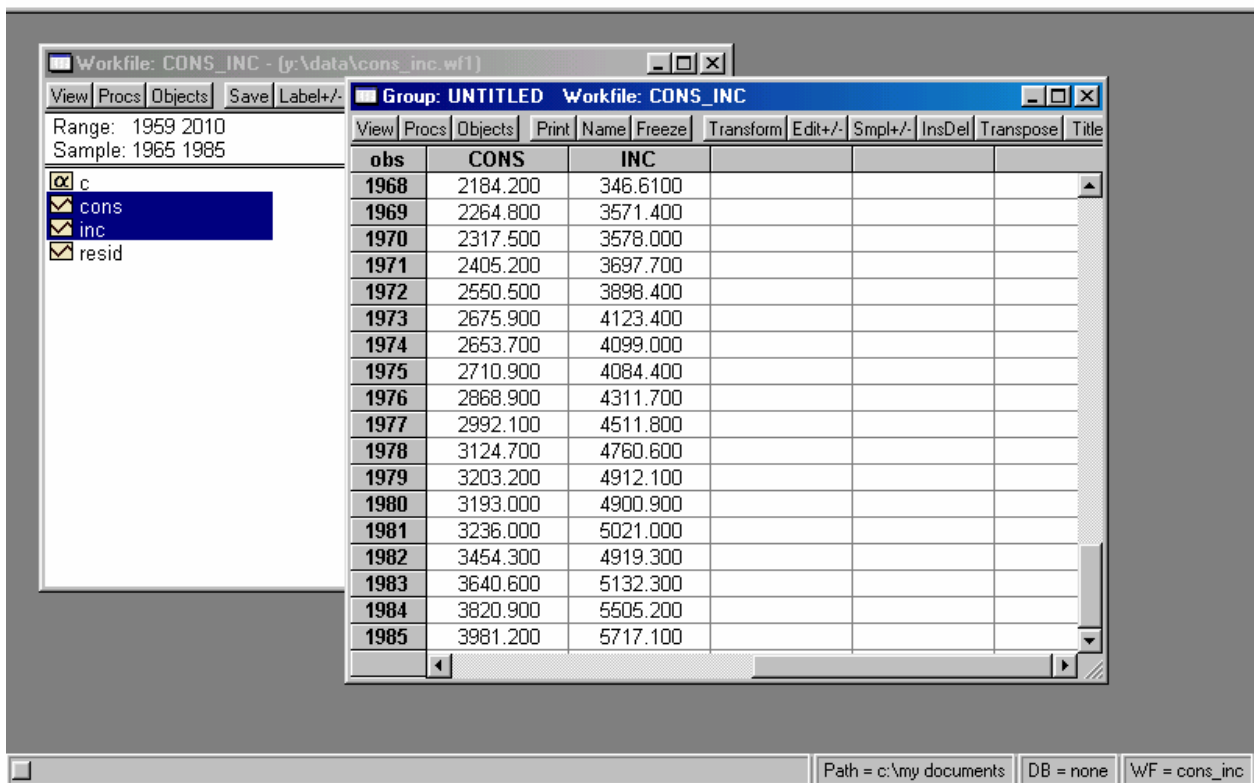
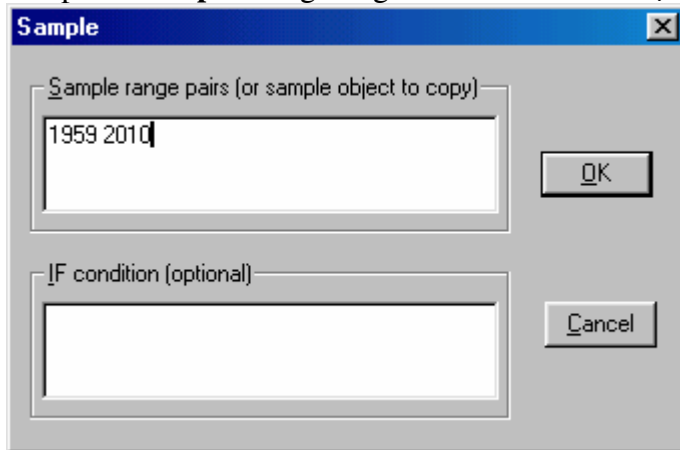
### Lựa chọn mẫu dữ liệu

Lựa chọn mẫu dữ liệu là chọn giai đoạn nghiên cứu trong dữ liệu hiện có trong workfile

Chúng ta thường muốn tập trung phân tích vào một mẫu phụ của dữ liệu. EViews cho phép chúng ta làm điều này theo hai cách: (1) bằng cách xác định khoảng mẫu mà chúng ta muốn xem xét, (2) hoặc là bằng cách xác định các quan sát thoả mãn một điều kiện logic (if) nào đó. Nếu chúng ta sử dụng cả hai phương pháp, thì mẫu kết quả là giao của các mẫu được tạo ra bởi hai phương pháp này.

Nhấp **File/Open/Workfile** sau đó nhấp đúp vào **cons\_inc.wf1**.

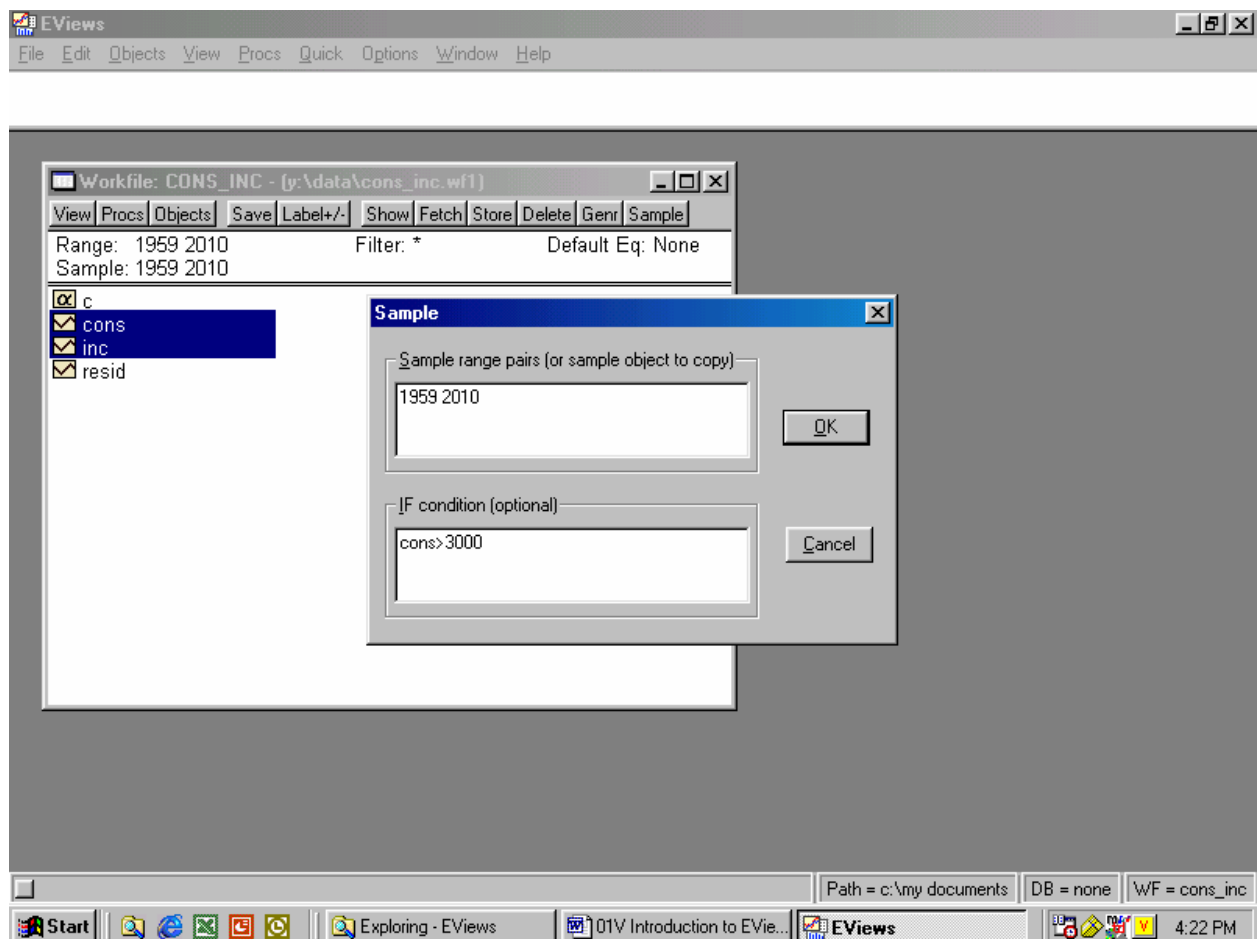
Nhấp nút **Sample** trong bảng tính và đánh vào một mẫu mới: **1965 1985**.



Mở **cons\_inc** thành bảng tính, sau đó nhấp **View/Graph/Line**. Hãy so sánh đồ thị này với đồ thị có tên **Graph\_2**.

Đóng đồ thị này và nhấp **Sample** trên menu của workfile và khôi phục toàn bộ mẫu **1959 1997**.

Chúng ta cũng có thể muốn xác định các quan sát thỏa mãn một tiêu chí nào đó. Ví dụ, chúng ta có thể muốn lựa chọn tập hợp dữ liệu con trong đó **cons** lớn hơn 3000. Trong trường hợp này, chúng ta nhấp **Sample** trên menu của workfile này; sau đó trong cửa sổ có tên **If condition** đánh máy **cons > 3000** và nhấp **OK**.



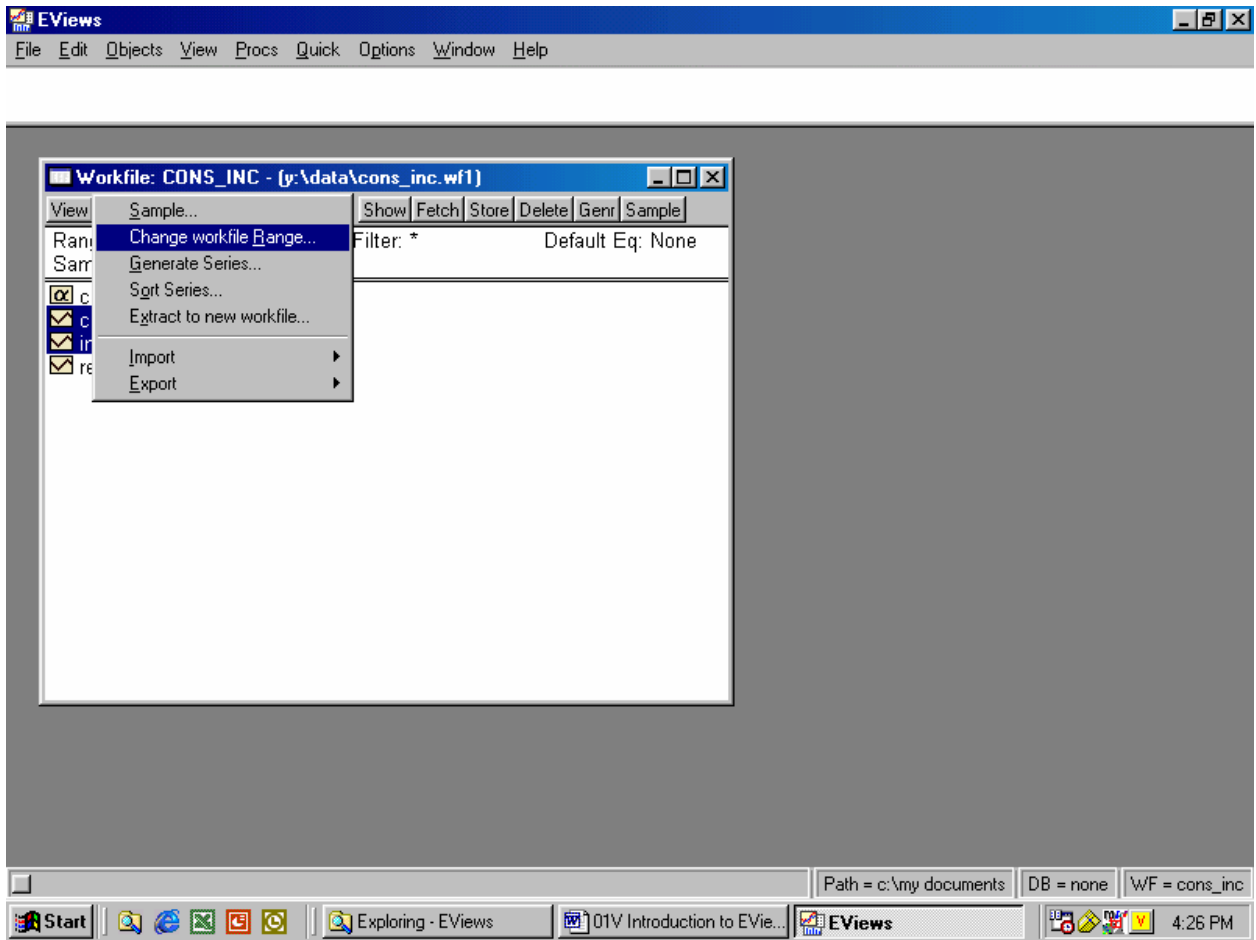
The screenshot shows the EViews software interface. The main window displays a workfile named 'CONS\_INC' with a range of 1959-2010. The data is presented in a table with columns for 'obs', 'CONS', and 'INC'. The data points are as follows:

obs	CONS	INC
1978	3124.700	4760.600
1979	3203.200	4912.100
1980	3193.000	4900.900
1981	3236.000	5021.000
1982	3454.300	4919.300
1983	3640.600	5132.300
1984	3820.900	5505.200
1985	3981.200	5717.100
1986	4113.400	5912.400
1987	4279.500	6113.300
1988	4393.700	6368.400
1989	4474.500	6591.800
1990	4466.600	6707.900
1991	4594.500	6676.400
1992	4748.900	6880.000
1993	4928.100	7062.600
1994	5075.600	7347.700
1995		

