



**ĐẠI HỌC KINH TẾ QUỐC DÂN**  
**KHOA TOÁN KINH TẾ**  
**BỘ MÔN TOÁN KINH TẾ**



**TÀI LIỆU HƯỚNG DẪN**  
**THỰC HÀNH KINH TẾ LƯỢNG**  
**BẰNG PHẦN MỀM EIEWS 4.0**

**Tác giả: Bùi Dương Hải**

**[www.mfe.edu.vn/buiduonghai](http://www.mfe.edu.vn/buiduonghai)**

**Lưu hành nội bộ**

*Có bổ sung so với bản năm 2013*

**HÀ NỘI, 1 / 2014**

# MỤC LỤC

<b>MỞ ĐẦU .....</b>	<b>3</b>
<b>§ 1 SỐ LIỆU DÙNG TRONG EIEWS .....</b>	<b>6</b>
<b>§ 2 NHẬP SỐ LIỆU TỪ BÀN PHÍM – XỬ LÝ SỐ LIỆU .....</b>	<b>9</b>
<b>§ 3 ƯỚC LƯỢNG MÔ HÌNH HỒI QUY ĐƠN .....</b>	<b>18</b>
<b>§ 4 ƯỚC LƯỢNG MÔ HÌNH HỒI QUY BỘI .....</b>	<b>23</b>
<b>§ 5 MÔ HÌNH VỚI BIẾN GIÁ.....</b>	<b>31</b>
<b>§ 6 ĐỊNH DẠNG PHƯƠNG TRÌNH HỒI QUY .....</b>	<b>37</b>
<b>§ 7 HIỆN TƯỢNG ĐA CỘNG TUYẾN.....</b>	<b>41</b>
<b>§ 8 HIỆN TƯỢNG PHƯƠNG SAI SAI SỐ THAY ĐỔI.....</b>	<b>44</b>
<b>§ 9 MÔ HÌNH VỚI SỐ LIỆU CHUỖI THỜI GIAN.....</b>	<b>53</b>
<b>§ 10 TỰ TƯƠNG QUAN VÀ MÔ HÌNH CÓ BIẾN TRỄ.....</b>	<b>61</b>

**Chương trình Eviews4 có thể tải về tại trang mạng Khoa Toán kinh tế - Đại học Kinh tế quốc dân:**

**[www.mfe.edu.vn](http://www.mfe.edu.vn) → Thư viện → Dữ liệu - phần mềm**


Địa chỉ trên cũng là nơi có thể tải về tài liệu hướng dẫn này, và các thư mục số liệu.

**Các thắc mắc, trao đổi xin vui lòng gửi thư cho tác giả theo địa chỉ thư điện tử: [haitkt@gmail.com](mailto:haitkt@gmail.com)**

## MỞ ĐẦU

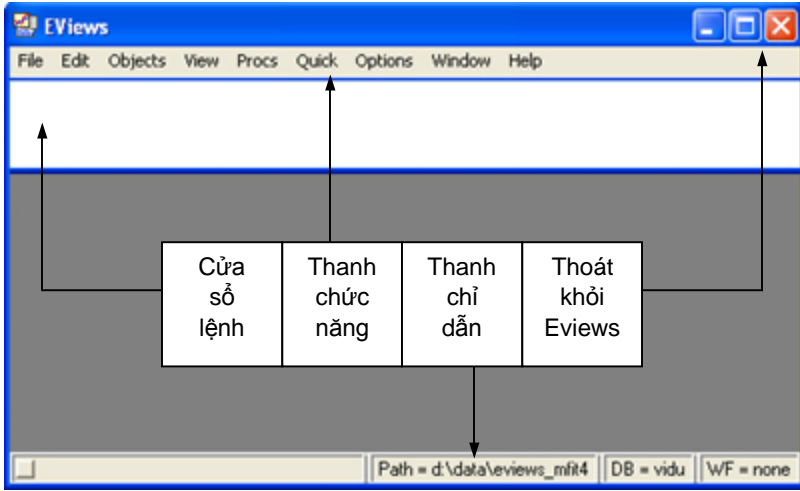
Eviews là phần mềm được thiết kế riêng cho các mô hình kinh tế lượng và chuỗi thời gian. Phần mềm này phù hợp cho giảng dạy và học tập kinh tế lượng cho đối tượng sinh viên đại học và sau đại học. Hiện nay (1/2013) đã có phiên bản thương mại Eviews7, từ phiên bản Eviews5 có yêu cầu cài đặt cũng như bản quyền. Eviews4 không yêu cầu cài đặt và bản quyền, có thể tải về từ trang mạng khoa Toán kinh tế. Khi đã nắm được các kỹ năng với Eviews4, việc chuyển sang thực hành với các phiên bản cao hơn là hoàn toàn tương tự. Do đó để phù hợp với thực trạng tại Việt Nam, với mục tiêu tạo điều kiện thuận lợi nhất cho sinh viên, tài liệu này được viết cho thực hành phiên bản Eviews4.

Với sự đổi mới của giáo trình và chương trình giảng dạy Kinh tế lượng tại Đại học Kinh tế quốc dân từ năm 2013, nội dung hướng dẫn thực hành cũng có sự thay đổi phù hợp. Nội dung trước hết sẽ sử dụng các số liệu chéo, sau đó thực hiện với số liệu chuỗi thời gian.

Tệp chạy chương trình Eviews có biểu tượng là . Nhấn vào biểu tượng của Eviews, cửa sổ chính của chương trình xuất hiện.

Cửa sổ chính của chương trình Eviews gồm các phần:

- Thanh chức năng: thực hiện các thao tác đã được chương trình định sẵn, tương tự như các chương trình chạy trong Windows.
- Cửa sổ lệnh: bên dưới thanh chức năng là nơi để viết lệnh trực tiếp. Có thể dùng chuột để kéo rộng cửa sổ lệnh tùy ý.
- Thanh chỉ dẫn: xác định đường dẫn đến tệp đang sử dụng.
- Các nút thu nhỏ, mở rộng cửa sổ, và thoát khỏi chương trình.



Có thể không cần sử dụng chuột mà dùng bàn phím để chọn lựa các nút. Ấn và giữ phím Alt trên bàn phím, trên dòng task bar các lựa chọn sẽ tự động gạch chân các chữ cái. Khi đó phím Alt và nhấn phím tương ứng với chữ cái tương ứng sẽ cho kết quả giống như khi dùng chuột chọn nút đó.






Ví dụ: Khi giữ phím Alt, gõ phím F tương đương với nhấn chuột vào nút File; chữ E tương đương với nút Edit.



- Ngăn cách phần nguyên và phần thập phân của một số, Eviews dùng dấu chấm “.”

## Một số chủ thể cơ bản của Eviews

Eviews làm việc với một số dạng chủ thể cơ bản, mỗi chủ thể có thể lưu lại và đặt tên để có thể dễ dàng sử dụng khi cần thiết. Một số chủ thể thông dụng gồm:

-  **Series:** Thông tin về từng biến số.
-  **Group:** Thông tin về một số biến số xét cùng lúc.
-  **Graph:** Thông tin đồ thị. Đồ thị có thể gồm đồ thị của một biến, đồ thị của nhiều biến, đồ thị của các biến theo nhau.
-  **Equation:** Thông tin về một phương trình hồi quy
-  **System:** Thông tin về một hệ nhiều phương trình hồi quy.

## Một số kí hiệu dùng cho thực hành

- : Thao tác, thực hiện thao tác yêu cầu.  
Ví dụ  Chọn View : dùng chuột nhấn vào nút View
- $\Rightarrow$  : Kết quả của thao tác.
- $\rightarrow$  : Các thao tác, chọn lựa kế tiếp nhau.  
Ví dụ: File  $\rightarrow$  Open: Chọn nút File rồi nút Open.
- [?] : Câu hỏi, cần nắm được lý thuyết để trả lời.  
Tất cả các câu hỏi kiểm định trong sách đều được thực hiện với mức ý nghĩa  $\alpha$  là 5%.

Cặp ngoặc vuông [...]: do trong Eviews có thể mở nhiều cửa sổ cùng lúc, do đó cặp ngoặc vuông để xác định cửa sổ cần thực hiện thao tác. Ví dụ [Eviews] là cửa sổ chính.

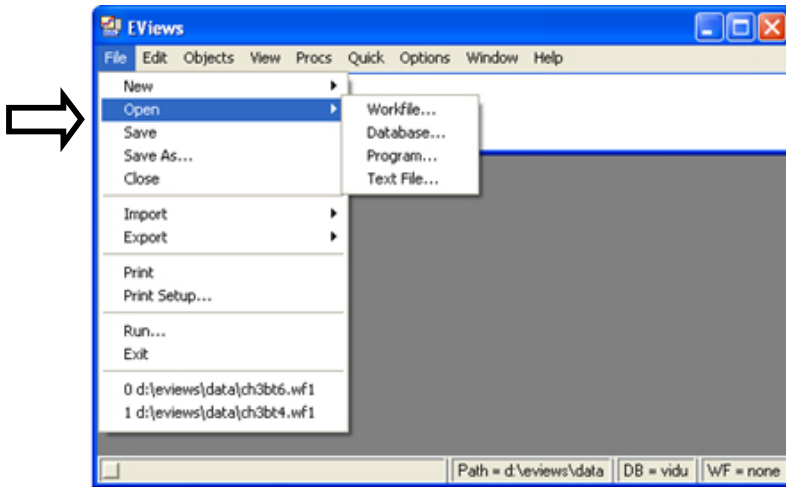
- Để ngắn gọn và thuận tiện về sau, các nút trên các cửa sổ được viết trong cặp dấu <>, ví dụ: <Ok>, <Cancel>

## § 1 SỐ LIỆU DÙNG TRONG EIEWS

Eviews là chương trình xử lý số liệu, ước lượng phương trình hồi quy, phân tích chuỗi thời gian, do đó việc hiểu rõ về số liệu là điều cần thiết. Để hiểu rõ cấu trúc số liệu được quản lý và xử lý bởi Eviews, mở một số bộ số liệu và quan sát các số liệu sau.

### Mở bộ số liệu chéo (cross-section)

□ Tại cửa sổ chính, chọn **File** → **Open**



Trong lựa chọn Open, có bốn dạng định dạng tệp có thể mở:

- Dạng Workfile: là tệp dữ liệu và thực hiện các phân tích thông thường. Đây là dạng cơ bản, trong tệp có thể lưu số liệu, các đồ thị, các phương trình hồi quy, kết quả ước lượng.
- Dạng Database: cơ sở dữ liệu, bao gồm nhiều định dạng.
- Dạng Program: các chương trình được lập trình trước đó
- Dạng Text File: các tệp lưu trữ dạng văn bản

- Chọn dạng **Workfile**, là dạng thông thường nhất để tính toán xử lý số liệu với Eviews.
- Chọn thư mục **DATA2012**, tệp số liệu **YWKM**, cửa sổ Workfile mở ra. Trên cửa sổ này có một số thông tin:
  - Dòng trên cùng: Tên của Workfile và đường dẫn
  - Các nút với các chức năng khác nhau, sẽ được đề cập sau
  - Khoảng số liệu và Mẫu từ 1 đến 100

Bên dưới của cửa sổ, liệt kê các chủ thể mà Workfile đang quản lý, gồm:  c  k  m  resid  tc  w  y.

Nhấn vào nút **Label+/-**, xuất hiện các thông tin về thời gian khởi tạo các số liệu này, và chú thích về ba biến. Hai chủ thể  c  resid không có chú thích, vì đây là hai chủ thể đặc biệt dùng để lưu các thông tin riêng.

Thông tin của một biến số thông thường bao gồm:

- Tên biến: Tối đa 24 ký tự chỉ gồm chữ và số, không có dấu cách, bắt đầu bởi chữ cái.
- Nhãn biến: chú thích về ý nghĩa của biến
- Tần số: nếu số liệu chéo thì đánh số thứ tự, nếu chuỗi thời gian thì theo thứ tự thời gian
- Giá trị của biến: đo lường bằng số, dấu ngăn cách với phần thập phân là dấu chấm. Khi chưa có giá trị thì kí hiệu là NA (not available)

c : chủ thể chứa các hệ số tính được từ các phương trình hồi quy, các mô hình. Khi chưa có kết quả hồi quy từ phương trình nào, các giá trị của C được gán bằng 0

resid: là chuỗi nhận sẽ nhận giá trị là phần dư từ có được từ việc ước lượng các phương trình hồi quy. Khi chưa có phương trình hồi quy, các giá trị Resid đều chưa có.

- Nhấn đúp chuột trái vào  **k**, cửa sổ [Series: K] mở ra. Cột ngoài cùng bên trái **obs** thứ tự của quan sát từ 1 đến 100, các giá trị của biến K được liệt kê theo các quan sát từ cột tiếp theo trong bảng.
- Chọn  **resid**, mở cửa sổ [Series: RESID] với các giá trị đều là NA vì chưa có kết quả tính toán nào được thực hiện.
- Tại cửa sổ [Workfile], sử dụng chuột đánh dấu (bôi đen) các biến từ K, M, TC, Y, W nháy chuột phải, chọn Open ▶ as Group, tất cả các biến đều được liệt kê trong cùng một cửa sổ [Group].

### **Mở bộ số liệu có tần số theo Quý**

Mở bộ số liệu mới, tại cửa sổ [Eviews] File → New, chọn tệp **VNQ\_GDP**.

Bộ số liệu được thể hiện từ 2004:1 đến 2012:3, với một chữ số sau dấu “:”. Với cách thể hiện này, số liệu là từ Quý 1 năm 2004 đến Quý 3 năm 2012.

- Chọn biến bất kỳ, chẳng hạn GDP, mở ra dưới dạng cửa sổ [Series]. Với cửa sổ này, tần suất biến có chu kỳ 1, 2, 3, 4, 1,... thể hiện đây là số liệu Quý.

### **Mở bộ số liệu có tần số theo Tháng**

Mở bộ số liệu mới, tại cửa sổ [Eviews] File → New, chọn tệp **VNM\_EXIM**. Bộ số liệu được thể hiện từ 2004:01 đến 2008:12, với hai chữ số sau dấu “:”, thể hiện là số liệu từ Tháng 1 năm 2004 đến Tháng 12 năm 2008.

Để biết ý nghĩa các biến, chọn nút Label trên thanh chức năng của cửa sổ [Workfile].



## § 2 NHẬP SỐ LIỆU TỪ BÀN PHÍM – XỬ LÝ SỐ LIỆU

Xét bộ số liệu ví dụ sau được mô phỏng về 12 hộ gia đình, với X2 là tổng tiền lương, X3 là tổng thu nhập khác, Y là tổng chi tiêu.

<b>STT</b>	<b>X<sub>2</sub></b> <b>(Lương)</b>	<b>X<sub>3</sub></b> <b>(TN khác)</b>	<b>Y</b> <b>(Chi tiêu)</b>
1	20	16	24.4
2	30	10	31.2
3	28	2	29.2
4	24	0	23.6
5	32	18	36.0
6	36	10	31.4
7	32	16	32.6
8	34	24	36.8
9	24	28	32.8
10	22	20	29.8
11	28	8	30.2
12	30	4	26.8

Thực hiện nhập bộ số liệu trên và lưu lại dưới dạng một tệp chuyên dụng của Eviews.

### 2.1 Định dạng tần số và nhập số liệu

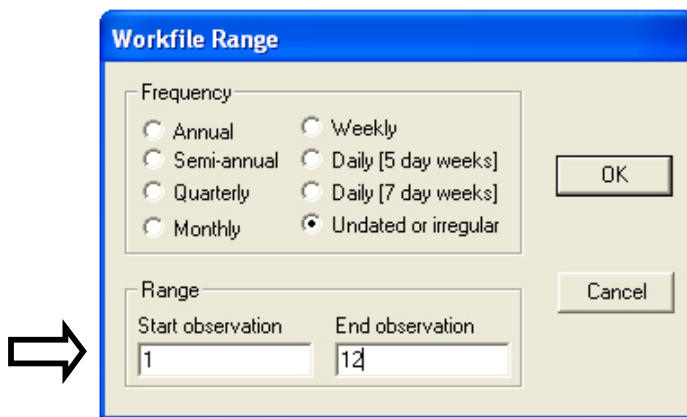
Tại cửa sổ chính của Eviews, để thuận tiện, đóng cửa sổ nhỏ đang mở (nếu có)

Chọn **File** → **New** : Cửa sổ [Workfile Range]: tần số của số liệu.  
Dấu lựa chọn được ngầm định đặt tại lựa chọn “Annual”

**Frequency - tần số của số liệu**

<b>Tần số và định dạng</b>	<b>Ví dụ</b>	
	<b>Đầu – cuối</b>	<b>Ý nghĩa</b>
Annual (Năm) <i>yyyy</i>	Start: 1991 End: 2005	15 quan sát theo năm, từ năm 1991 đến năm 2005
Semi-annual (Nửa năm) <i>yyyy:h</i>	Start: 1991:1 End: 2005:2	30 quan sát theo nửa năm, từ nửa đầu năm 1991 đến nửa sau năm 2005
Quarterly (Quý) <i>yyyy:q</i>	Start: 1991:1 End: 2005:4	60 quan sát theo quý, từ quý 1 năm 1991 đến quý 4 năm 2005
Monthly (Tháng) <i>yyyy:mm</i>	Start: 1991:01 End: 2005:12	180 quan sát theo tháng, từ tháng 1 năm 1991 đến tháng 12 năm 2005
Weekly (Tuần) <i>mm/dd/yyyy</i>	Start: 01/01/2008 End: 12/01/2008	49 quan sát theo tuần, từ tuần có ngày 1 tháng 1 năm 2008 đến tuần có ngày 1 tháng 12 năm 2008
Daily [5day] (Ngày: tuần 5 ngày) <i>mm/dd/yyyy</i>	Start: 11/12/2008 End: 12/11/2008	22 quan sát theo ngày, từ ngày 12 tháng 11 năm 08 đến ngày 11 tháng 12 năm 08, không có ngày cuối tuần
Daily [7day] (Ngày: tuần 7 ngày) <i>mm/dd/yyyy</i>	Start: 11/12/2008 End: 12/11/2008	29 quan sát theo ngày, từ ngày 12 tháng 11 năm 08 đến ngày 11 tháng 12 năm 08, có ngày cuối tuần
Undated or Irregular (Quy tắc khác)	Start: 1 Start: 30	30 quan sát không theo thời gian, hoặc theo thời gian nhưng quy tắc khác

Với bộ số liệu ví dụ trên là số liệu chéo, số quan sát từ 1 đến 12.



□ Frequency:  **Undated or irregular** → Start date: **1** → End date: **12** : mở cửa sổ **Workfile**

Cửa sổ [Workfile] là cửa sổ quản lý việc nhập, lưu, xử lý số liệu. Trong cửa sổ Workfile chỉ có hai biến là các hệ số **C** và phần dư **Resid**. Cần tạo ba biến số mới và nhập số liệu.

□ Tại cửa sổ [Eviews], chọn **Quick** → **Empty Group (Edit Series)**

Mở cửa sổ [Group]

□ Chọn ô đầu tiên bên phải ô **obs**, nhập tên biến là **X2**, các ô bên dưới tự động chuyển thành **NA**, nhập các giá trị của biến X2 ứng với các số liệu đã có. Tiếp tục với cột biến **X3** và biến **Y**.

Tại cửa sổ [Workfile], biểu tượng của X2, X3, và Y xuất hiện.

### Nhập nhãn biến

Các biến X2, X3 và Y đã nhập chưa có nhãn, cần nhập nhãn để chú thích ý nghĩa của các biến.

□ Tại cửa sổ [Workfile], nhấn chuột đúp vào biến **X2**, mở cửa sổ [Series: X2] chọn **Name**, mở cửa sổ [Object Name]

Trong cửa sổ này, ô trống ở trên là tên biến, ô trống bên dưới để nhập nhãn biến. Với ví dụ đang xét, nhãn cho biến X2 là: *Luong*.

Tương tự, có thể nhập nhãn cho biến X3 là *Thu nhap khac* và biến Y là *Chi tieu*.

## 2.2 Sửa đổi số liệu và lưu số liệu

Trường hợp cần sửa đổi số liệu, ví dụ biến X2 có thể thực hiện theo trình tự:

Chọn biến **X2**, mở cửa sổ [Series: X2], chọn nút **Edit+/-** và thay đổi các giá trị cần thiết.

Số liệu đã nhập có thể lưu lại dưới dạng tệp chuyên dụng của Eviews, để có thể mở và sử dụng khi cần thiết.

Tại cửa sổ [Eviews] Chọn **File** → **Save** (hoặc Save as)

⇒ Cửa sổ [SaveAs] : Chọn vị trí và tên tệp.

Các tệp có đuôi ngầm định là .wfl.

Số liệu đã nhập ở phần 2.1 là số liệu thô. Để thấy được rõ hơn các thông tin chứa đựng trong các biến đó, cần thực hiện tính các thống kê với từng biến, xem xét tương quan, đồ thị mô tả về các biến và mối liên hệ giữa chúng.

## 2.3 Vẽ đồ thị

Mô tả số liệu qua đồ thị và các thống kê đặc trưng cơ bản là xử lý ban đầu cần thiết đối với các biến số.