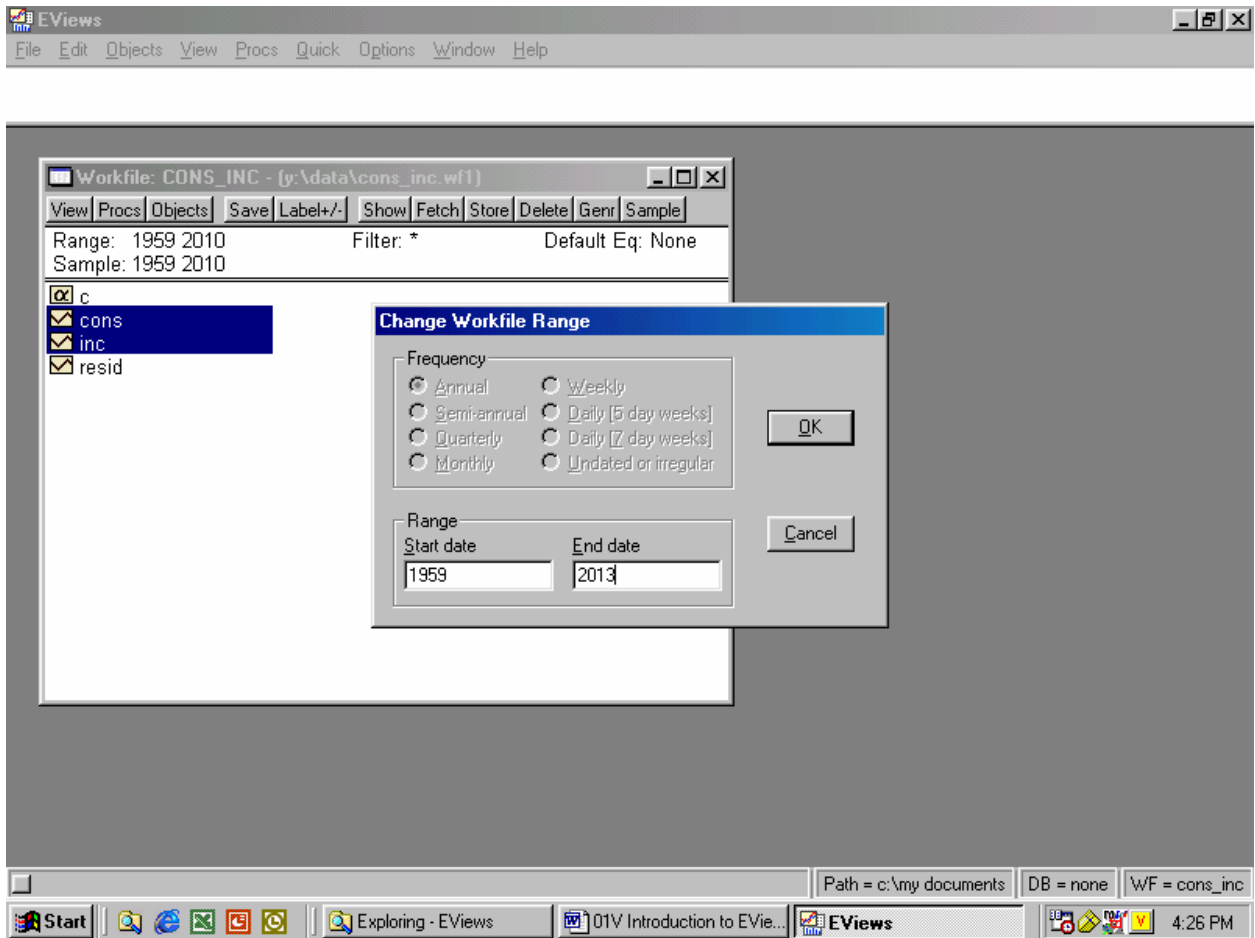
Bây giờ hãy kiểm tra cửa sổ tình trạng của workfile, dạng của bảng tình dành cho cons\_inc, và đồ thị dạng đường của **cons** và **inc**.

### Mở rộng Workfile

Hãy tưởng tượng rằng chúng ta mong muốn nhận được dữ liệu bổ sung của 3 năm cho workfile này. Chúng ta có thể mở rộng workfile này với mục đích bổ sung thêm dữ liệu này.







Nhấp **Proc/Change Workfile Range** và đánh máy một **End Date (dữ liệu cuối)** mới. Bây giờ Anh/Chị có thể biên tập bảng tính này để bổ sung thêm dữ liệu mới. Nên nhớ rằng khi muốn nhập thêm dữ liệu bổ xung, thì sau lệnh này chúng ta cũng cần mở rộng thêm cả phạm vi mẫu bằng lệnh **sample**

obs	CONS	INC
1996	5423.900	7813.200
1997	8159.500	
1998	NA	NA
1999	NA	NA
2000	NA	NA
2001	NA	NA
2002	NA	NA
2003	NA	NA
2004	NA	NA
2005	NA	NA
2006	NA	NA
2007	NA	NA
2008	NA	NA
2009	NA	NA
2010	NA	NA
2011	NA	NA
2012	NA	NA
2013	NA	NA

### Copy và dán dữ liệu mới

Hãy tưởng tượng rằng chúng ta có được một tập tin Excel có tên **Cons\_Inc.xls** bao gồm dữ liệu từ 1959 - 2000.

Hãy mở tập tin này, bôi đen dữ liệu dành cho năm 1998, 1999, và 2000, sau đó nhấn **Ctrl-C**, sau đó mở chức năng Edit trong EViews, đặt con trỏ vào ô trên cùng bên trái của dữ liệu MỚI, sau đó dán (**Ctrl-V**).

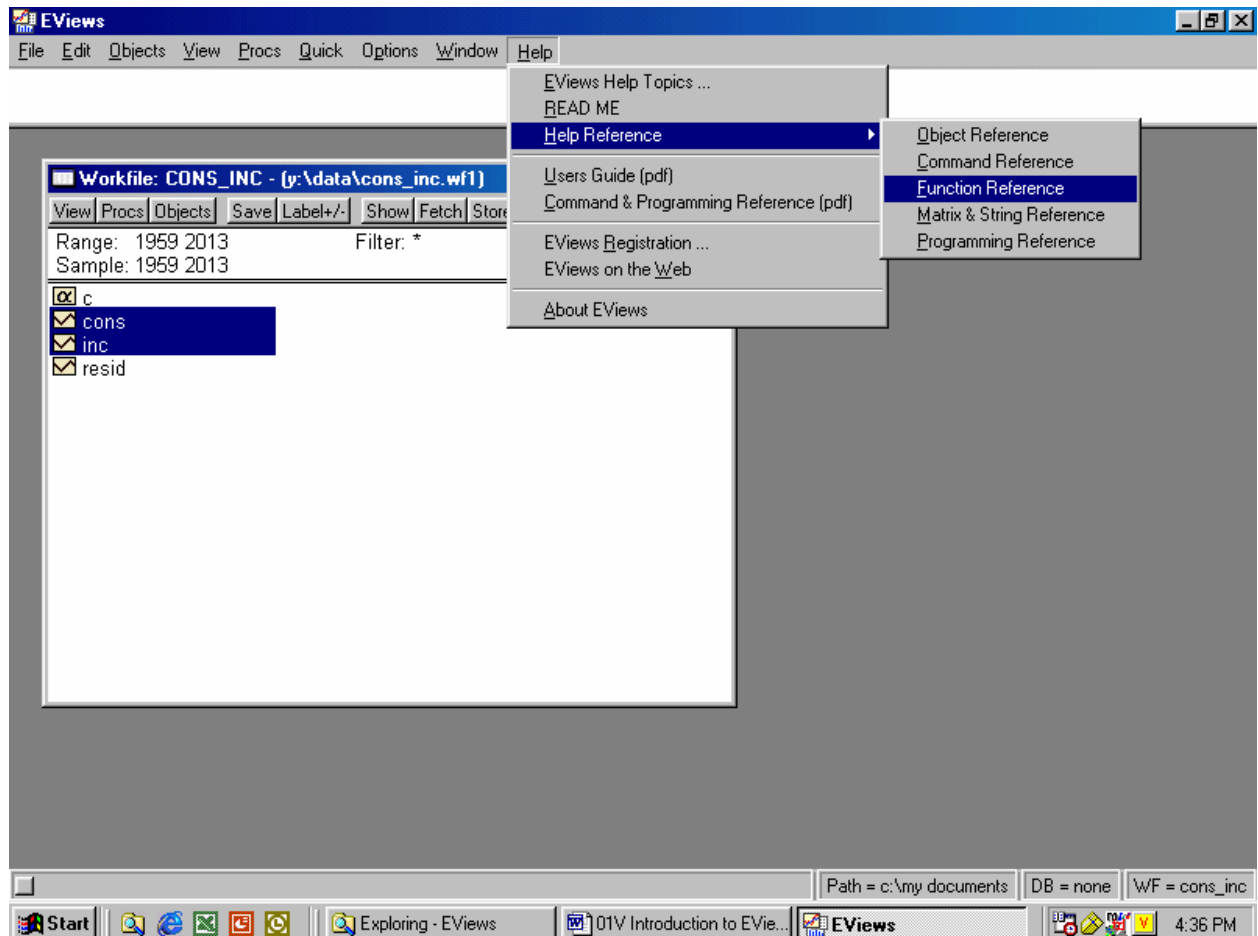
Dữ liệu mới này nhảy vào các ô mong muốn.

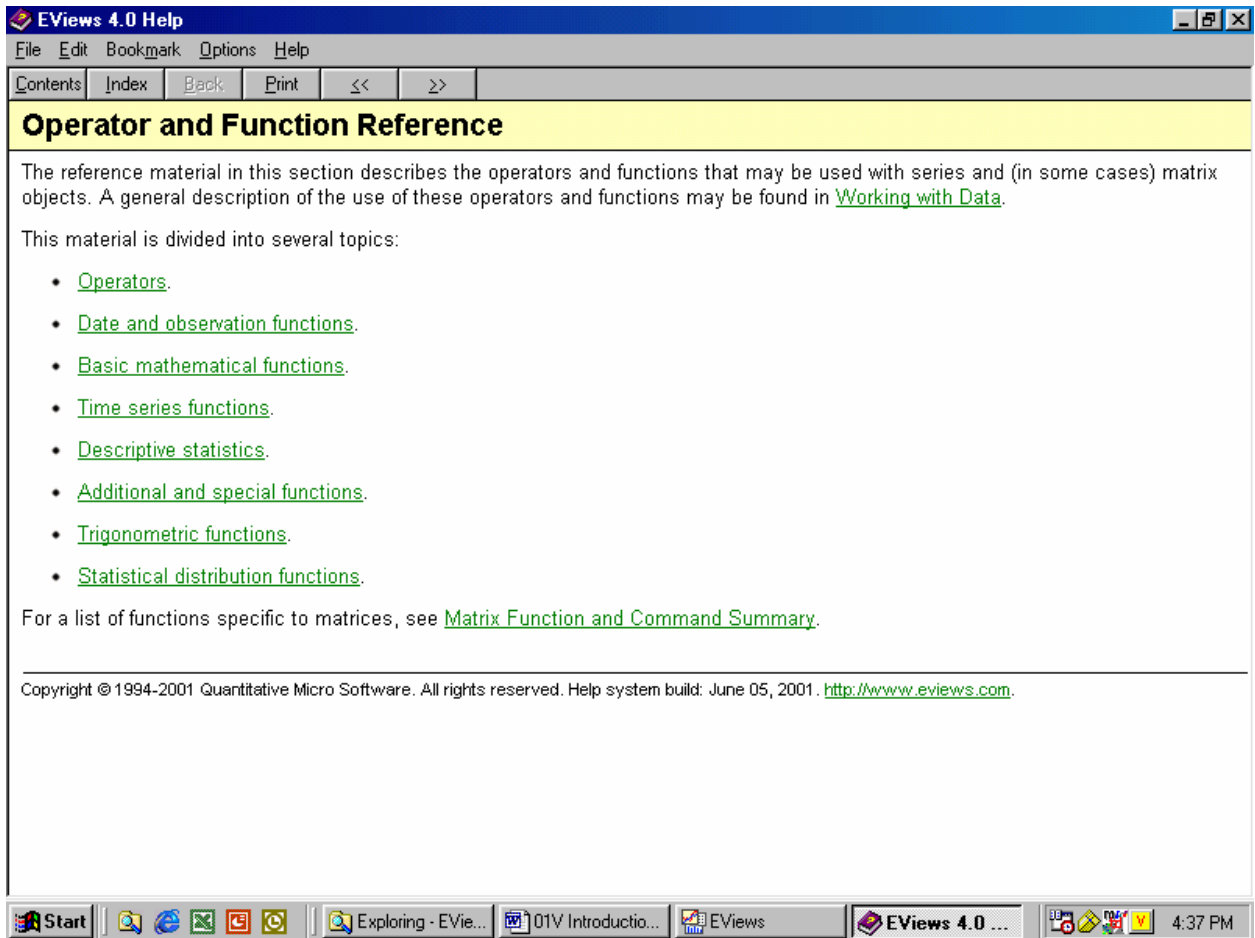
Lưu ý rằng các cột phải ở trong thứ tự chính xác.

## Tạo các biến mới, các độ trễ, các sai phân và các biến giả

### Tạo các biến mới

Nếu chúng ta muốn tạo ra biến mới là một hàm số của các biến hiện hữu, thì chúng ta sử dụng chức năng **Genr** trên **EViews**. Ví dụ, để tạo ra **Y** như một logarit tự nhiên của **x** chúng ta sẽ nhấp vào **Genr** trên menu của workfile, sau đó đánh máy **Y = log(x)**. EViews có một số lớn các chức năng mà Anh/Chị có thể khám phá dưới địa chỉ **Help/Function Reference**.





**EViews 4.0 Help**

File Edit Bookmark Options Help

Contents Index Back Print << >>

## Operators

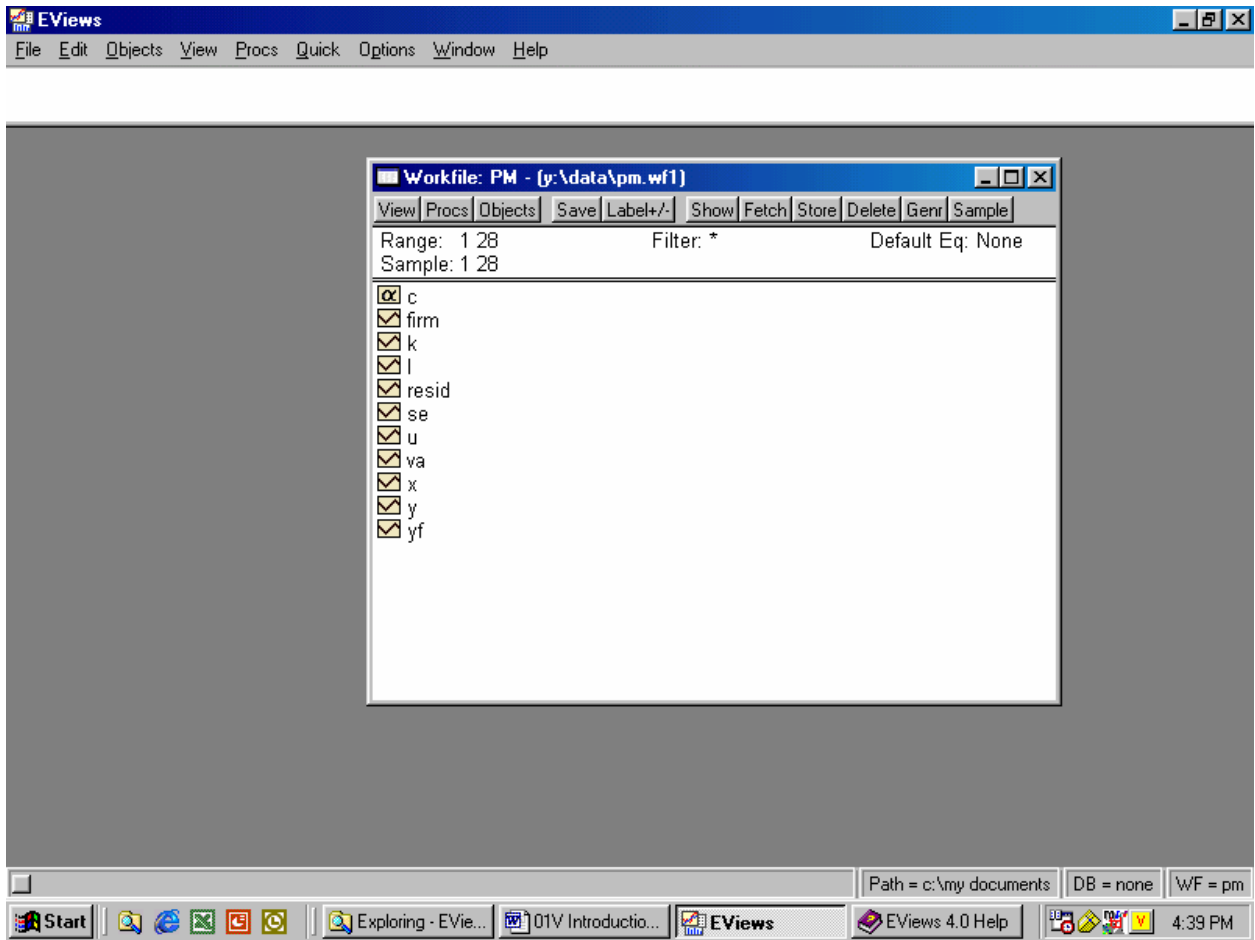
All of the operators described below may be used in expressions involving series and scalar values. When applied to a series expression, the operation is performed for each observation in the current sample. The precedence of evaluation is listed in [Operators](#). Note that you can enforce order-of-evaluation using parentheses.

Expression	Operator	Description
+	add	$x+y$ adds the contents of X and Y.
-	subtract	$x-y$ subtracts the contents of Y from X.
*	multiply	$x*y$ multiplies the contents of X by Y.
/	divide	$x/y$ divides the contents of X by Y.
^	raise to the power	$x^y$ raises X to the power of Y.
>	greater than	$x>y$ takes the value 1 if X exceeds Y, and 0 otherwise.
<	less than	$x<y$ takes the value 1 if Y exceeds X, and 0 otherwise.
=	equal to	$x=y$ takes the value 1 if X and Y are equal, and 0 otherwise.
<>	not equal to	$x<>y$ takes the value 1 if X and Y are not equal, and 0 if they are equal.
<=	less than or equal to	$x<=y$ takes the value 1 if X does not exceed Y, and 0 otherwise.
>=	greater than	$x>=y$ takes the value 1 if Y does not

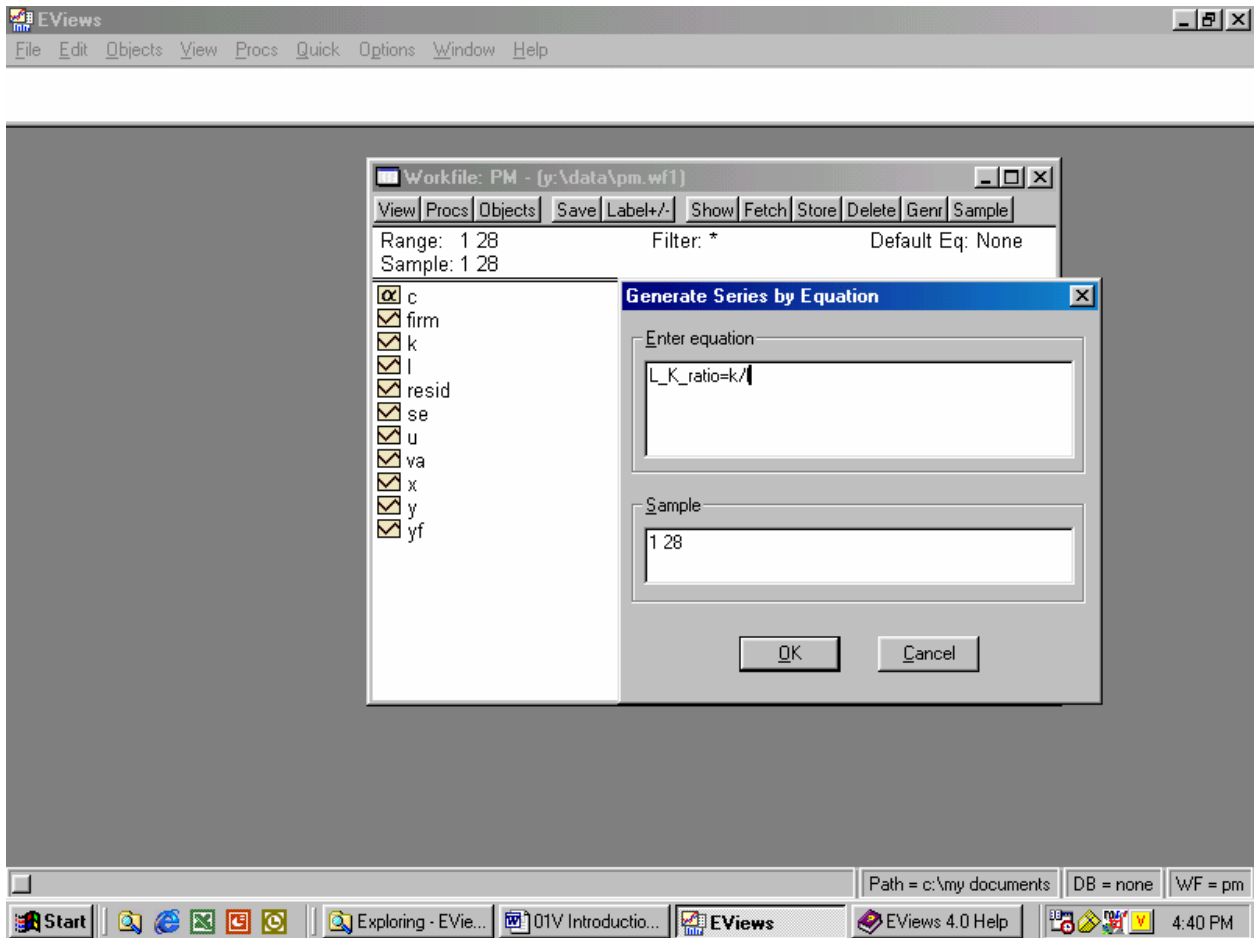
Lưu ý rằng Anh/Chị không bắt buộc phải tạo ra các biến mà Anh/Chị định sử dụng trong một phép hồi qui (hoặc qui trình khác); biểu thức có thể được đưa trực tiếp vào phần xác định phương trình hồi qui. Đặc tính này giúp ta giữ lại một workfile nhỏ bé. Chúng ta sẽ sớm minh họa điều này.

Giả sử chúng ta quan tâm tới tỉ lệ nhân công/vốn đối với 27 hãng tạo thành tập hợp dữ liệu các loại kim sơ cấp.

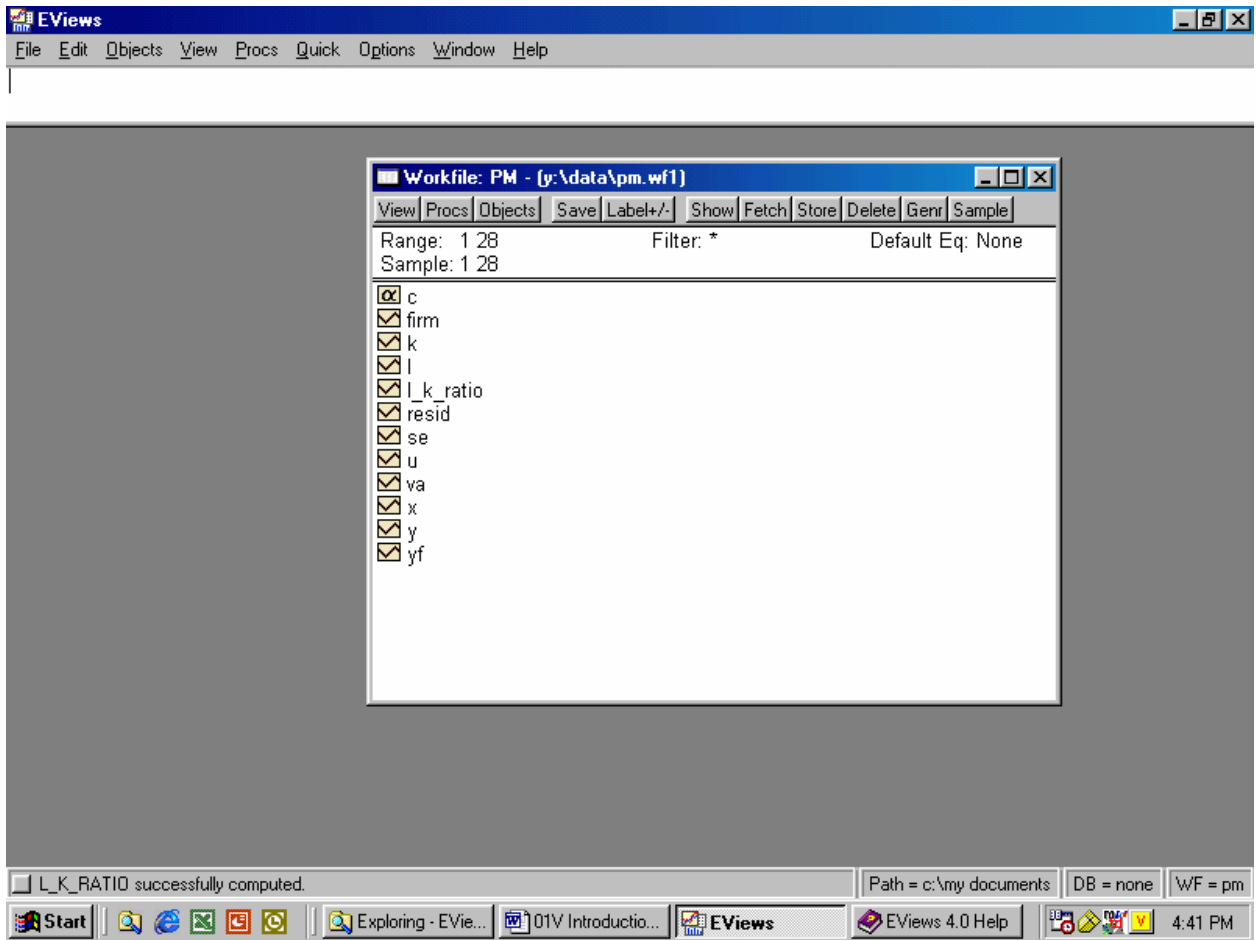
Hãy mở **pm.wf1**



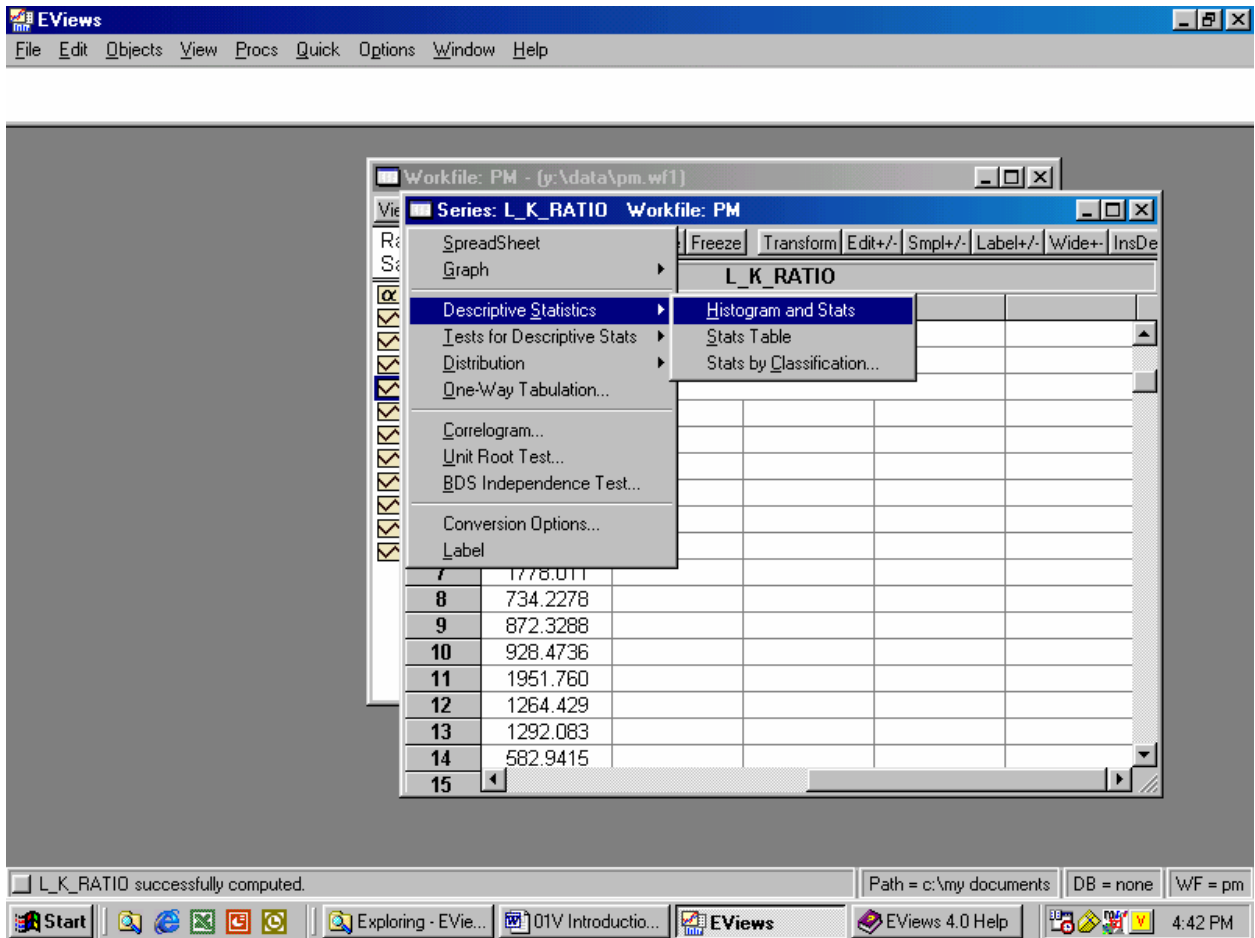
Nhấp **Genr** và đánh máy **L\_K\_Ratio = L / K**