Chọn **X2** và **Y**, mở cửa sổ [Group]

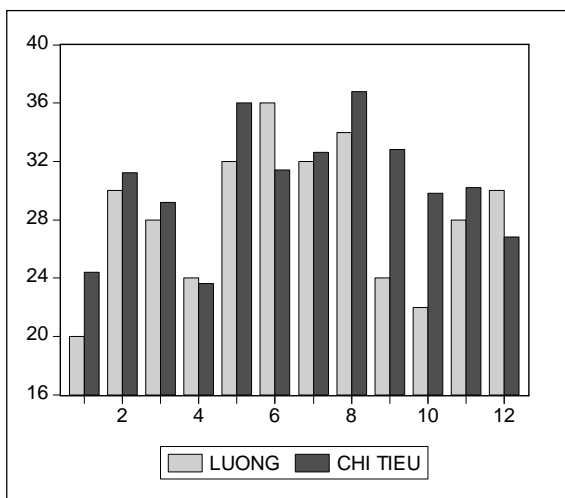
(Để chọn riêng các biến cách nhau, giữ phím Ctrl)

## Vẽ đồ thị các biến theo quan sát

Lưu ý: dạng đồ thị đường (line) thường chỉ dùng cho số liệu chuỗi thời gian, số liệu chéo nên chọn đồ thị cột)

[Group] **View** → **Graph** → **Bar**

Kết quả cho đồ thị cột của X2 và Y theo quan sát trên cùng hệ tọa độ.



Để vẽ đồ thị của biến này theo biến kia, mỗi biến trên một trục tọa độ, lựa chọn:

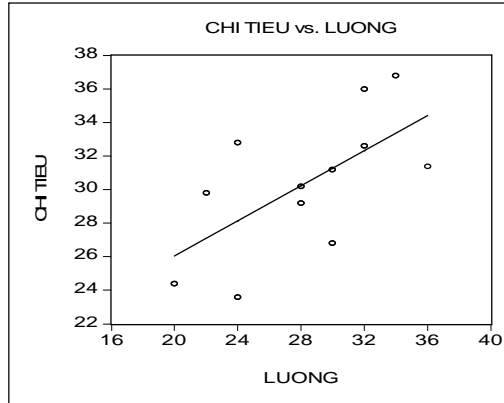
[Group] **View** → **Graph** → **Scatter** → **Simple Scatter**

Kết quả cho đồ thị điểm của biến Y trên trục tung và X2 trên trục hoành. Eviews ngầm định biến xếp sau nằm trên trục tung, biến xếp trước nằm ở trục hoành.

Xác định được hình ảnh của đường hồi quy bằng cách chọn:

[Group] **View** → **Graph** → **Scatter** → **Scatter with Regression**

Cửa sổ [Global Fit Option], nếu không định dạng đặc biệt, chọn OK, kết quả là đồ thị điểm với đường hồi quy



### Lưu đồ thị

Có thể lưu đồ thị để chèn vào các chương trình soạn thảo văn bản

Cửa sổ [Group] có đồ thị, nhấn tổ hợp Ctrl+C, mở cửa sổ [Graph Metafile], chọn OK, mở văn bản và nhấn Ctrl+V để dán đồ thị đã lưu.

Thay đổi định dạng đồ thị:

Tại cửa sổ [Group] có đồ thị, chọn **Freeze**, mở cửa sổ [Graph].

Với lựa chọn [Graph] **Proc** → **Options** : cửa sổ [Graph Option] với các lựa chọn cho định dạng.

[Graph] **Proc** → **Add text**: Thêm dòng chữ vào đồ thị

[Graph] **Proc** → **Save Graph**: lưu đồ thị để chèn vào văn bản

[Graph] **Name**: lưu đồ thị dưới dạng một chủ thể, đặt tên.

Ngoài ra còn nhiều lựa chọn khác với đồ thị

Bên cạnh cách chọn vẽ đồ thị từ cửa sổ [Graph], có thể có cách khác để vẽ đồ thị: Cửa sổ [Eviews] **Quick** → **Graph**: chọn loại đồ thị, và thứ tự của các biến để vẽ đồ thị.

**Vẽ đồ thị trên nhiều hệ tọa độ**: mỗi biến trên một hệ tọa độ riêng

[Group] **Multiple Graphs**.

## 2.4 Thống kê mô tả

Các thống kê mô tả về các biến và tương quan giữa các biến là những thông tin cơ bản để đánh giá về biến.

### Thống kê mô tả của các biến

Cửa sổ [Group] View → Descriptive Stats → Common Sample

		X2	X3	Y
Trung bình	Mean	28.33333	13.00000	30.40000
Trung vị	Median	29.00000	13.00000	30.70000
Tối đa	Maximum	36.00000	28.00000	36.80000
Tối thiểu	Minimum	20.00000	0.00000	23.60000
Độ lệch chuẩn	Std. Dev.	4.960450	8.800826	4.063138
Hệ số bất đối xứng	Skewness	-0.198458	0.110321	-0.163204
Hệ số nhọn	Kurtosis	1.965517	1.967268	2.300945
Thống kê JB	Jarque-Bera	0.613848	0.557609	0.297610
Mức xác suất	Probability	0.735706	0.756688	0.861737
Tổng	Sum	340.0000	156.0000	364.8000
Tổng bình phương chênh lệch	Sum Sq. Dev.	270.6667	852.0000	181.6000
Số quan sát	Observations	12	12	12

Thống kê JB và P-value dùng để kiểm định về cặp giả thuyết:

$H_0$ : biến phân phối theo quy luật Chuẩn

$H_1$ : biến không phân phối theo quy luật Chuẩn

[?] - Trung bình, phương sai biến nào lớn nhất, nhỏ nhất

- Trong mẫu, các biến có phân phối lệch trái hay lệch phải

- Qua kiểm định JB, các biến có phân phối chuẩn không?

### Hệ số tương quan

[Group] View → Correlations → Common Sample

	X2	Y
X2	1.000000	0.636883
Y	0.636883	1.000000

## Phương sai - hiệp phương sai

[Group] **View** → **Covariances** → **Common Sample**

	X2	Y
X2	22.55556	11.76667
Y	11.76667	15.13333

## Kiểm định so sánh các biến

Kiểm định so sánh trung bình, trung vị, phương sai của hai biến

[Group] **View** → **Test of Equality**, mở cửa sổ [Test Between Series], có ba lựa chọn kiểm định

Kiểm định sự bằng nhau của hai trung bình: **Mean**

Test for Equality of Means Between Series			
Sample: 1 12			
Included observations: 12			
Method	df	Value	Probability
t-test	22	1.116504	0.2763
Anova F-statistic	(1, 22)	1.246580	0.2763

Kiểm định sự bằng nhau của hai phương sai: **Variance**

Test for Equality of Variances Between Series			
Included observations: 12			
Method	df	Value	Probability
F-test	(11, 11)	1.490455	0.5190
Siegel-Tukey		0.086621	0.4045
Bartlett	1	0.416188	0.5188
Levene	(1, 22)	0.780029	0.3867
Brown-Forsythe	(1, 22)	0.747508	0.3966

[?] - Có thể cho rằng trung bình của X2 và Y là bằng nhau không?

- Có thể cho rằng phương sai của X2 và Y bằng nhau không?



## 2.5 Đặt biến mới

Bên cạnh các biến đã nhập số liệu là  $X_2$ ,  $X_3$  và  $Y$ , có thể đặt các biến mới từ các biến đã có hoặc nhập biến số mới.

Ví dụ: cần đặt biến mới tổng thu nhập:  $Z = X_2 + X_3$

□ [Eviews] **Quick** → **Generates Series**: mở cửa sổ [Generate Series by Equation]

Tại ô Enter Equation gõ:  $Z = X2 + X3$

Tại ô Sample, ngầm định mẫu mà lệnh đặt biến có tác dụng là toàn bộ. Khi cần có thể thay đổi mẫu này. Nhấn OK để chấp nhận.

Khi không cần thay đổi mẫu, có thể sử dụng lệnh trong Cửa sổ lệnh.

□ [Cửa sổ lệnh] **GENR Z = X2 + X3**

Biến  $Z$  được tạo ra nằm trong cửa sổ [Workfile].

### Các hàm và lệnh cơ bản

Ngoài các phép toán cơ bản: cộng [+], trừ [-], nhân [\*], chia [/], lũy thừa [^], các hàm cơ bản của Eviews như sau:

Tên	Ý nghĩa	Ví dụ	Kết quả
LOG	Logarit tự nhiên	$LX = \text{LOG}(X)$	$LX_i = \ln(X_i)$
EXP	Hàm mũ cơ số tự nhiên	$EX = \text{EXP}(X)$	$EX_i = e^{X_i}$
ABS	Giá trị tuyệt đối	$AX = \text{ABS}(X)$	$AX_i =  X_i $
SIN	Hàm sin	$\text{SIN}(X)$	$\text{Sin}(X)$
COS	Hàm cosin	$\text{COS}(X)$	$\text{Cosin}(X)$
(-J)	Trễ bậc J	$X4 = X(-4)$	$X4_i = X_{i-4}$
D(X)	Sai phân bậc 1	$DX = \text{D}(X)$	$DX_i = X_i - X_{i-1}$

### § 3 ƯỚC LƯỢNG MÔ HÌNH HỒI QUY ĐƠN

(Tiếp theo § 2)

Với bộ số liệu từ bài 2, nhận xét qua đồ thị giữa Y và X<sub>2</sub>, thấy giữa hai biến có xu thế cùng chiều, hệ số tương quan giữa hai biến bằng 0,636883 dương, do đó có thể xác định một mô hình hồi quy trong đó Chi tiêu (Y) phụ thuộc vào Lương (X<sub>2</sub>) dạng tuyến tính với hai hệ số.

Hàm hồi quy tổng thể PRF:

$$E(Y | X_2) = \beta_1 + \beta_2 X_2$$

Dạng ngẫu nhiên  $Y = \beta_1 + \beta_2 X_2 + u$  Mô hình (3.1)

Nếu xét biến hằng số C luôn bằng 1, có thể viết mô hình (3.1):

$$Y = \beta_1 C + \beta_2 X_2 + u$$

Với mẫu kích thước là 12 gồm 12 quan sát đã nhập, hàm hồi quy mẫu SRF có dạng:

$$\hat{Y}_i = \hat{\beta}_1 + \hat{\beta}_2 X_{2i}$$

$$Y_i = \hat{\beta}_1 + \hat{\beta}_2 X_{2i} + e_i$$

Dùng phương pháp Bình phương nhỏ nhất (Least Squares - LS) với bộ số liệu đã nhập, tính toán các ước lượng, và các thống kê cần thiết dành cho phân tích.