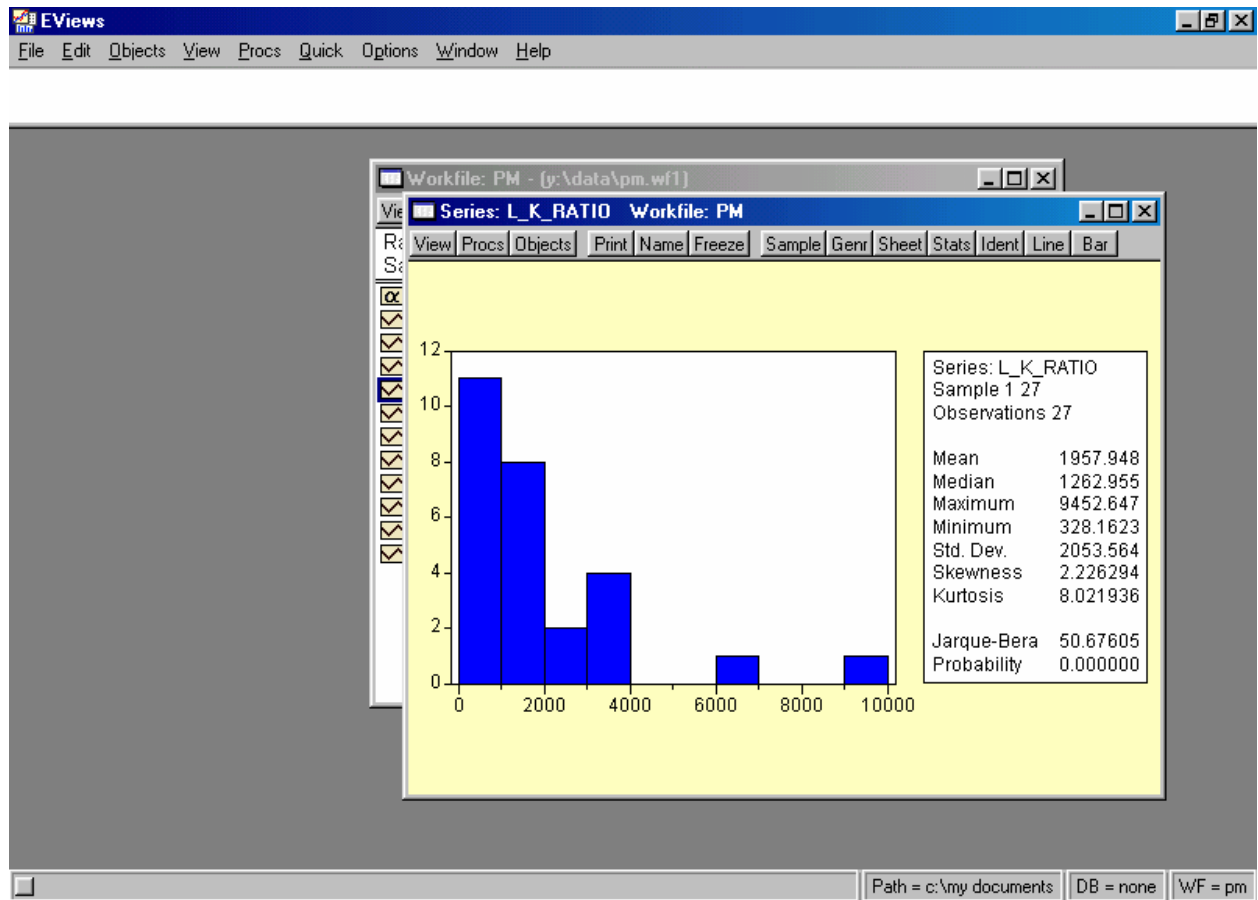
Nhấp đúp vào **L\_K\_RATIO** và chọn **View / Descriptive Statistics / Histogram and Stats**.







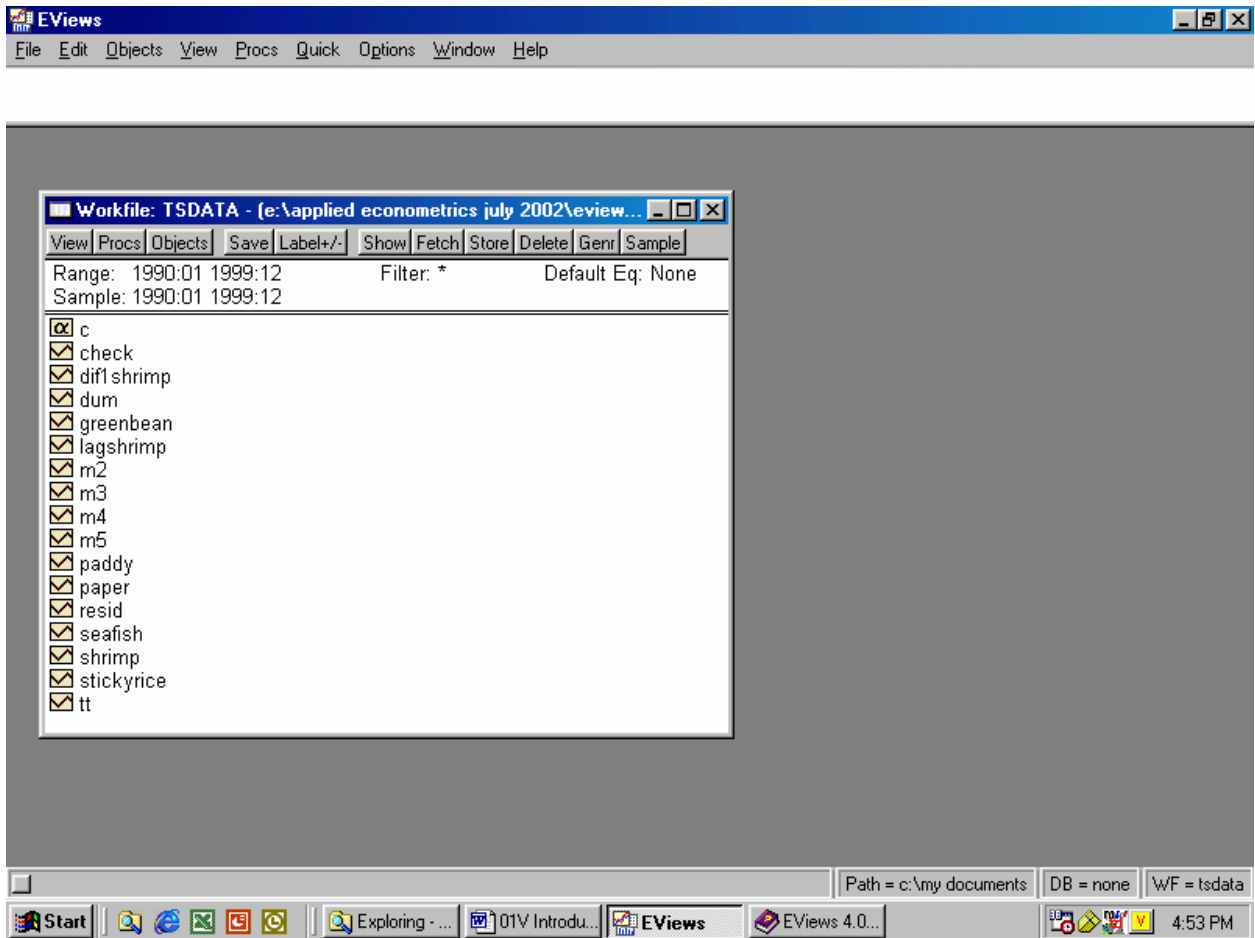
Điều này cho thấy chúng ta có thể kiểm tra phân bố xác suất của một biến đơn như thế nào; cửa sổ này sẽ rất hữu ích khi chúng ta xem xét các phần dư hồi qui .



### Tạo biến độ trễ, sai phân và xu hướng

Trong nhiều ứng dụng chuỗi thời gian chúng ta đã sử dụng các biến có độ trễ, các sai phân bậc I (và bậc cao hơn), các biến giả có tính mùa vụ (hoặc theo tháng, quý), và các xu hướng thời gian. EViews có các lệnh rất dễ để tạo ra chuỗi theo ý muốn.

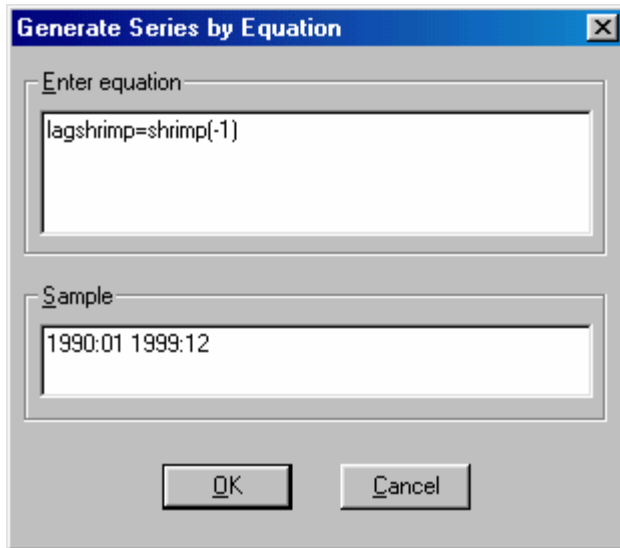
Workfile **tsdata.wf1** có chứa dữ liệu về giá của một số mặt hàng nông sản của Việt Nam.



Quan sát đầu tiên là 1990:01 và quan sát cuối cùng là 1999:12.

Giá trị có độ trễ của **tôm** có thể được tạo ra bằng cách nhấp **Genr**, sau đó đánh máy vào phương trình:

$$\text{LAGSHRIMP} = \text{SHRIMP}(-1)$$



Tất nhiên, Anh/Chị có thể đặt cho biến này bất cứ tên nào mà Anh/Chị thích ; Tôi đã chọn một tên gợi gợi cho tôi nhớ tới định nghĩa của biến này.

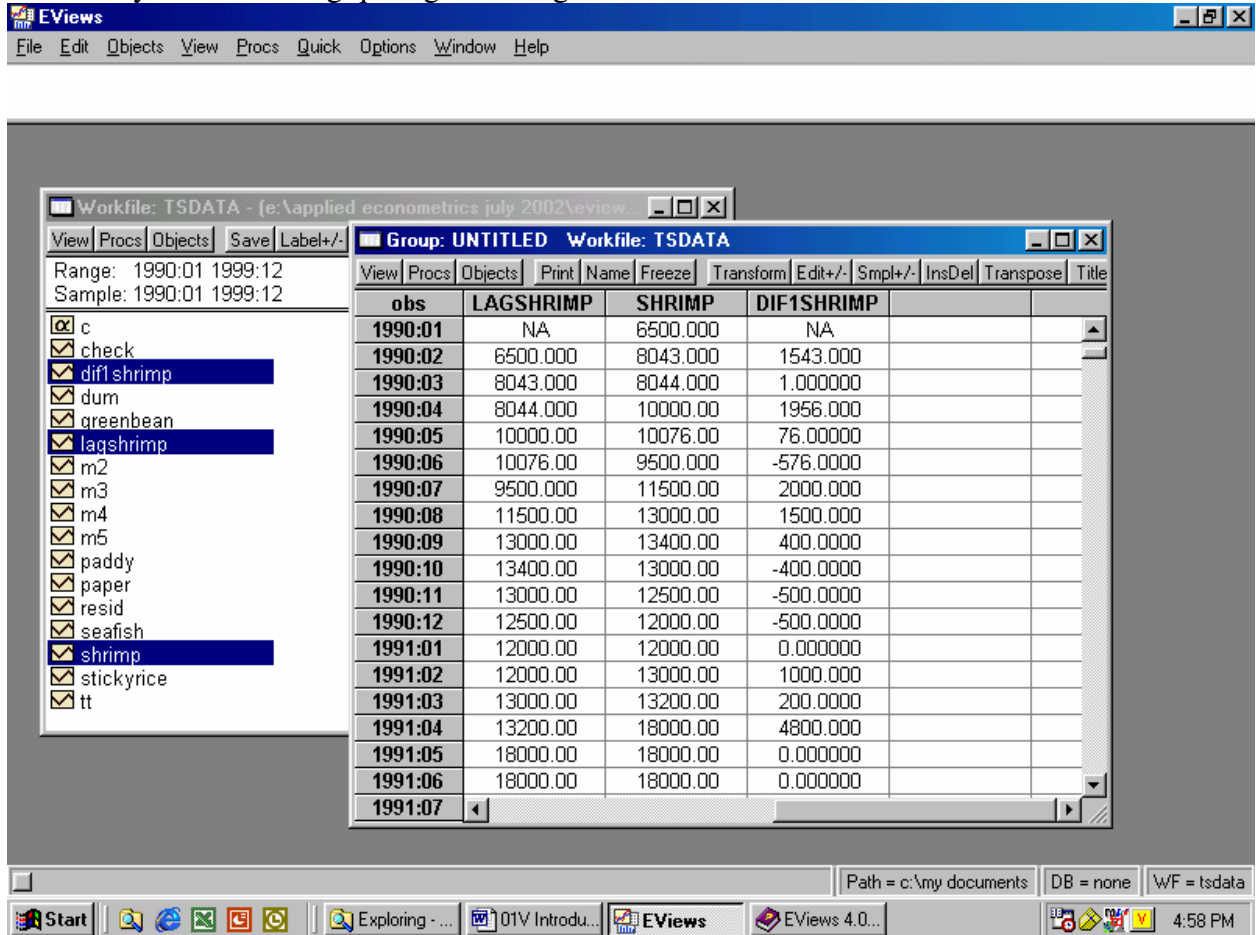
Để tạo ra sai phân thứ nhất của giá tôm, hãy nhấp **Genr**, và đưa vào :



Hoặc là

$$\mathbf{DIF1SHRIMP = SHRIMP - SHRIMP(-1)}$$

Bôi đen **SHRIMP**, **LAGSHRIMP** và **DIF1SHRIMP** và kiểm tra dạng bảng tính của chúng; rất dễ thấy các mối tương quan giữa chúng.



Các độ trễ có độ dài  $j$  có thể được tạo ra bằng cách sử dụng **SHRIMP(-j)** và các sai phân bậc  $j$  có thể được tạo ra bằng cách sử dụng

$$D_j\text{SHRIMP} = \text{SHRIMP} - \text{SHRIMP}(-j).$$

Các sai phân bậc cao hơn được tạo ra bởi toán tử sai phân **D(SHRIMP, j)**. Ví dụ nếu  $j = 2$ , thì toán tử này sẽ tạo ra sai phân của các sai phân bậc nhất.

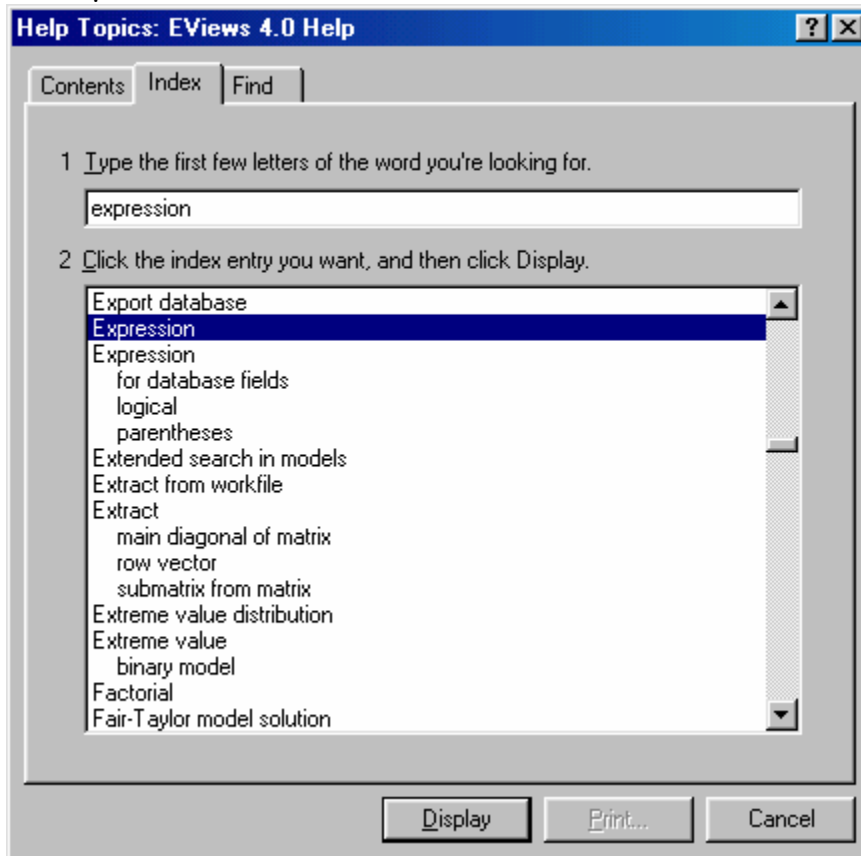
$$D(\text{SHRIMP}, 2) = [\text{SHRIMP} - \text{SHRIMP}(-1)] - [\text{SHRIMP}(-1) - \text{SHRIMP}(-2)]$$

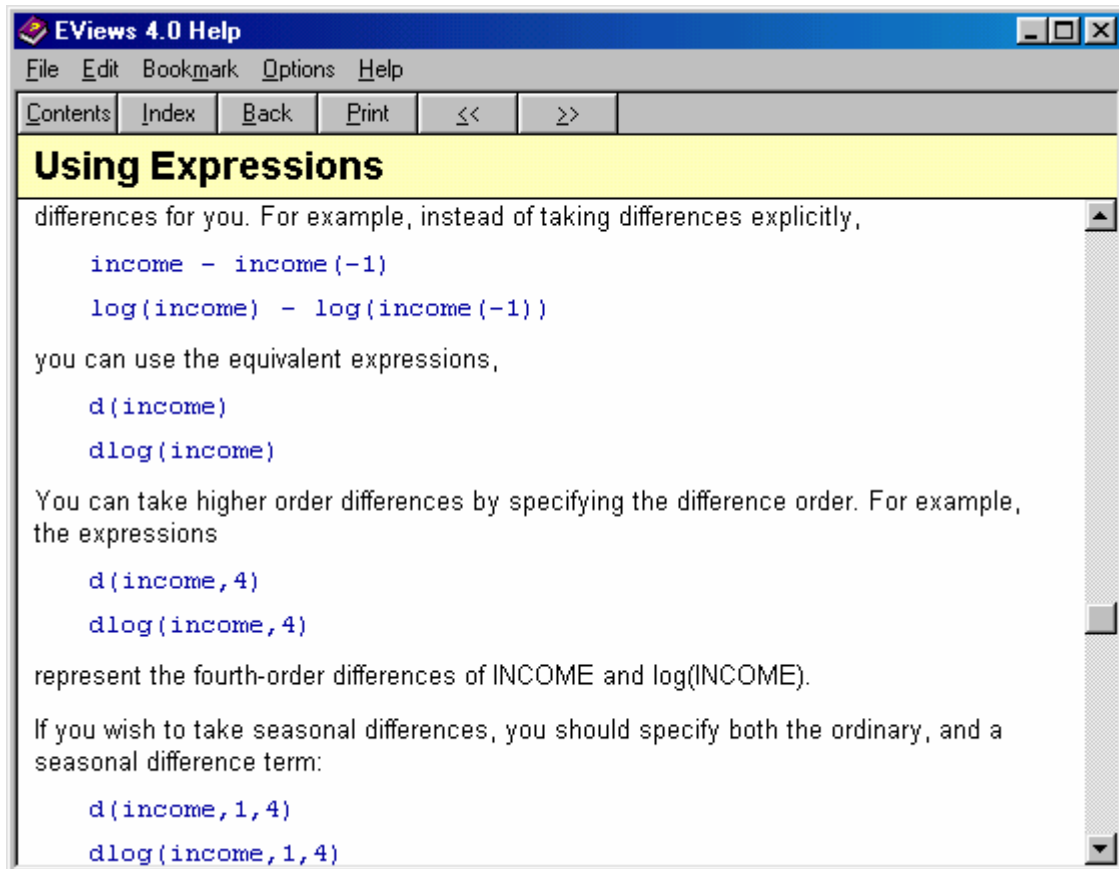
### Công thức tạo biến mới chi tiết trong EViews

Nhằm đi xa hơn nữa các công thức tạo biến phức tạp, nếu không có hướng dẫn của giáo viên hoặc các chuyên gia phần mềm Eviews anh (chị) cần sử dụng lệnh sau đây:

## Help / EViews Help Topics / Index

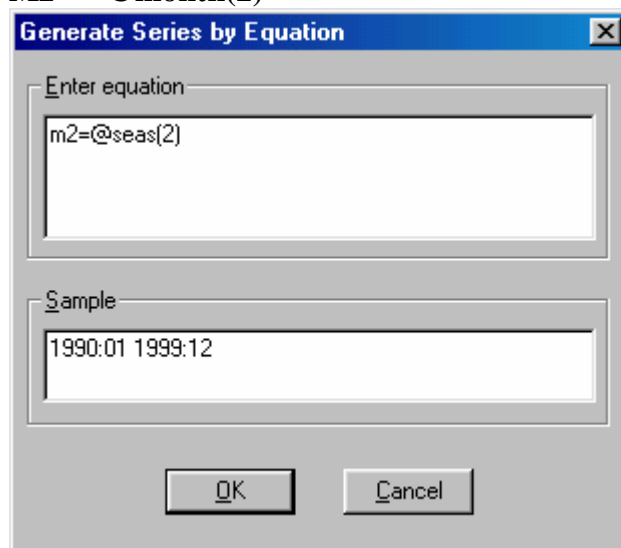
Sau đó đánh vào Expressions. Khi chữ Expressions được bôi đen, nhấn **Display** và đọc các chỉ dẫn hiện ra.





**EViews** có một số chức năng đặc biệt để tạo ra các biến giả có tính mùa vụ và các xu hướng thời gian. Nhấn **Genr** trên menu của workfile, sau đó đánh máy phương trình

**M2 = @month(2)**



Việc này tạo ra một biến giả theo tháng có giá trị là 1 đối với tháng thứ hai của mỗi năm và giá trị 0 đối với các tháng khác.

Tạo **M3, M4, M5** và **M6** bằng cách sử dụng cùng phương pháp y như vậy, sau đó nhấp **Genr** trên menu của workfile, sau đó đưa vào phương trình

$$TT = @trend(1989:12)$$

Việc này tạo ra một xu hướng thời gian có giá trị là 1 trong tháng đầu của năm 1990 và tăng thêm 1 sau mỗi quý. Khi Anh/Chị sử dụng toán tử **@trend(p)**, tham số *p* chỉ ra giai đoạn mà đối với nó giá trị của biến xu hướng bằng zero.

Hãy bôi đen **M2, M3, M4, M5, M6** và **TT**, rồi kiểm tra dạng bảng tính của nhóm này.

The screenshot shows the EViews software interface. On the left, a list of objects is displayed, with **m2, m3, m4, m5, and tt** selected. The main window shows a data table for the workfile 'TSDATA'.

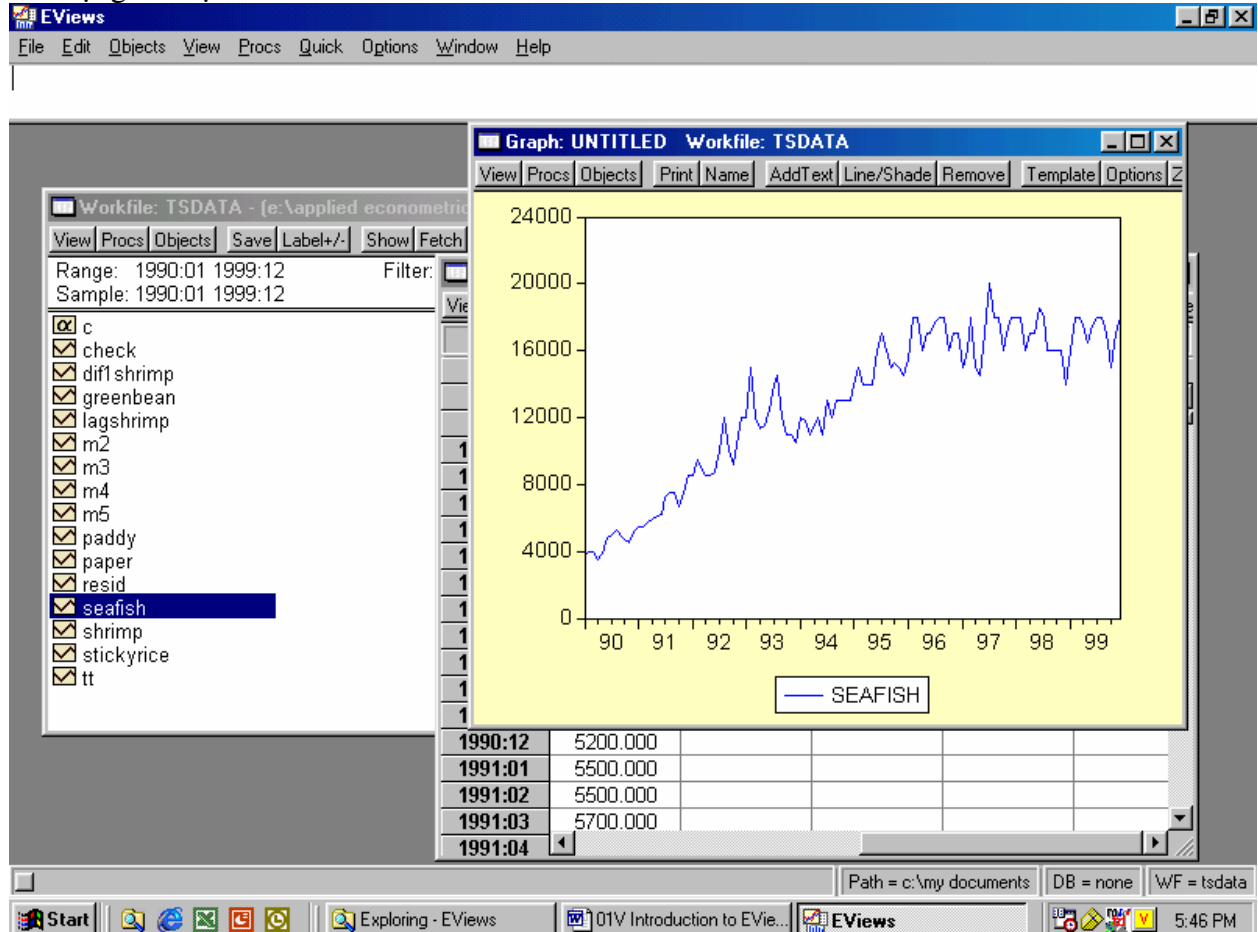
obs	M2	TT	M5	M4	M3
1990:01	0.000000	1.000000	0.000000	0.000000	0.000000
1990:02	1.000000	2.000000	0.000000	0.000000	0.000000
1990:03	0.000000	3.000000	0.000000	0.000000	1.000000
1990:04	0.000000	4.000000	0.000000	1.000000	0.000000
1990:05	0.000000	5.000000	1.000000	0.000000	0.000000
1990:06	0.000000	6.000000	0.000000	0.000000	0.000000
1990:07	0.000000	7.000000	0.000000	0.000000	0.000000
1990:08	0.000000	8.000000	0.000000	0.000000	0.000000
1990:09	0.000000	9.000000	0.000000	0.000000	0.000000
1990:10	0.000000	10.000000	0.000000	0.000000	0.000000
1990:11	0.000000	11.000000	0.000000	0.000000	0.000000
1990:12	0.000000	12.000000	0.000000	0.000000	0.000000
1991:01	0.000000	13.000000	0.000000	0.000000	0.000000
1991:02	1.000000	14.000000	0.000000	0.000000	0.000000
1991:03	0.000000	15.000000	0.000000	0.000000	1.000000
1991:04	0.000000	16.000000	0.000000	1.000000	0.000000
1991:05	0.000000	17.000000	1.000000	0.000000	0.000000
1991:06	0.000000	18.000000	0.000000	0.000000	0.000000
1991:07					



## Các biến giả

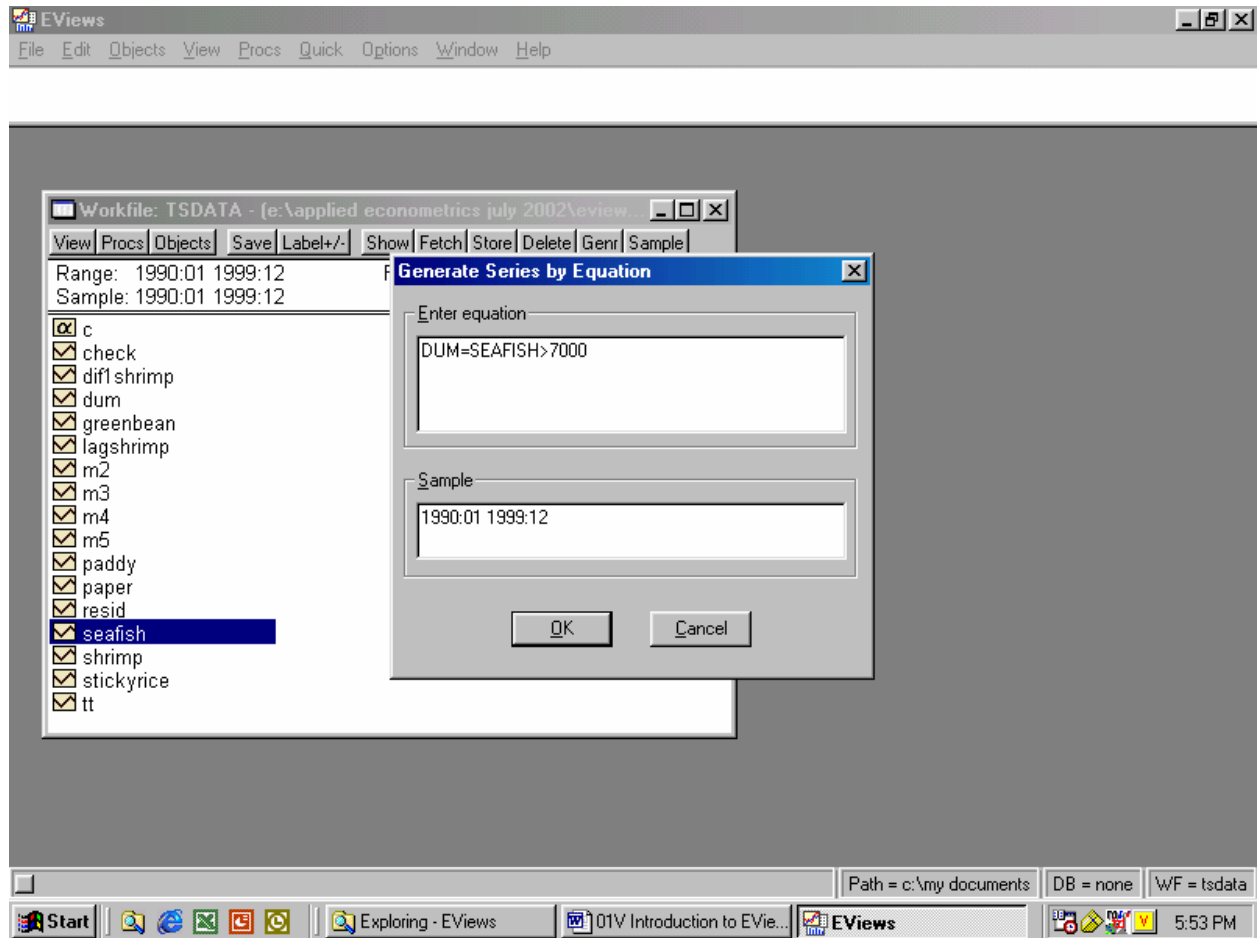
Các biến giả được sử dụng để đại diện cho sự có mặt hoặc là vắng mặt của các thuộc tính định lượng. Trong phần tiếp theo, chúng ta sử dụng một biến giả để xác định các quan sát mà đối với giá gạo nếp (**sticky rice**) giảm so với giá trị trước đó của nó.