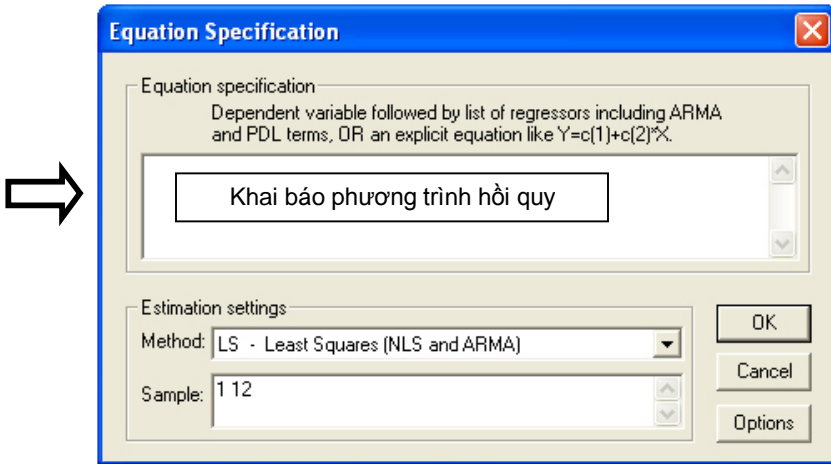
#### 3.1 Ước lượng mô hình

□ [Eviews] **Quick** → **Estimate Equation**

(Ngoài ra còn 4 cách khác, được đề cập sau)

Cửa sổ định dạng phương trình xuất hiện.

Cửa sổ [Equation Specification] gồm ô khai báo phương trình hồi quy, phương pháp ước lượng, mẫu để ước lượng.



Với mô hình hồi quy Y theo X2 có hệ số chặn, có hai kiểu khai báo:

Kiểu 1:  $Y = C(1) + C(2)*X2$

Kiểu 2:  $Y \ C \ X2$

Sử dụng kiểu khai báo thứ hai, phương pháp LS – bình phương nhỏ nhất, mẫu 1 đến 12, được kết quả ở cửa sổ [Equation]

Dependent Variable: Y					
Method: Least Squares					
Date: ...Time: ...					
Sample: 1 12					
Included observations: 12					
Variable	Coefficient	Std. Error	t-Statistic	Prob.	
C	15.61921	5.737035	2.722523	0.0215	
X2	0.521675	0.199698	2.612324	0.0259	
R-squared	0.405619	Mean dependent var		30.40000	
Adjusted R-squared	0.346181	S.D. dependent var		4.063138	
S.E. of regression	3.285415	Akaike info criterion		5.367875	
Sum squared resid	107.9395	Schwarz criterion		5.448692	
Log likelihood	-30.20725	F-statistic		6.824238	
Durbin-Watson stat	1.680815	Prob(F-statistic)		0.025937	

Nếu sử dụng cách khai báo thứ nhất:  $Y = C(1) + C(2)*X2$  thì kết quả chỉ khác phần thể hiện các hệ số như sau

$Y = C(1) + C(2)*X2$				
	Coefficient	Std. Error	t-Statistic	Prob.
C(1)	15.61921	5.737035	2.722523	0.0215
C(2)	0.521675	0.199698	2.612324	0.0259

Các thông tin khác vẫn giữ nguyên

Bên cạnh Cách 1 [Eviews] Quick → Estimate Equation như trên, còn một số cách sau:

**Cách 2:** [Eviews] **Objects** → **New Object** → **Equation** → OK

**Cách 3:** Chọn X2, Y thành cửa sổ [Group] **Procs** → **Make Equation**

**Cách 4:** Chọn X2 và Y, nhấn chuột phải → **Open** → **As Equation**

**Cách 5:** [Cửa sổ lệnh] **LS Y C X2**

Cách sử dụng lệnh LS Y C X2 là đơn giản nhất, sẽ được sử dụng trong các phần sau.


Các kiểu thể hiện kết quả hồi quy

[Equation] **View** → **Representations** : được các kết quả

Estimation Command: ===== LS Y C X	Câu lệnh để ước lượng (viết trong cửa sổ lệnh)
Estimation Equation: ===== $Y = C(1) + C(2)*X2$	Phương trình hồi quy
Substituted Coefficients: ===== $Y = 15.61921182 + 0.5216748768*X2$	Kết quả ước lượng

[Equation] **View** → **Estimation Output** : bảng kết quả chi tiết

### 3.2 Lưu kết quả

□ [Equation] **Name:** cửa sổ [Object Name], tên ngầm định cho kết quả ước lượng phương trình hồi quy là eq01. Nếu chọn OK, thì chủ thể  eq01 được tạo ra trong Workfile.

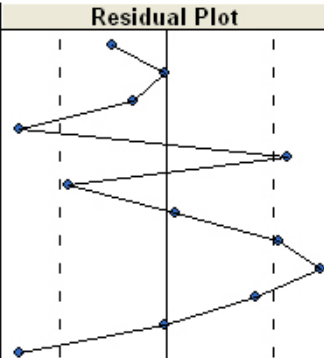
Muốn sao và dán kết quả ước lượng để chèn vào các văn bản, đánh dấu toàn bộ bảng kết quả chi tiết, nhấn chuột phải, chọn Copy, chọn loại giữ nguyên định dạng hoặc không giữ định dạng, rồi tại nơi cần dán chọn Paste; hoặc sử dụng tổ hợp Ctrl+C và Ctrl+V.

### 3.3 Xem phần dư và giá trị ước lượng (giá trị tương hợp)

Sau khi ước lượng một mô hình hồi quy, có thể đánh giá kết quả thông qua các giá trị phần dư (*residuals*) và các giá trị ước lượng của biến phụ thuộc (còn gọi là giá trị tương hợp – *fitted values*).

□ [Equation] **View** → **Actual, Fitted, Residual** → **Actual, Fitted, Residual Table**

Được bảng giá trị và đồ thị phần dư, Actual là giá trị thực tế bộ số liệu:  $Y_i$ , Fitted là giá trị ước lượng bởi hàm hồi quy  $\hat{Y}_i$ , Residual là giá trị phần dư:  $e_i$  hay  $\hat{u}_i$ .

obs	Actual	Fitted	Residual	Residual Plot
1	24.4000	26.0527	-1.65271	
2	31.2000	31.2695	-0.06946	
3	29.2000	30.2261	-1.02611	
4	23.6000	28.1394	-4.53941	
5	36.0000	32.3128	3.68719	
6	31.4000	34.3995	-2.99951	
7	32.6000	32.3128	0.28719	
8	36.8000	33.3562	3.44384	
9	32.8000	28.1394	4.66059	
10	29.8000	27.0961	2.70394	
11	30.2000	30.2261	-0.02611	
12	26.8000	31.2695	-4.46946	

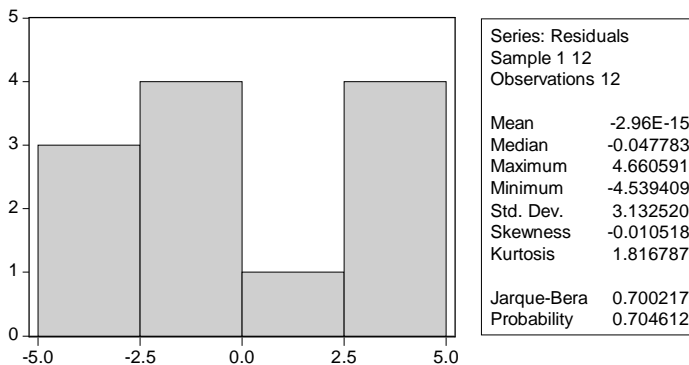
[?] - Quan sát nào có giá trị thực tế và giá trị ước lượng gần nhau nhất, xa nhau nhất? Khi đó giá trị thực tế cao hơn hay thấp hơn giá trị ước lượng?

- Với những quan sát nào thì đường hồi quy mô tả gần đúng nhất sự biến động của biến phụ thuộc Y?

Có thể xem đồ thị của giá trị thực tế, giá trị ước lượng, và phần dư trên cùng hệ tọa độ hoặc đồ thị của riêng phần dư bằng cách chọn Actual, Fitted, Residual Graph. Các đồ thị này có thể lưu lại hoặc cắt dán vào các văn bản tương tự như đồ thị của các biến X2, Y.

### Kiểm định tính phân phối chuẩn của phần dư

[Equation] View → Residual Tests → Histogram-Normality Test



[?] - Phần dư phân phối lệch trái hay lệch phải hay đối xứng?

- Sai số có phân phối chuẩn không?

Thực hiện hồi quy Chi tiêu (Y) theo Thu nhập khác (X3) và giải thích ý nghĩa; so sánh kết quả với mô hình Y phụ thuộc X2. Hồi quy Chi tiêu theo Tổng thu nhập (X2 + X3), khi đó ước lượng điểm tiêu dùng tự định và khuynh hướng tiêu dùng cận biên bằng bao nhiêu?

## § 4 ƯỚC LƯỢNG MÔ HÌNH HỒI QUY BỘỊ

Bài này sẽ sử dụng bộ số liệu **YWKM** có sẵn nằm trong thư mục **DATA2012** để thực hành. Đây là bộ số liệu về sản lượng và chi phí sản xuất của một số doanh nghiệp

### 4.1 Mở tệp số liệu có sẵn

[Eviews] **File** → **Open**

Chọn tệp số liệu **YWKM** trong thư mục tương ứng

Cửa sổ [Workfile] gồm các biến số xếp theo thứ tự K, M, TC, W, Y, với Y là sản lượng, W là chi cho công nhân, M là chi cho quản lý, K là chi phí khác gồm máy móc thiết bị nguyên nhiên vật liệu – hay còn gọi là chi phí vốn, TC là tổng chi phí. Như vậy TC là tổng của W, M, K.

### 4.2 Đánh giá chung về các biến

Xem thống kê đặc trưng mẫu của tất cả các biến (Descriptive Stats)

[?] - Trung bình của các biến bằng bao nhiêu?

- Biến nào có sự biến động tuyệt đối đo bằng phương sai lớn nhất, nhỏ nhất? Biến nào có sự biến động tương đối đo bằng hệ số biến thiên lớn nhất, nhỏ nhất?

- Biến nào đối xứng nhất, bất đối xứng nhất?

- Biến nào có phân phối chuẩn?

Xem ma trận tương quan của các biến (Correlations)

[?] - Cặp biến nào có tương quan ngược chiều?

- Cặp biến nào có tương quan chặt nhất, lỏng nhất?

- Sản lượng tương quan với biến nào chặt chẽ nhất?

### 4.3 Ước lượng mô hình

Xét mô hình  $E(Y | W, K) = \beta_1 + \beta_2 W + \beta_3 K$  (MH 4.1)

Hay  $Y_i = \beta_1 C_i + \beta_2 W_i + \beta_3 K_i + u_i$

□ [Cửa sổ lệnh] **LS Y C W K** ↵

Dependent Variable: Y				
Sample: 1 100				
Included observations: 100				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-485.9608	95.85601	-5.069695	0.0000
W	2.214092	0.050943	43.46253	0.0000
K	1.292811	0.044404	29.11470	0.0000
R-squared	0.964118	Mean dependent var	3707.680	
Adjusted R-squared	0.963378	S.D. dependent var	1425.836	
S.E. of regression	272.8616	Akaike info criterion	14.08535	
Sum squared resid	7221985.	Schwarz criterion	14.16350	
Log likelihood	-701.2674	F-statistic	1303.136	
Durbin-Watson stat	2.090510	Prob(F-statistic)	0.000000	

[?] - Giải thích ý nghĩa các ước lượng hệ số? Kết quả có phù hợp với lý thuyết kinh tế không?

- Hàm hồi quy có phù hợp không? Ý nghĩa của điều này là gì?

- Trong mẫu, các biến độc lập W và K giải thích bao nhiêu % sự biến động của biến Y?

- Các hệ số góc có ý nghĩa thống kê không? Điều này có ý nghĩa như thế nào?

Xem giá trị ước lượng biến phụ thuộc (giá trị tương hợp) và phần dư

□ [Equation] **Actual, Fitted, Residual** → **Actual, Fitted, Residual Table**: để xem giá trị ước lượng biến phụ thuộc và phần dư.

□ [Equation] **Actual, Fitted, Residual** → **Actual, Fitted, Residual Graph**: để xem đồ thị các giá trị.



#### 4.4 Phương sai, hiệp phương sai các ước lượng hệ số

Để có thể ước lượng, kiểm định về hơn một hệ số hồi quy, cần có ước lượng phương sai, hiệp phương sai của các ước lượng hệ số.

□ [Equation] **Covariance Matrix**

	C	W	K
C	9188.374	-2.461994	-3.582358
W	-2.461994	0.002595	0.000123
K	-3.582358	0.000123	0.001972

[?] - Phương sai của các ước lượng hệ số bằng bao nhiêu?

- Hiệp phương sai ước lượng hai hệ số góc bằng bao nhiêu?

- Hiệp phương sai các ước lượng hệ số được dùng trong trường hợp nào?

- Sử dụng kết quả hiệp phương sai, và bảng kết quả hồi quy mô hình từ phần đầu, cho biết nếu W và K cùng tăng một đơn vị thì trung bình của Y thay đổi như thế nào?

- Kiểm định giả thuyết cho rằng: hệ số của biến W lớn gấp đôi hệ số của biến K.

- Nếu W tăng một đơn vị, nhưng K giảm một đơn vị, thì Y sẽ tăng lên hay giảm đi, tăng lên (giảm đi) tối đa bao nhiêu, tối thiểu bao nhiêu?

#### 4.5 Kiểm định về các hệ số - kiểm định Wald

Eviews thực hiện kiểm định ràng buộc về các hệ số hồi quy bằng kiểm định Wald, sử dụng thống kê F và thống kê  $\chi^2$ .

Xét cặp giả thuyết

$$H_0: \beta_2 = 2,3 \quad H_1: \beta_2 \neq 2,3$$

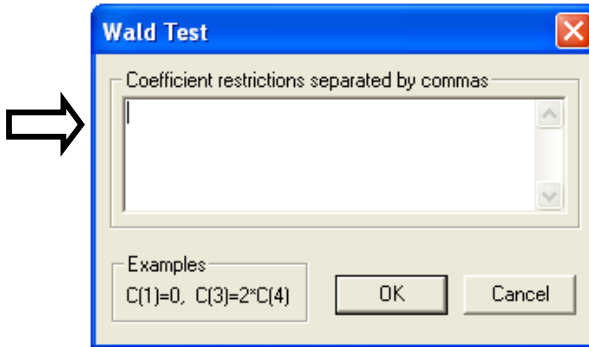
$$\text{Mô hình ban đầu: } Y_i = \beta_1 + \beta_2 W_i + \beta_3 K_i + u_i \quad (4.1)$$

$$\text{Mô hình có ràng buộc: } Y_i - 2.3W_i = \beta_1' + \beta_3' K_i + u_i \quad (4.2)$$

Thống kê F: 
$$F = \frac{(RSS_{(4,2)} - RSS_{(4,1)}) / m}{RSS_{(4,1)} / (n - k_{(4,1)})}$$

Tại cửa sổ kết quả hồi quy

□ [Equation] → View → Coefficient Tests → Wald Coefficient Restrictions ⇒ Cửa sổ khai báo giả thuyết cần kiểm định về các hệ số:



Quy ước đánh số các hệ số C(1), C(2),... theo thứ tự khai báo của phương trình hồi quy, do đó nếu khi khai báo biến là C W K thì hệ số chặn là C(1), hệ số  $\beta_w$  là C(2),  $\beta_K$  là C(3).

Tại cửa sổ Wald Test: **C(2) = 2.3** <Ok>

⇒ Kết quả kiểm định bằng kiểm định Wald, kiểm định F và  $\chi^2$ .

Wald Test:			
Equation: Untitled			
Null Hypothesis: C(2) = 2.3			
F-statistic	2.843870	Probability	0.094937
Chi-square	2.843870	Probability	0.091723

[?] - Bậc tự do của kiểm định F bằng bao nhiêu?

- Dùng thống kê F và P-value để kết luận về giả thuyết  $\beta_2 = 2.3$
- Kiểm định  $\chi^2$  cho kết luận thế nào về giả thuyết trên?

Tương tự, kiểm định giả thuyết: hệ số của biến W gấp đôi hệ số biến K, hay C(2) = 2\*C(3)

## 4.6 Kiểm định bỏ bớt biến

$$Y_i = \beta_1 + \beta_2 W_i + \beta_3 K_i + u_i \quad (4.1)$$

Muốn kiểm định bỏ biến K:

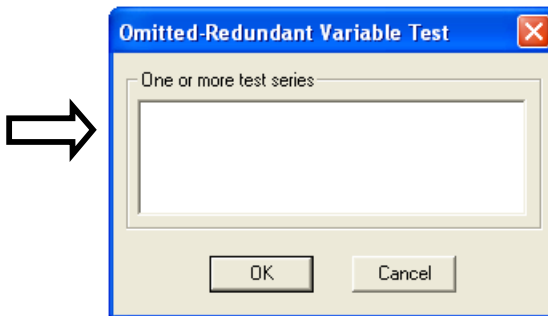
$$Y_i = \beta_1' + \beta_2' W_i + u_i \quad (4.3)$$

$H_0: \beta_3 = 0$  : có thể bỏ biến K

$H_1: \beta_3 \neq 0$  : không thể bỏ biến K

□ [Equation] View → Coefficient Tests → Redundant Variables

⇒ Cửa sổ khai báo biến **Omitted-Redundant Variable Test**



→ Gõ biến muốn bỏ: **K** <Ok>

⇒ Kết quả hồi quy sau khi đã bỏ biến K.

Redundant Variables: K				
F-statistic	847.6660	Probability	0.000000	
Log likelihood ratio	227.6120	Probability	0.000000	
Test Equation:				
Dependent Variable: Y		Included observations: 100		
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	1862.909	160.7195	11.59105	0.0000
W	2.133128	0.157928	13.50698	0.0000
R-squared	0.650547	Mean dependent var	3707.680	
Sum squared resid	70333643	Schwarz criterion	16.39357	
Log likelihood	-815.0734	F-statistic	182.4384	
Durbin-Watson stat	2.022167	Prob(F-statistic)	0.000000	

[?] - Thống kê  $F = 847.666$  tính cụ thể như thế nào?

- Bậc tự do của kiểm định  $F$  bằng bao nhiêu?

- Dùng thống kê  $F$  và  $P$ -value kiểm định giả thuyết  $H_0$ : “nên bỏ biến  $K$  khỏi mô hình”.

Thực hiện kiểm định tương tự, có nên bỏ biến  $W$  khỏi mô hình hay không?

## 4.6 Kiểm định thêm biến

Thêm hai biến  $W^2$  và  $K^2$  vào mô hình

$$Y_i = \beta_1 + \beta_2 W_i + \beta_3 K_i + u_i \quad (4.1)$$

$$Y_i = \beta_1' + \beta_2' W_i + \beta_3' K_i + \beta_4 W_i^2 + \beta_5 K_i^2 + u_i \quad (4.4)$$

$$H_0: \beta_3 = \beta_4 = 0$$

$$H_1: \beta_3^2 + \beta_4^2 \neq 0$$

Tại cửa sổ kết quả ước lượng mô hình [Equation]

□ [Equation] **View** → **Coefficient Tests** → **Omitted Variables**

⇒ Cửa sổ khai báo bớt (thêm) biến

→ Gõ biến muốn thêm: **W^2 K^2** <Ok>

⇒ Kết quả hồi quy sau khi thêm biến.

Omitted Variables: W^2 K^2			
F-statistic	35.30513	Probability	0.000000
Log likelihood ratio	55.57603	Probability	0.000000

[?] - Nêu cách tính thống kê  $F = 35.04656$ , bậc tự do tương ứng với kiểm định  $F$  là bao nhiêu?.

- Việc cho thêm hai biến vào mô hình có ý nghĩa thống kê không?

- Dùng kiểm định  $T$  để kiểm định xem có cần đưa biến mới vào mô hình không?

- Những kết luận trên có gì thay đổi nếu mức ý nghĩa bằng 1%?

### 4.7 Mô hình đa thức của biến độc lập

Mô hình 
$$Y_i = \beta_1 + \beta_2 W_i + \beta_3 K_i + \beta_2 W_i^2 + \beta_3 K_i^2 + u_i \quad (4.5)$$

□ Cửa sổ lệnh: **LS Y C W K W^2 K^2** ↵

Dependent Variable: Y				
Included observations: 100				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-1304.324	202.4025	-6.444208	0.0000
W	3.196177	0.131148	24.37084	0.0000
K	1.791658	0.198742	9.014987	0.0000
W^2	-0.000462	5.92E-05	-7.806547	0.0000
K^2	-0.000124	4.58E-05	-2.703958	0.0081
R-squared	0.979417	Mean dependent var	3707.680	
Durbin-Watson stat	2.252379	Prob(F-statistic)	0.000000	

Đây chính là kết quả của mô hình thêm biến trong mục 4.6

[?] - Có thể cho rằng mô hình thể hiện quy luật lợi ích cận biên giảm dần trong kinh tế hay không? Vì sao

- Dựa vào kết quả ước lượng, viết phương trình thể hiện tác động của biến W đến sản lượng, theo kết quả đó, đâu là mức chi cho công nhân làm tối đa hóa sản lượng?