### Mở dạng đồ thị của seafish

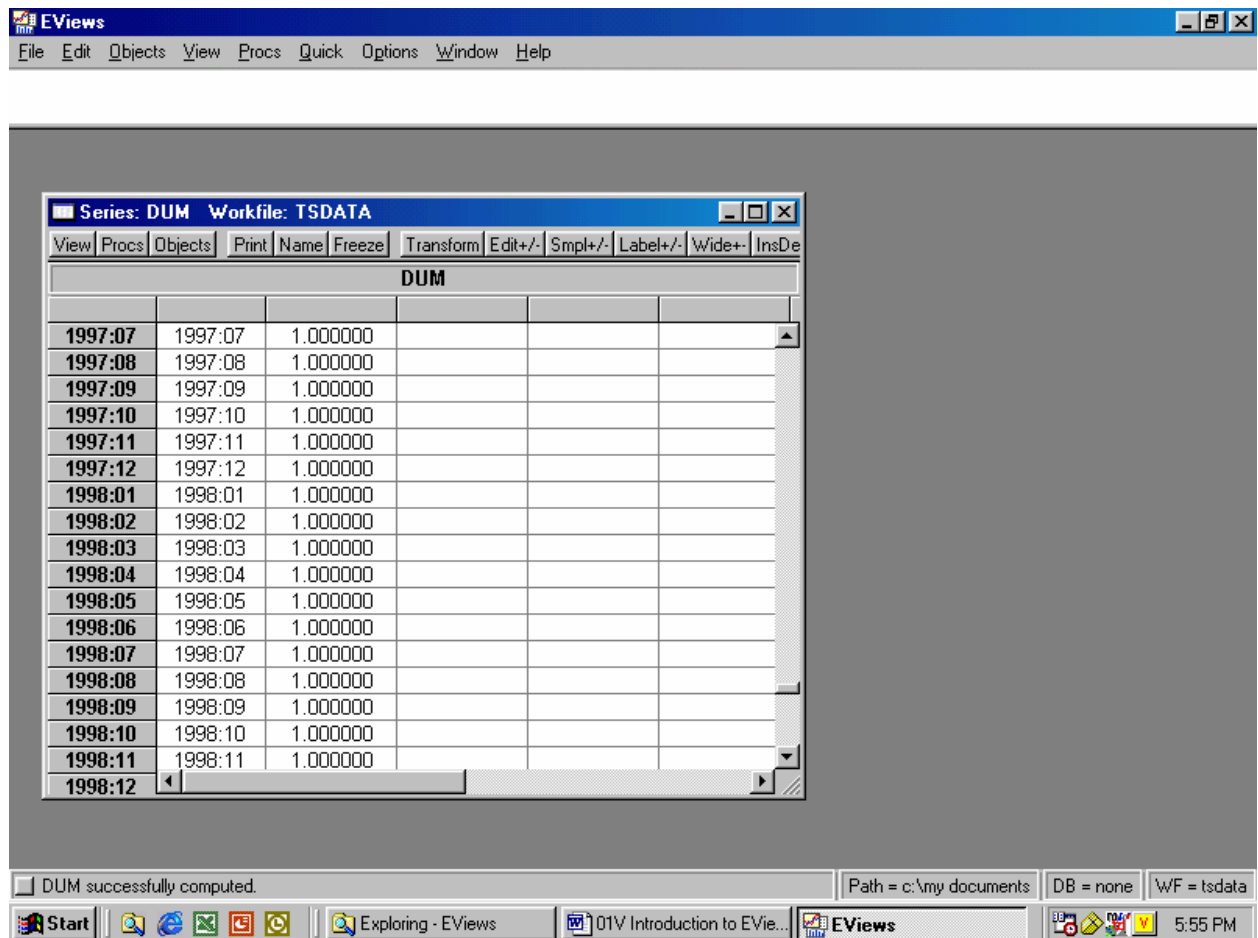


Chúng ta muốn kiểm định rằng giá có xu hướng tăng trong giai đoạn 90-95, do đó chúng ta có thể tạo ra biến giả như sau

$$\text{DUM} = \text{SEAFISH} > 7000$$



Việc này tạo ra một biến có giá trị bằng 1 khi giá SEAFISH > 7000 và BẰNG zero nếu không phải như vậy.



The screenshot shows the EViews software interface. A window titled 'Series: DUM Workfile: TSDATA' is open, displaying a table of data. The table has columns for time periods and values. The values are all 1.000000. The status bar at the bottom indicates 'DUM successfully computed.' and the path is 'c:\my documents'. The taskbar shows the Start button and several open applications, including 'Exploring - EViews' and '01V Introduction to EVie...'. The system clock shows 5:55 PM.

Time	Value
1997:07	1.000000
1997:08	1.000000
1997:09	1.000000
1997:10	1.000000
1997:11	1.000000
1997:12	1.000000
1998:01	1.000000
1998:02	1.000000
1998:03	1.000000
1998:04	1.000000
1998:05	1.000000
1998:06	1.000000
1998:07	1.000000
1998:08	1.000000
1998:09	1.000000
1998:10	1.000000
1998:11	1.000000
1998:12	1.000000

## Ước lượng phép hồi qui đơn biến trên dữ liệu chéo

### Hàm sản xuất

Trong phần này chúng ta sẽ xem xét kết quả của một hàm sản xuất đơn biến được ước lượng bằng cách sử dụng tập hợp dữ liệu các kim loại sơ cấp. Hiện thời chúng ta tránh phép hồi qui chuỗi thời gian, vì các hồi qui chuỗi thời gian đòi hỏi việc kiểm định thực tiễn đối với giả thiết về tính dừng, là điều mà chúng ta còn chưa sẵn sàng.

Chúng ta hãy viết một hàm sản xuất như sau, trong đó sản lượng là một hàm số của nhân công và vốn:  $q = f(L, K)$ .

Một dạng phổ biến của hàm  $f$  này có tên là hàm sản xuất Cobb-Douglas:

$$\ln(q) = \beta_1 + \beta_2 \ln(L) + \beta_3 \ln(K)$$

Giả thuyết về lợi thế kinh tế không đổi theo qui mô (CRS) hàm ý rằng  $\beta_2 + \beta_3 = 1$ , nó cho phép ta thay  $\beta_3 = 1 - \beta_2$  và viết lại phương trình này như sau :

$$\ln(q/K) = \beta_1 + \beta_2 \ln(L/K)$$

Nói cách khác,  $\log(\text{tỉ lệ sản lượng/vốn})$  là một hàm tuyến tính của  $\log(\text{tỉ lệ nhân công/vốn})$ .

Để hoàn tất phần xác định về mặt kinh tế lượng cho mô hình này, chúng ta đặt một chỉ số cho các quan sát, và chúng ta bổ sung thêm số hạng nhiễu ngẫu nhiên,  $\varepsilon$ . Chúng ta giả định rằng  $\varepsilon$  thoả mãn các giả định chuẩn cổ điển.

Mô hình kinh tế lượng này là :

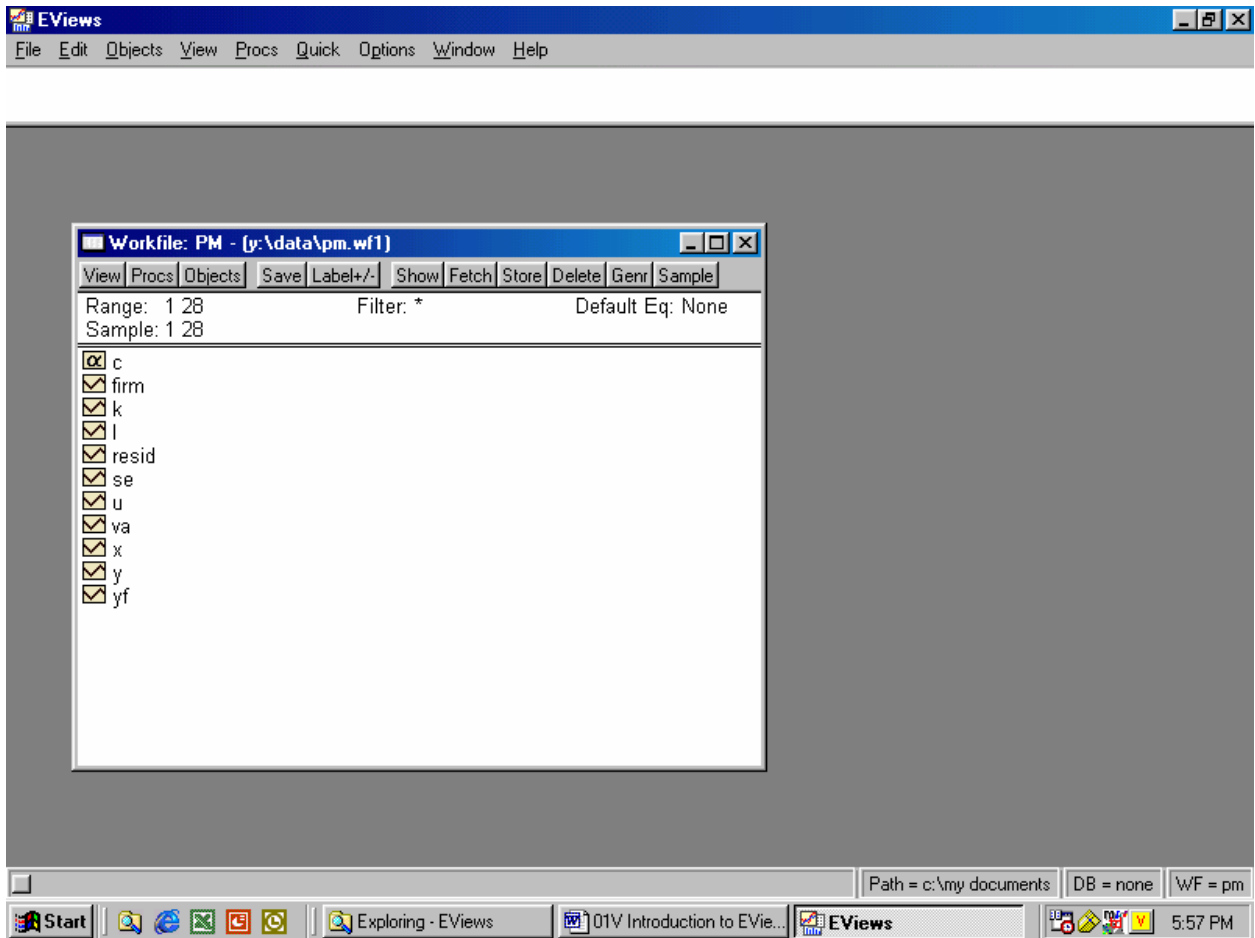
$$\ln(q_i / K_i) = \beta_1 + \beta_2 \ln(L_i / K_i) + \varepsilon_i$$

Với các định nghĩa  $Y = \ln(q / K)$  và  $X = \ln(L / K)$  chúng ta có thể viết mô hình này như trong sách giáo khoa :

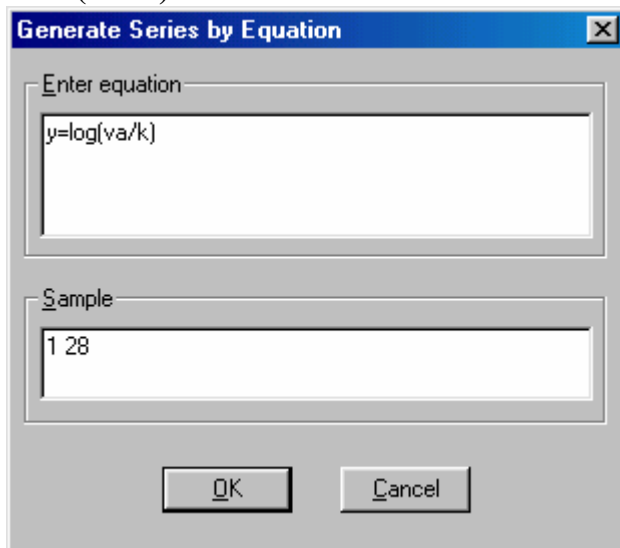
$$Y_i = \beta_1 + \beta_2 X_i + \varepsilon_i.$$

Dữ liệu của chúng ta về ngành công nghiệp kim loại sơ cấp bao gồm một số đo sản lượng (giá trị gia tăng), một số đo nhập lượng nhân công, và một số đo nhập lượng vốn.

Hãy khởi động EViews và mở workfile các kim loại sơ cấp **pm.wf1**.