### 4.8 Mô hình có logarit

#### Mô hình dạng hàm mũ Cobb-Douglas

Mô hình 
$$E(Y | W, K) = e^{\beta_1} W^{\beta_2} K^{\beta_3}$$

$$Y_i = e^{\beta_1} W_i^{\beta_2} K_i^{\beta_3} e^{u_i} \quad (4.6)$$

$$\hat{Y}_i = e^{\hat{\beta}_1} W_i^{\hat{\beta}_2} K_i^{\hat{\beta}_3}$$

Logarit hai vế 
$$\ln Y_i = \beta_1 + \beta_2 \ln W_i + \beta_3 \ln K_i + u_i$$

□ Cửa sổ lệnh: **LS LOG(Y) C LOG(W) LOG(K)** ↵

Dependent Variable: LOG(Y)				
Included observations: 100				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.416571	0.114175	3.648529	0.0004
LOG(W)	0.476395	0.005883	80.97390	0.0000
LOG(K)	0.621661	0.014506	42.85566	0.0000
R-squared	0.988628	Mean dependent var	8.136574	
Sum squared resid	0.204993	Schwarz criterion	-3.213916	
Log likelihood	167.6036	F-statistic	4216.348	
Durbin-Watson stat	1.927110	Prob(F-statistic)	0.000000	

Theo kết quả này  $\hat{Y}_i = e^{0.409494} W^{0.478281} K^{0.620866}$

[?] - Giải thích ý nghĩa kết quả hồi quy

- Các hệ số có ý nghĩa thống kê không?

- Mô hình dạng hàm mũ giải thích bao nhiêu % sự biến động của biến GDP, có giải thích được nhiều hơn dạng hàm tuyến tính không?

- Kiểm định giả thuyết cho rằng quá trình sản xuất là không đổi theo quy mô của chi cho công nhân và chi phí vốn.

### Mô hình dạng nửa logarit

□ Hồi quy mô hình  $Y_i = \beta_1 + \beta_2 \ln W_i + \beta_3 \ln K_i + u_i$

[?] - Giải thích ý nghĩa kết quả ước lượng, tác động của biến nào là mạnh hơn đến sự thay đổi của Y?

□ Hồi quy mô hình  $Y_i = e^{\beta_1 + \beta_2 W_i + \beta_3 K_i + u_i}$

Hay  $\ln Y_i = \beta_1 + \beta_2 W_i + \beta_3 K_i + u_i$

## § 5 MÔ HÌNH VỚI BIẾN GIÁ

Sử dụng tệp **WORLD\_EXIM\_2010** trong thư mục **DATA2012**.

Đây là bộ số liệu nguồn từ Worldbank năm 2010 với 220 quốc gia và vùng lãnh thổ, trong đó EX là xuất khẩu, IM là nhập khẩu, GDP là tổng sản phẩm quốc nội.

RANK là xếp hạng về mức thu nhập của các nước, giá trị như sau:

- (1) Các nước thu nhập cao thuộc khối OECD
- (2) Các nước thu nhập cao không thuộc khối OECD
- (3) Các nước thu nhập trên trung bình
- (4) Các nước thu nhập trung bình
- (5) Các nước thu nhập dưới trung bình

Nhóm (1)(2) là các nước phát triển, còn lại là các nước đang phát triển. Việt Nam thuộc nhóm (4).

### 5.1 Mô hình có biến định tính

Để đánh giá về xuất khẩu (EX) của các nước OECD so sánh với các nước khác, cần đặt biến giả có dạng:

$D1 = 1$ : nếu là nước thuộc OECD

$D1 = 0$ : nếu là nước không thuộc OECD

Các mô hình sẽ xét như sau:

#### Mô hình chỉ có biến độc lập là biến giả

Xem xét Xuất khẩu trung bình của các nước OECD có khác các nước còn lại hay không, có dạng:

$$E(EX | D1) = \beta_1 + \beta_2 D1 \quad (5.1)$$

Với nước thuộc OECD:  $E(EX | D1 = 1) = \beta_1 + \beta_2$

Với nước khác:  $E(EX | D1 = 0) = \beta_1$

Nếu  $\beta_2 \neq 0$  thì có sự khác biệt.

## Mô hình có biến độc lập là biến định lượng và biến giả

Biến độc lập định lượng là GDP. Xét mối quan hệ Xuất khẩu phụ thuộc GDP có dạng

$$E(EX | GDP) = [\text{hệ số chặn}] + [\text{hệ số góc}]GDP$$

$$\text{hay } EX_i = [\text{hệ số chặn}] + [\text{hệ số góc}]GDP + u_i$$

### Biến giả tác động đến hệ số chặn

$$E(EX | GDP, D1) = \beta_1 + \beta_2 GDP + \beta_3 D1 \quad (5.2)$$

$$\text{Nước OECD: } E(EX | GDP, D1 = 1) = (\beta_1 + \beta_3) + \beta_2 GDP$$

$$\text{Nước khác: } E(EX | GDP, D1 = 0) = \beta_1 + \beta_2 GDP$$

### Biến giả tác động đến hệ số góc

$$E(EX | GDP, D1) = \beta_1 + \beta_2 GDP + \beta_3 D1 \times GDP \quad (5.3)$$

$$\text{Nước OECD: } E(EX | GDP, D1 = 1) = \beta_1 + (\beta_2 + \beta_3)GDP$$

$$\text{Nước khác: } E(EX | GDP, D1 = 0) = \beta_1 + \beta_2 GDP$$

### Biến giả tác động đến cả hai hệ số

$$E(EX | GDP, D1) = \beta_1 + \beta_2 GDP + \beta_3 D1 + \beta_4 D1 \times GDP \quad (5.4)$$

$$\text{Nước OECD: } E(EX | GDP, D1 = 1) = (\beta_1 + \beta_3) + (\beta_2 + \beta_4)GDP$$

$$\text{Nước khác: } E(EX | GDP, D1 = 0) = \beta_1 + \beta_2 GDP$$

## 5.2 Đặt biến giả

Cần đặt biến  $D1 = 0$  với mọi quan sát, sau đó riêng với các nước OECD (có  $RANK = 1$ ) thì đổi lại giá trị  $D1 = 1$ .

□ Tại Cửa sổ lệnh gõ lần lượt các lệnh, sau mỗi lệnh gõ Enter

**GENR D1 = 0** ↵ : đặt biến mới  $D1 = 0$  với mọi quan sát

**SMPL IF RANK = 1** ↵ : xét riêng các quan sát có  $RANK=1$

**D1 = 1** ↵ : đặt lại  $D1 = 1$

**SMPL 1 220** ↵ : lấy lại toàn bộ các quan sát



Sau câu lệnh SMPL IF, sẽ thấy trên cửa sổ Workfile có dòng:

**Sample: 1 220 if rank = 1**

Lưu ý phải có câu lệnh cuối, vì nếu thiếu, các thao tác về sau sẽ chỉ làm với các quan sát có RANK = 1. Nếu đủ câu lệnh cuối, trên cửa sổ Workfile phải xuất hiện lại là: **Sample 1 220**.

Có thể mở riêng cửa sổ biến D1, sẽ thấy giá trị bằng 1 tại các quan sát thứ 11, 12, 19..., còn lại bằng 0.

### 5.3 Hồi quy mô hình có biến giả

#### Mô hình (5.1)

□ Lệnh: **LS EX C D1** ↵

Dependent Variable: EX				
Sample(adjusted): 1 213				
Included observations: 158				
Excluded observations: 55 after adjusting endpoints				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	57078.39	21931.16	2.602616	0.0101
D1	295998.5	49511.86	5.978335	0.0000
R-squared	0.186400	Mean dependent var		115154.0
Log likelihood	-2185.192	F-statistic		35.74048
Durbin-Watson stat	2.472372	Prob(F-statistic)		0.000000

Số quan sát được điều chỉnh do có một số nước không có số liệu.

[?] - Viết hàm hồi quy với hai nhóm nước và giải thích ý nghĩa?

- Xuất khẩu trung bình của các nước OECD có khác các nước còn lại không? Cao hơn hay thấp?

- Ước lượng điểm và khoảng mức chênh lệch của xuất khẩu trung bình giữa hai nhóm nước?

### Mô hình (5.2): Biến giả tác động đến hệ số chặn

□ Lệnh: **LS EX C GDP D1** ↵

Dependent Variable: EX				
Included observations: 158				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	30735.85	11838.70	2.596218	0.0103
GDP	0.160118	0.008136	19.67955	0.0000
D1	111465.3	28162.60	3.957921	0.0001
R-squared	0.767451	Mean dependent var	115154.0	
Log likelihood	-2086.255	F-statistic	255.7626	
Durbin-Watson stat	2.640003	Prob(F-statistic)	0.000000	

- [?] - Viết hàm hồi quy với hai nhóm nước và giải thích ý nghĩa?
- Hệ số chặn có thực sự khác nhau giữa hai nhóm nước không?
  - Nếu cùng mức GDP thì xuất khẩu trung bình của hai nước thuộc hai nhóm bên nào nhiều hơn, nhiều hơn trong khoảng nào?

### Mô hình (5.3): Biến giả tác động đến hệ số góc

□ Lệnh: **LS EX C GDP D1\*GDP** ↵

Dependent Variable: EX				
Included observations: 158				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	39614.77	10638.14	3.723842	0.0003
GDP	0.266546	0.018845	14.14414	0.0000
D1*GDP	-0.110621	0.020052	-5.516777	0.0000
R-squared	0.785973	Mean dependent var	115154.0	
Log likelihood	-2079.698	F-statistic	284.6038	
Durbin-Watson stat	2.472840	Prob(F-statistic)	0.000000	

- [?] - Viết hàm hồi quy với hai nhóm nước và giải thích ý nghĩa?
- Hệ số góc có thực sự khác nhau giữa hai nhóm nước không?
  - Nếu cùng tăng GDP thì xuất khẩu trung bình nhóm nước nào tăng lên nhiều hơn? Nhiều hơn trong khoảng nào?

**Mô hình (5.4): Biến giả tác động đến cả hai hệ số**

Lệnh: **LS EX C GDP D1 D1\*GDP ↵**

Dependent Variable: EX				
Included observations: 158				
Excluded observations: 55 after adjusting endpoints				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	11209.38	10388.57	1.079010	0.2823
GDP	0.278806	0.016826	16.57005	0.0000
D1	162879.1	24876.43	6.547525	0.0000
D1*GDP	-0.142900	0.018463	-7.740000	0.0000
R-squared	0.832579	Mean dependent var	115154.0	
Log likelihood	-2060.296	F-statistic	255.2792	
Durbin-Watson stat	2.473535	Prob(F-statistic)	0.000000	

[?] - Viết hàm hồi quy với hai nhóm nước và giải thích ý nghĩa?  
 - Hệ số chặn và hệ số góc có khác nhau giữa hai nhóm nước không? Phân tích ý nghĩa của kết luận.