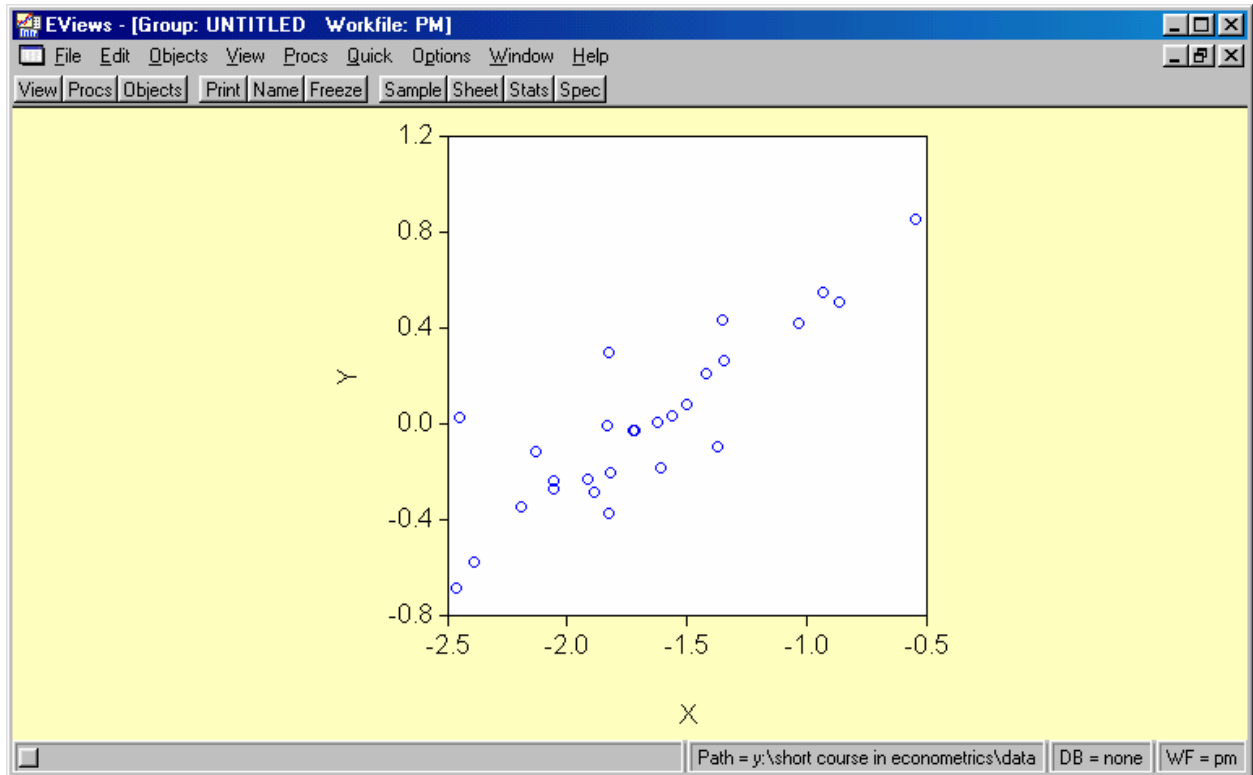


Hãy tạo ra các biến mới bằng cách sử dụng hàm **GENR** :  $Y = \text{LOG}(VA / K)$  và  $X = \text{LOG}(L / K)$ .





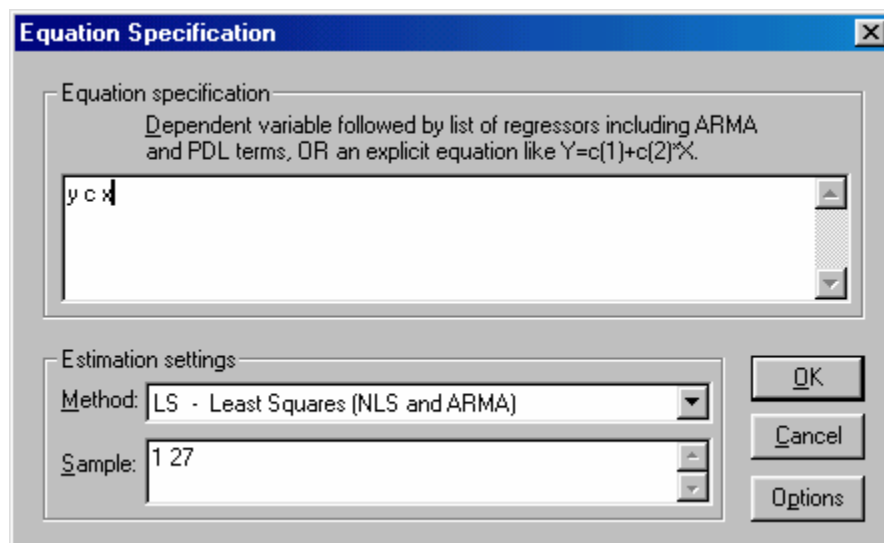
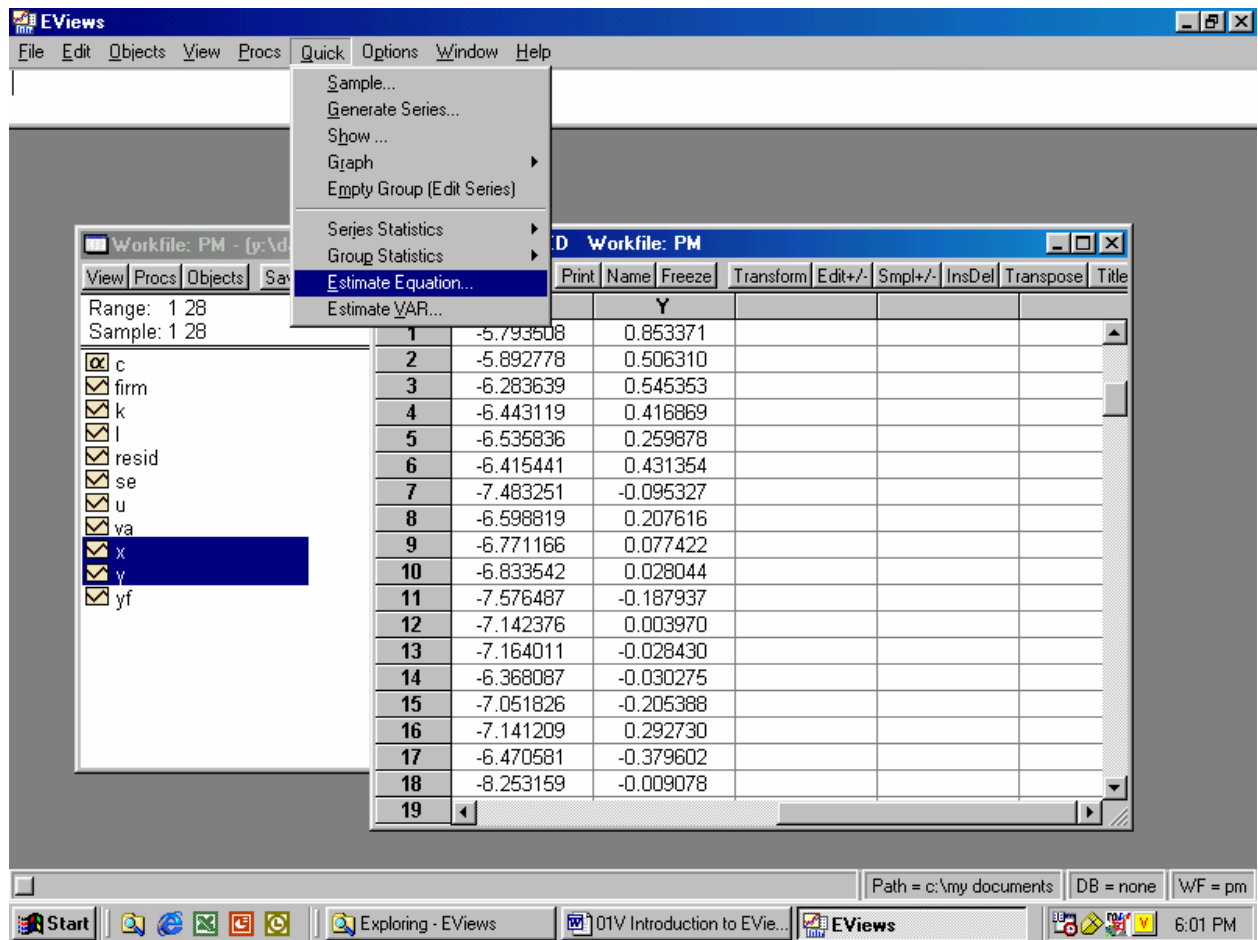
Kết quả trông như thế này:



Sau khi chiêm ngưỡng đồ thị của mình, Anh/Chị đã sẵn sàng chạy một phép hồi qui đơn biến.

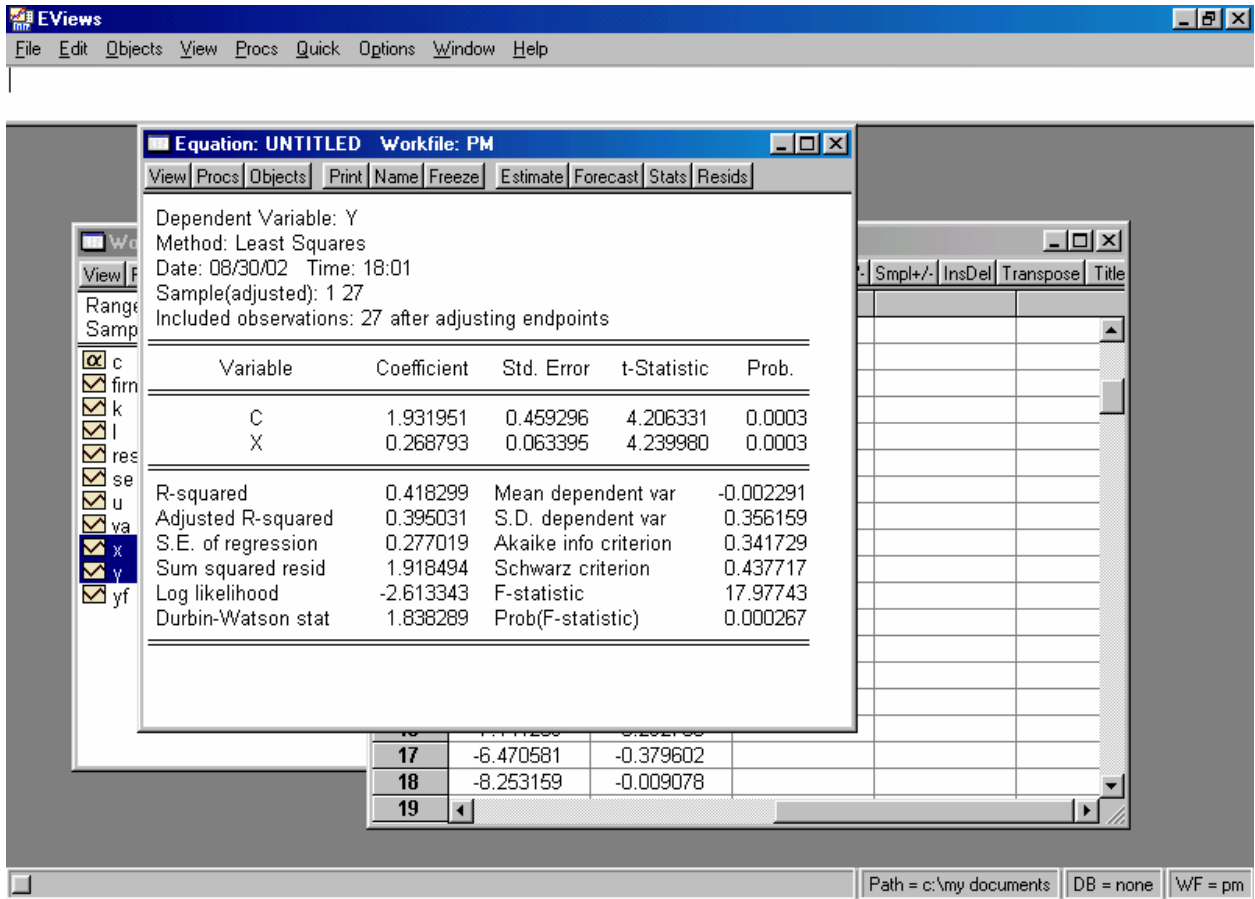
### Ước lượng phép hồi qui đơn biến

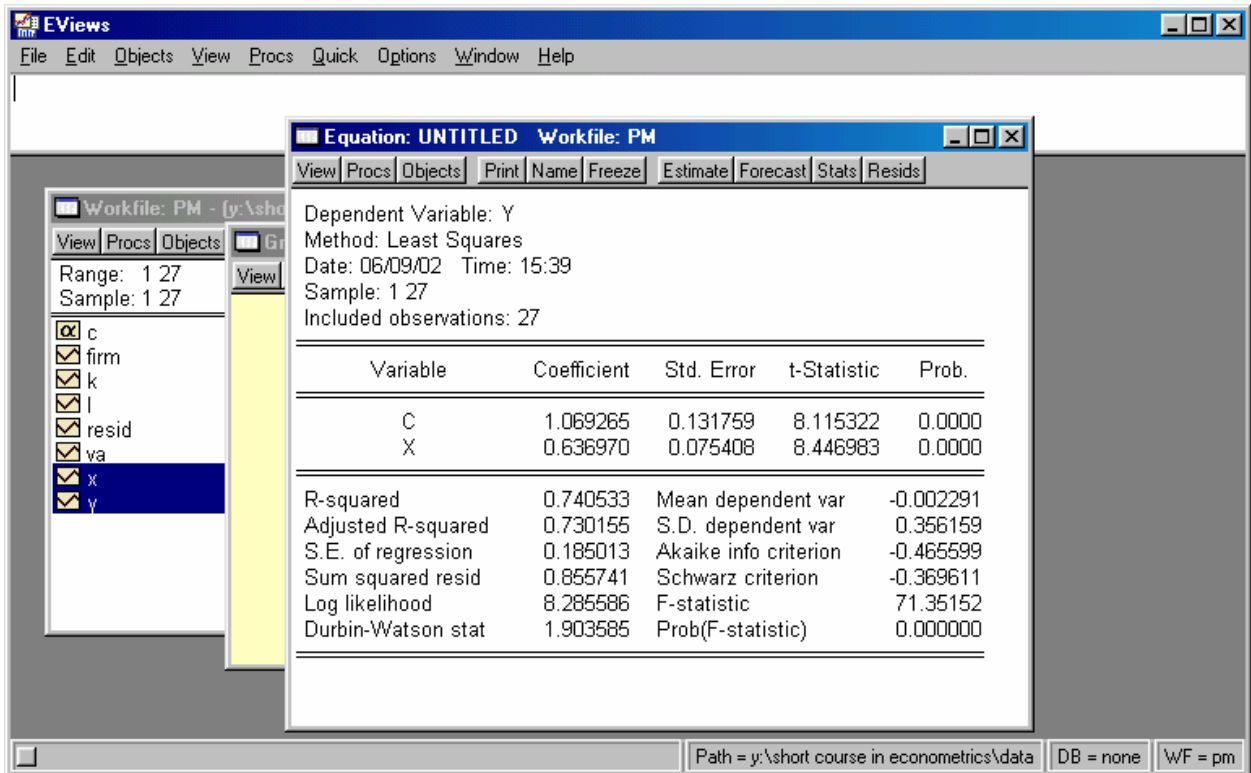
Nhấp **Quick / Estimate Equation**. Khi ô xác định phương trình (Equation Specification) mở ra, hãy đánh máy **Y C X** và nhấp **OK**.



Kết quả hồi qui xuất hiện!







Hãy giải thích từng mục trong bảng trên và liên hệ nó với phần ghi bài giảng của Anh/Chị.