**Kiểm định sự đồng nhất bằng kiểm định F**

Với mô hình (5.4):

$$E(EX | GDP, D1) = \beta_1 + \beta_2 GDP + \beta_3 D1 + \beta_4 D1 \times GDP$$

Nếu  $\beta_3 = \beta_4 = 0$  thì hàm hồi quy là đồng nhất giữa hai nhóm nước. Qua kiểm định T với riêng từng hệ số thấy chúng khác 0, tuy nhiên vẫn cần kiểm định F về sự đồng thời bằng 0

[Equation] **View** → **Coefficient Tests** → **Redundant Variables**

Biến cần bỏ: **D1 D1\*GDP** <Ok>

⇒ Kết quả sau khi bỏ hai biến giả.

Redundant Variables: D1 D1*GDP			
F-statistic	40.76313	Probability	0.000000
Log likelihood ratio	67.12943	Probability	0.000000
Test Equation:			
Dependent Variable: EX			

Included observations: 158				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	48417.17	11467.23	4.222220	0.0000
GDP	0.170840	0.008025	21.28970	0.0000
R-squared	0.743948	Mean dependent var		115154.0
Durbin-Watson stat	2.622065	Prob(F-statistic)		0.000000

- [?] - Thống kê F của kiểm định thu hẹp hồi quy bằng bao nhiêu?  
 - Hàm hồi quy có đồng nhất giữa hai nhóm nước không?

## 5.4 Mô hình nhiều biến giả

Tương tự với các nước OECD, đặt các biến giả D2, D3, D4 tương ứng với các nước thuộc nhóm (2), nhóm (3), nhóm (4), và lấy các nước thuộc nhóm (5) làm trạng thái gốc.

Hồi quy các mô hình xem xét sự khác biệt về mức xuất khẩu giữa các nhóm nước so với nhóm nước nghèo nhất, có dạng:

Đánh giá về mức xuất khẩu trung bình:

$$E(EX | D1, D2, D3, D4) = \beta_1 + \beta_2 D1 + \beta_3 D2 + \beta_4 D3 + \beta_5 D4$$

Đánh giá về hệ số chặn

$$E(EX | GDP, D1, D2, D3, D4) = \beta_1 + \beta_2 D1 + \beta_3 D2 + \beta_4 D3 + \beta_5 D4 + \beta_6 GDP$$

Đánh giá về hệ số góc

$$E(EX | GDP, D1, D2, D3, D4) = \beta_1 + \beta_2 GDP + \beta_3 D1 * GDP + \beta_4 D2 * GDP + \beta_5 D3 * GDP + \beta_6 D4 * GDP$$

Tương tự, có thể đánh giá tác động đến cả hai hệ số.

Ngoài ra, có thể phân tích về Nhập khẩu và so sánh giữa các nhóm nước bằng phương pháp sử dụng biến giả.

## § 6 ĐỊNH DẠNG PHƯƠNG TRÌNH HỒI QUY

Thực hành với bộ số liệu **YWKM**

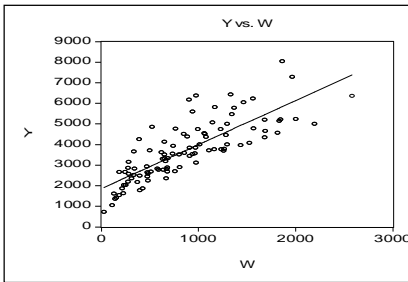
### 6.1 Phân tích qua đồ thị

Xét mối quan hệ sản lượng (Y) phụ thuộc vào chi phí cho công nhân (W), chọn biến W trước, rồi mới chọn biến Y, mở cửa sổ Group

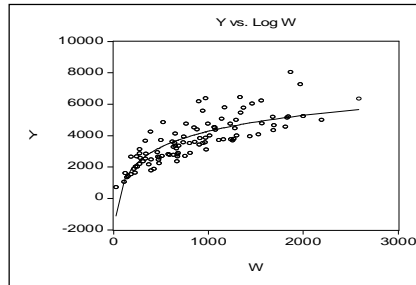
□ [Group] **View** → **Graph** → **Scatter with Regression**

Với bốn trường hợp:

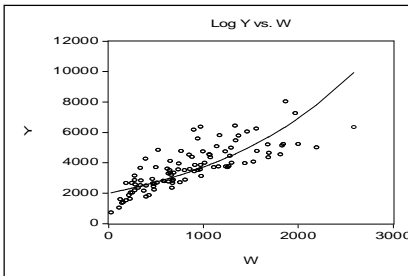
- (1) None-None: đường thẳng
- (2) None-Logathmic: đường cong logarit
- (3) Logathmic-None: hàm e mũ
- (4) Logathmic-Logathmic: hàm Cobb-Douglas



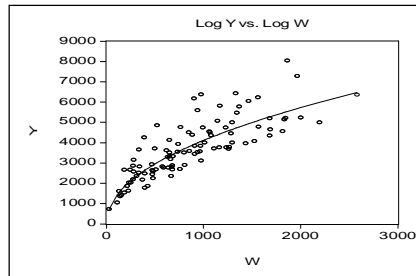
(1) Tuyến tính – tuyến tính



(2) Tuyến tính – Logarit



(3) Logarit – Tuyến tính



(4) Logarit – Logarit

Qua đó, nhận xét dạng đồ thị phù hợp với bộ số liệu, để xác định dạng của phương trình hồi quy.

## 6.2 Kiểm định Ramsey RESET

Xét mô hình tuyến tính – tuyến tính

$$Y_i = \beta_1 + \beta_2 W_i + u_i \quad (6.1)$$

□ **LS Y C W** ↵

⇒ Bảng kết quả hồi quy.

Dependent Variable: Y				
Included observations: 100				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	1862.909	160.7195	11.59105	0.0000
W	2.133128	0.157928	13.50698	0.0000
R-squared	0.650547	Mean dependent var	3707.680	
Durbin-Watson stat	2.022167	Prob(F-statistic)	0.000000	

Kiểm định Ramsey RESET sử dụng hồi quy phụ, khi thêm một biến (là bình phương của ước lượng biến phụ thuộc) vào mô hình gốc:

$$Y_i = \beta_1' + \beta_2' W_i + \alpha_1 \hat{Y}_i^2 + u_i \quad (6.2)$$

$H_0: \alpha_1 = 0$ : Mô hình gốc không thiếu biến, dạng hàm đúng

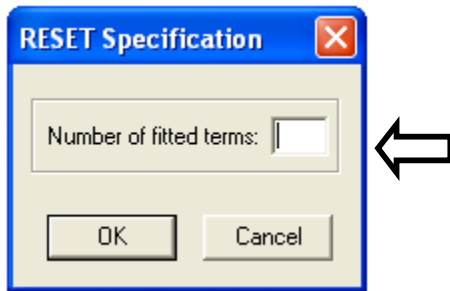
$H_1: \alpha_1 \neq 0$ : Mô hình gốc thiếu biến, dạng hàm sai

$$F_{qs} = \frac{R_{(6.2)}^2 - R_{(6.1)}^2}{1 - R_{(6.2)}^2} \times \frac{n - k_{(6.2)}}{1} \text{ so sánh với } F_{\alpha}^{(1, n - k_{(6.2)})}$$

□ [Equation] **View** → **Stability Test** → **Ramsey RESET Test...**

□ [RESET Specification] Number of fitted terms: **1**

⇒ Số phần tử muốn thêm vào, nhập là 1, rồi <OK>



Ramsey RESET Test:				
F-statistic	9.815365	Probability	0.002289	
Log likelihood ratio	9.639081	Probability	0.001905	
Test Equation:				
Dependent Variable: Y				
Included observations: 100				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	1837.378	154.1603	11.91862	0.0000
W	4.898633	0.895584	5.469762	0.0000
FITTED^2	-0.000157	5.02E-05	-3.132948	0.0023
R-squared	0.682659	Mean dependent var	3707.680	
Sum squared resid	63870618	Schwarz criterion	16.34323	
Log likelihood	-810.2538	F-statistic	104.3323	
Durbin-Watson stat	1.885336	Prob(F-statistic)	0.000000	

[?] - Thống kê F-statistic = 9,815 trong bảng có bậc tự do tương ứng bằng bao nhiêu?

- Theo kiểm định này, mô hình (6.1) có thiếu biến không, có thể nói dạng của phương trình hồi quy là đúng được không?

Khi muốn kiểm định bằng cách thêm 2 biến vào mô hình gốc, phương trình hồi quy phụ có dạng:

$$Y_i = \beta_1' + \beta_2'W_i + \alpha_1\hat{Y}_i^2 + \alpha_2\hat{Y}_i^3 + u_i \quad (6.3)$$

Giả thuyết  $H_0$ :  $H_0: \alpha_1 = \alpha_2 = 0$

[?] - Thực hiện kiểm định khi thêm 2 biến, và kết luận về dạng hàm của mô hình gốc?

### 6.3 Thay đổi phương trình hồi quy

Nhận thấy mô hình (6.1)  $Y_i = \beta_1 + \beta_2 W_i + u_i$  có dạng hàm không đúng, thực hiện một số việc thay đổi phương trình hồi quy, sau đó kiểm định bằng Ramsey để đưa ra phương trình phù hợp

**Thêm biến K:**  $Y_i = \beta_1 + \beta_2 W_i + \beta_3 K_i + u_i$

Dependent Variable: Y		Included observations: 100		
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-485.9608	95.85601	-5.069695	0.0000
W	2.214092	0.050943	43.46253	0.0000
K	1.292811	0.044404	29.11470	0.0000
R-squared	0.964118	Prob(F-statistic)		0.000000
Ramsey RESET Test:				
F-statistic	0.006172	Probability		0.937544
Log likelihood ratio	0.006429	Probability		0.936093

**Đổi dạng bậc hai:**  $Y_i = \beta_1 + \beta_2 W_i + \beta_3 W_i^2 + u_i$

Dependent Variable: Y		Included observations: 100		
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	1291.989	238.5519	5.415967	0.0000
W	3.649636	0.507138	7.196537	0.0000
W^2	-0.000715	0.000228	-3.132948	0.0023
R-squared	0.682659	Prob(F-statistic)		0.000000
Ramsey RESET Test:				
F-statistic	0.670094	Probability		0.415046
Log likelihood ratio	0.695590	Probability		0.404269

Tương tự, có thể sử dụng mô hình dạng Cobb-Douglas:  $Y_i = e^{\beta_1} W_i^{\beta_2} e^{u_i}$ .