### Ước lượng phép hồi qui bội

Nếu hàm sản xuất của chúng ta không có lợi thế kinh tế không đổi theo qui mô, thì ràng buộc  $\beta_2 + \beta_3 = 1$  không hiệu lực, và chúng ta không thể đơn giản hoá hàm sản xuất Cobb-Douglas như chúng ta đã làm ở trên. Thay vào đó, chúng ta có thể viết phần xác định kinh tế lượng như sau:

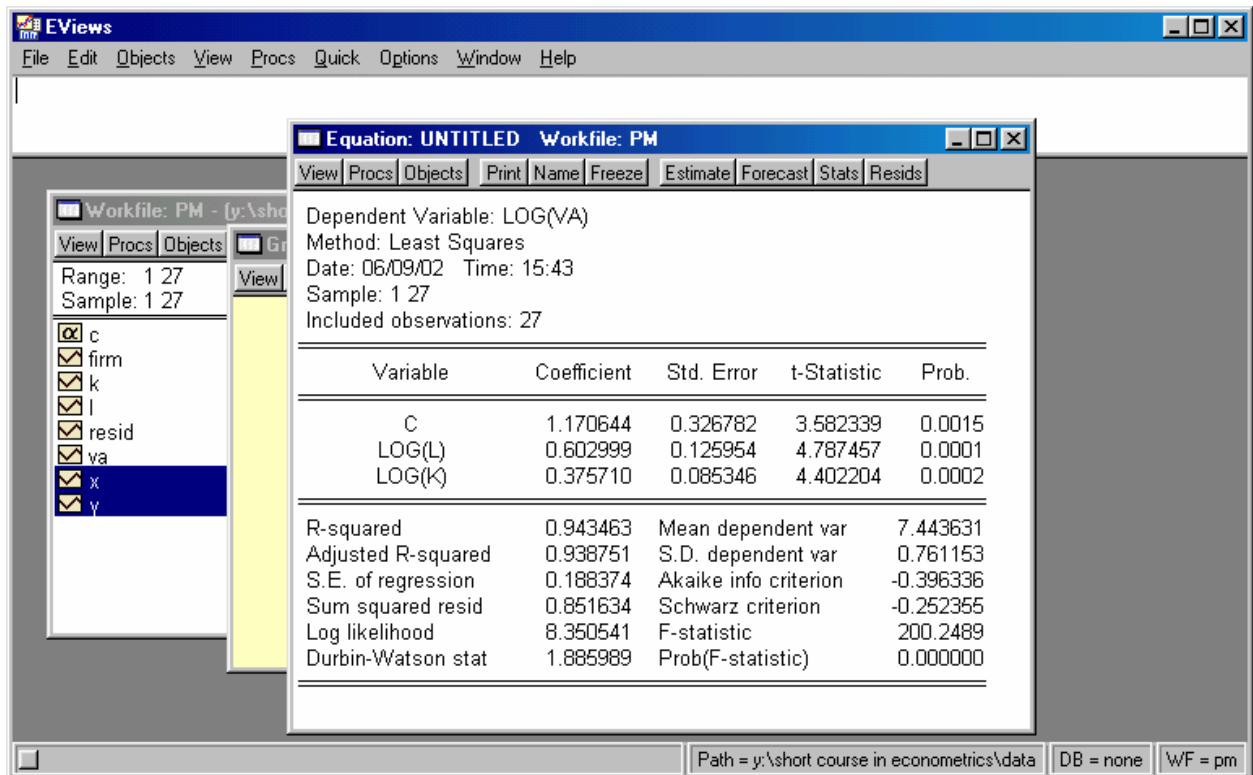
$$\ln(q_i) = \beta_1 + \beta_2 \ln(L_i) + \beta_3 \ln(K_i) + \varepsilon_i$$

trong đó  $q = VA$ .

Với mục đích ước tính phương trình này, hãy nhấp **Quick / Estimate Equation** và sau đó đánh máy vào hộp Xác định phương trình (Equation Specification) như sau :

$$\log(VA) \ C \ \log(L) \ \log(K)$$

Hãy lưu ý tới sự khác biệt giữa điều chúng ta đã làm ở đây với điều chúng ta đã làm trong ví dụ trước. Trong ví dụ trước, tôi muốn sử dụng ký hiệu đúng như các ký hiệu trong bài giảng, nên tôi đã tạo ra các biến mới có tên là X và Y. Tuy nhiên, chúng ta không cần phải làm điều đó trên EViews. Việc này có một lợi điểm lớn. Trong kết quả nêu ra sau đây, những biến đổi này của các biến được chỉ rõ !



Giải thích kết quả .

Cuối cùng, xét giả thiết về lợi thế kinh tế không đổi theo qui mô. Chúng ta có thể phát biểu điều này như là :

$$H_0 : \beta_2 + \beta_3 = 1$$

Cách tốt nhất để kiểm định giả thiết này trên EViews là thực hiện một kiểm định Wald. Sự trình bày đơn giản nhất của kiểm định Wald giả định rằng các giả định chuẩn cổ điển là đúng.

Sự áp đặt một ràng buộc bắt buộc (giả thiết ) lên một phép hồi qui làm tăng tổng các bình

phương phần dư (ESR). Nếu gia tăng này lớn, thì chúng ta có thể phán quyết rằng những ràng buộc này là không tương hợp với dữ liệu, vì thế chúng ta cần phải bác bỏ giả thiết này. Câu hỏi đặt ra là một gia tăng “lớn” là “lớn” tới mức nào?

Để xây dựng kiểm định Wald, hãy chạy hai phép hồi qui : một phép không có giới hạn và một phép có giới hạn. Những phép này thu được tổng các bình phương phần dư không có giới hạn ( $ESSR_U$ ) và tổng các bình phương phần dư có giới hạn ( $ESSR_R$ ). Trị thống kê kiểm định đối với kiểm định Wald được xác định như sau :

$$\hat{F} = \frac{(SSR_R - SSR_U)/q}{SSR_U/(n - k_u)}$$

Ở đây , q, bậc tự do của tử số , bằng số các ràng buộc . Nó có thể được tính như là chênh lệch giữa các bậc tự do của các phép hồi qui có giới hạn và không có giới hạn. Bậc tự do của mẫu số bằng số bậc tự do của phép hồi qui không có giới hạn.

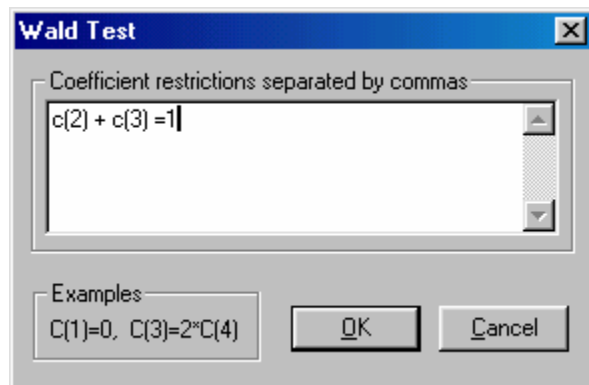
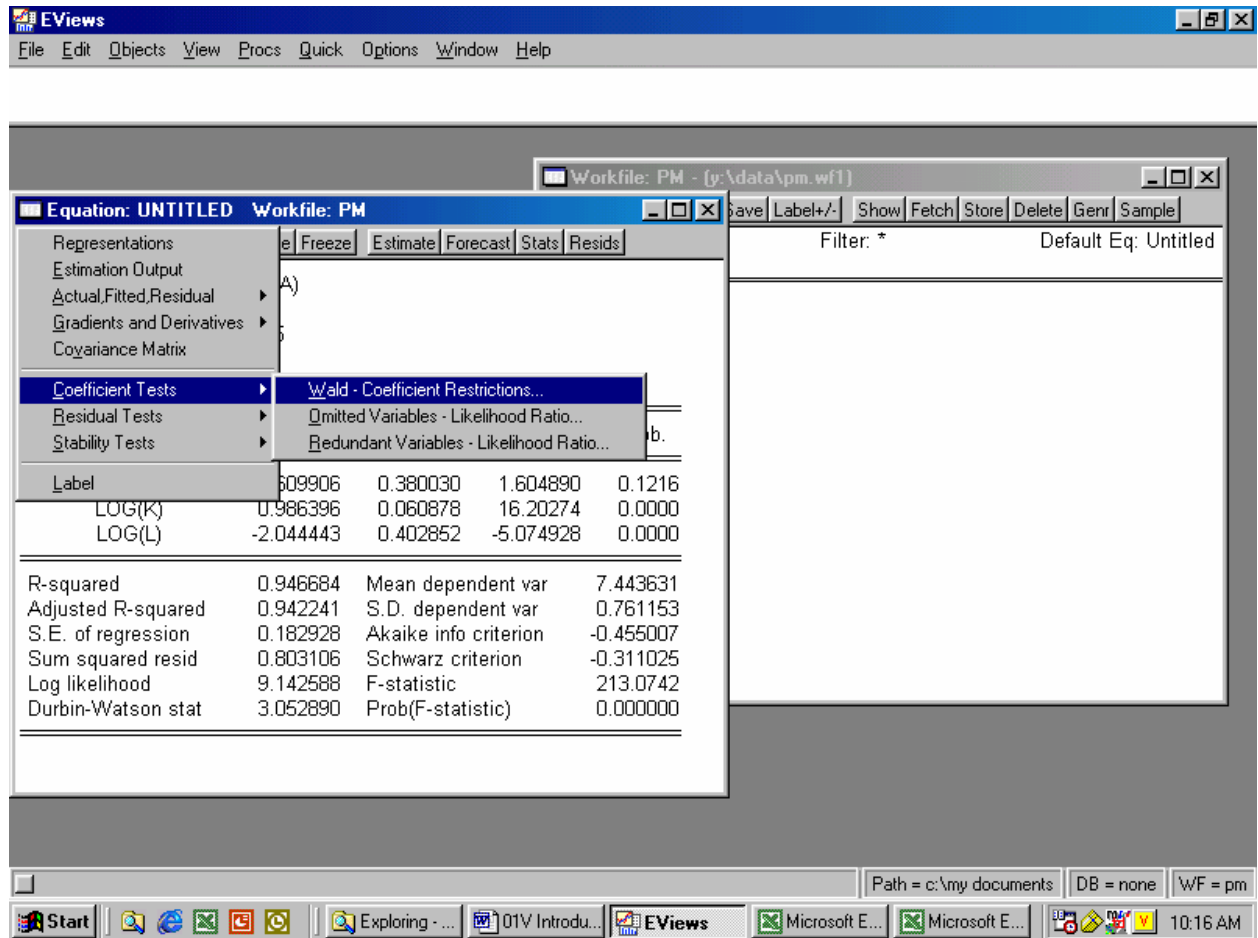
Các giá trị lớn của  $\hat{F}$  xảy ra khi những ràng buộc này dẫn tới một gia tăng lớn trong ESR. Quy tắc quyết định là bác bỏ giả thuyết không nếu F vượt quá giá trị tới hạn phù hợp hoặc nếu giá trị **Prob** đối với kiểm định này nhỏ hơn mức độ ý nghĩa đã lựa chọn trước đây.

Sự thể hiện này cho ta một cảm nhận trực giác nào đó, nhưng trên **EViews** không cần thiết phải chạy hai phép hồi qui với mục đích thực hiện kiểm định Wald này.

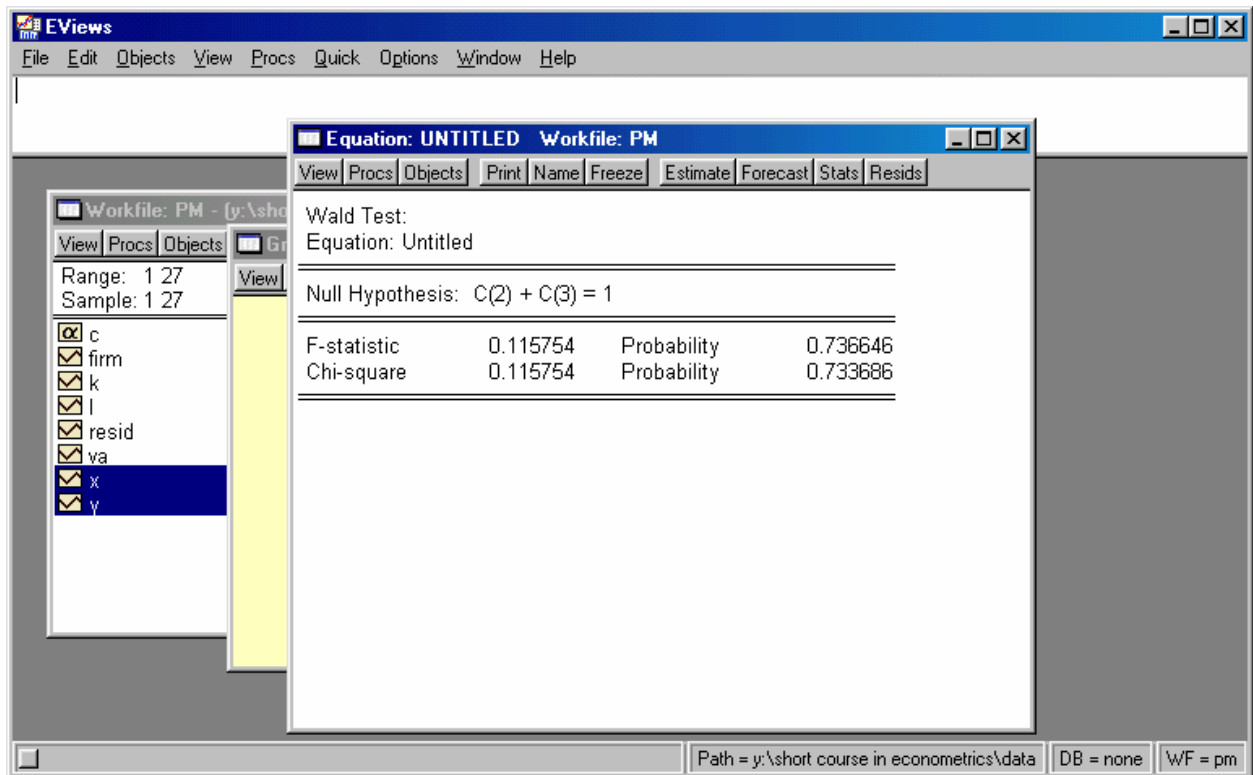
Trở lại với hàm hồi qui của Anh/Chị  $\log(Y) = C + \log(L) + \log(K)$

Bây giờ hãy nhấp **View / Coefficient Tests / Wald - Coefficient Restrictions**. Trong hộp thoại mở ra , đánh máy những ràng buộc bằng cách sử dụng ký hiệu liên quan tới việc đặt tên nội bộ cho các hệ số trong EViews:

$$c(2) + c(3) = 1$$



Kết quả như sau xuất hiện :



Liên kết giữa ký hiệu của EViews và ký hiệu trên lớp là :

$$\begin{aligned}c(1) &= \hat{\beta}_1 \\c(2) &= \hat{\beta}_2 \\&\vdots \\c(k) &= \hat{\beta}_k\end{aligned}$$

Trong trường hợp này trị thống kê kiểm định là  $\hat{F} = 0.115754$  và giá trị **prob** là 0.7366. **Quyết định của chúng ta là Không Bác Bỏ Được Giả Thiết Này.**

### **EViews còn ứng dụng nhiều hơn nữa**

Như vậy chúng ta thấy rằng Eviews rất hiệu quả trong việc sử dụng phân tích dữ liệu. Chúng ta sẽ sử dụng EViews cho nhiều tình huống khác nữa trong các bài giảng của môn học này và trong thực tế.

**Chúc các anh chị vui vẻ và thích thú khi tìm tòi nhiều điều mới lạ và thú vị trong EViews**