## § 7 HIỆN TƯỢNG ĐA CỘNG TUYẾN

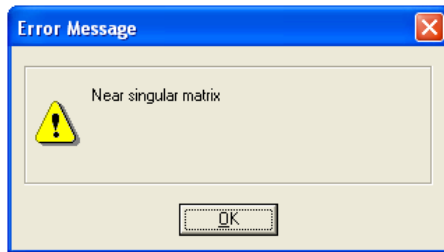
Sử dụng bộ số liệu **YWKM**

### 7.1 Đa cộng tuyến hoàn hảo

Xét mô hình  $Y_i = \beta_1 + \beta_2 W_i + \beta_3 K_i + \beta_4 M_i + \beta_5 TC_i + u_i$  (7.1)

□ **LS Y C W K M TC ↵**

⇒ Cửa sổ Error Message với thông báo:



Mô hình có hiện tượng đa cộng tuyến hoàn hảo nên các tính toán không thực hiện được.

Singular matrix là ma trận suy biến: các cột của ma trận các biến độc lập có liên hệ tuyến tính với nhau.

[?] - Tại sao mô hình có đa cộng tuyến hoàn hảo, cách loại trừ đa cộng tuyến hoàn hảo trong mô hình này thế nào?

### 7.2 Đa cộng tuyến không hoàn hảo

Xét mô hình sau khi bỏ biến TC (tổng chi phí) khỏi mô hình (7.1)

$$Y_i = \beta_1 + \beta_2 W_i + \beta_3 K_i + \beta_4 M_i + u_i \quad (7.2)$$

□ **LS Y C W K M ↵**

Dependent Variable: Y				
Included observations: 100				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-488.5271	96.19136	-5.078701	0.0000
W	0.531746	2.452609	0.216808	0.8288
K	0.875197	0.610312	1.434016	0.1548
M	8.406298	12.25247	0.686090	0.4943
R-squared	0.964293	Mean dependent var		3707.680
Log likelihood	-701.0228	F-statistic		864.1738
Durbin-Watson stat	2.095488	Prob(F-statistic)		0.000000

[?] - Với các kiểm định T, các biến độc lập có giải thích cho biến phụ thuộc hay không?

- Hệ số xác định bằng bao nhiêu?
- Bằng kiểm định F, cho biết hàm hồi quy có phù hợp không?
- Mâu thuẫn giữa các kiểm định cho thấy điều gì?

### 7.3 Đánh giá mức độ đa cộng tuyến

Xem hệ số tương quan giữa các biến độc lập của mô hình (7.2)

Chọn **K, M, W** → **Group** → **View** → **Correlation**

	K	M	W
K	1.000000	0.224602	-0.054588
M	0.224602	1.000000	0.960526
W	-0.054588	0.960526	1.000000

[?] - Cặp biến nào có tương quan với nhau chặt nhất? Theo đánh giá này thì mô hình có những biến nào dễ có đa cộng tuyến nhất?

Với mô hình (7.2) có đa cộng tuyến, sử dụng các hồi quy phụ để đánh giá mức độ đa cộng tuyến giữa các biến độc lập. Hồi quy phụ một biến độc lập theo các biến còn lại, có thể có ba trường hợp:

**LS W C K M** ↙ (7.2a)

**LS K C W M** ↙ (7.2b)

**LS M C W K** ↙ (7.2c)

[?] - Qua các hồi quy phụ, có kết luận gì về quan hệ giữa các biến độc lập của (7.2)?

- Nêu cách khắc phục đơn giản hiện tượng trong mô hình (7.2)

### 7.3 Khắc phục hiện tượng đa cộng tuyến

Để giảm bớt mức độ đa cộng tuyến, có thể bỏ bớt biến độc lập khỏi mô hình (7.2). Khi đó cũng có thể có ba trường hợp: bỏ M, bỏ K, bỏ W. Với mỗi trường hợp, đánh giá lại mức độ đa cộng tuyến trong các mô hình sau khi bỏ biến bằng hồi quy phụ tương ứng.

Bỏ biến M:  $LS \quad Y \quad C \quad W \quad K \quad \downarrow$  (7.3)

Hồi quy phụ đánh giá:  $LS \quad W \quad C \quad K \quad \downarrow$  (7.3a)

Bỏ biến K:  $LS \quad Y \quad C \quad W \quad M \quad \downarrow$  (7.4)

Hồi quy phụ đánh giá:  $LS \quad W \quad C \quad M \quad \downarrow$  (7.4a)

Bỏ biến W:  $LS \quad Y \quad C \quad K \quad M \quad \downarrow$  (7.5)

Hồi quy phụ đánh giá:  $LS \quad K \quad C \quad M \quad \downarrow$  (7.5a)

[?] - Bỏ biến độc lập nào thì có được mô hình mới tốt hơn?

Có thể đổi dạng mô hình sang logarit, mâu thuẫn giữa các kiểm định trong mô hình (7.2) không còn nữa.

Dependent Variable: LOG(Y)				
Included observations: 100				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.618638	0.086769	7.129678	0.0000
LOG(W)	0.317445	0.017914	17.72070	0.0000
LOG(K)	0.517653	0.015590	33.20453	0.0000
LOG(M)	0.293691	0.032121	9.143369	0.0000
R-squared	0.993921	Mean dependent var	8.136574	
Durbin-Watson stat	1.987288	Prob(F-statistic)	0.000000	

## § 8 HIỆN TƯỢNG PHƯƠNG SAI SAI SỐ THAY ĐỔI

Sử dụng bộ số liệu **WORLD\_EXIM\_2010**

Chỉ xét riêng các nước phát triển, biến RANK bằng 1 và 2. Do đó để giới hạn số liệu với các nước này, tại cửa sổ lệnh:

**SMPL IF RANK < 3** ↵

Tại cửa sổ Workfile có dòng: **Sample 1 220 if rank < 3**

### 8.1 Hiện tượng

Xét mô hình  $IM_i = \beta_1 + \beta_2 GDP_i + \beta_3 EX_i + u_i$  (MH 8.1)

**LS IM C GDP EX** ↵

⇒ Bảng kết quả hồi quy.

Dependent Variable: IM				
Sample(adjusted): 11 203 IF RANK < 3				
Included observations: 45 after adjusting endpoints				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	1879.134	7997.118	0.234976	0.8154
GDP	0.050309	0.005070	9.922462	0.0000
EX	0.813882	0.030802	26.42261	0.0000
R-squared	0.990809	Mean dependent var	277422.0	
Sum squared resid	6.96E+10	Schwarz criterion	24.25097	
Durbin-Watson stat	0.669835	Prob(F-statistic)	0.000000	

Để nhận định về phương sai sai số, đánh giá qua giá trị phần dư

[Equation] **View** → **Actual, Fitted, Residual** → **Actual, Fitted, Residual Table**

Phần dư dao động quanh giá trị trung bình (bằng 0) không đều, tuy nhiên cảm nhận chưa đủ cơ sở đánh giá.

## 8.2 Kiểm định White

Kiểm định White dựa trên hồi quy bình phương phần dư (kí hiệu là RESID) theo bậc nhất và bậc hai của các biến độc lập. Có hai trường hợp: kiểm định không và có tích chéo giữa các biến độc lập.

### Kiểm định không có tích chéo

$$e_i^2 = \alpha_1 + \alpha_2 GDP_i + \alpha_3 GDP_i^2 + \alpha_4 EX_i + \alpha_5 EX_i^2 + v_i \quad (8.1a)$$

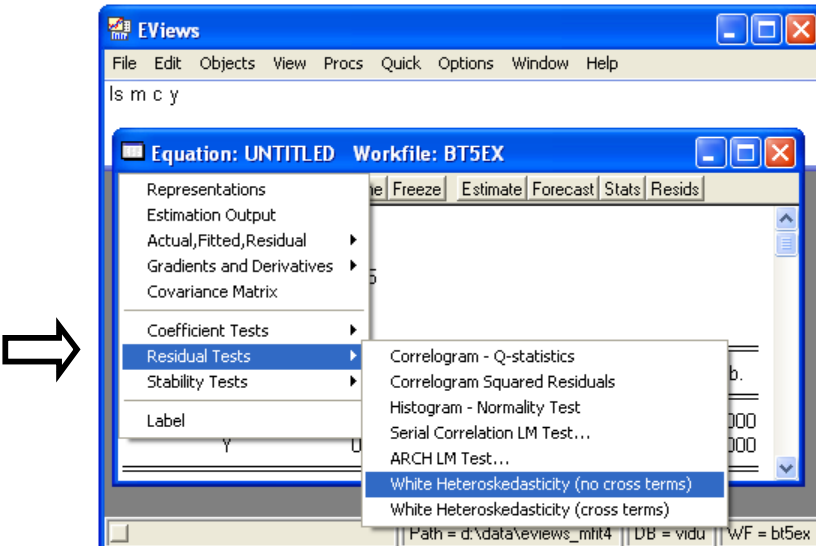
$H_0$ : Mô hình gốc có phương sai sai số không đổi (đồng đều)

$H_1$ : Mô hình gốc có phương sai sai số thay đổi (không đồng đều)

$$F_{qs} = \frac{R_{(8.1a)}^2}{1 - R_{(8.1a)}^2} \times \frac{n - k_{(8.1a)}}{k_{(8.1a)} - 1} \text{ so sánh với } F_{\alpha}^{(k_{(8.1a)} - 1, n - k_{(8.1a)})}$$

$$\chi_{qs}^2 = nR_{(8.1a)}^2 \text{ so sánh với } \chi_{\alpha}^2(k_{(8.1a)} - 1)$$

□ [Equation] View → Residual Test → White heteroskedasticity (no cross terms)



White Heteroskedasticity Test:				
F-statistic	15.59507	Probability	0.000000	
Obs*R-squared	27.41849	Probability	0.000016	
Test Equation:				
Dependent Variable: RESID^2				
Included observations: 45				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-3.72E+08	7.59E+08	-0.489639	0.6271
GDP	6930.669	1149.842	6.027496	0.0000
GDP^2	-0.000380	7.50E-05	-5.064444	0.0000
EX	-7843.361	5341.569	-1.468363	0.1498
EX^2	-0.000419	0.003244	-0.129023	0.8980
R-squared	0.609300	Mean dependent var	1.55E+09	
Log likelihood	-1046.442	F-statistic	15.59507	
Durbin-Watson stat	1.126117	Prob(F-statistic)	0.000000	

[?] - Qua các thống kê và P-value tương ứng, thực hiện kiểm định để kết luận về phương sai sai số của mô hình gốc?

- Hồi quy phụ cho thấy bình phương phần dư phụ thuộc biến độc lập nào?

- Các ước lượng hệ số của mô hình (8.1) có phải là ước lượng tốt nhất không? Sai số các ước lượng hệ số có phải là tốt nhất không?

- Nếu ước lượng trên chưa phải tốt nhất, hãy nêu cách để ước lượng được kết quả tốt hơn.

### Kiểm định White có tích chéo

[Equation] **View** → **Residual Test** → **White heter. (cross terms)**

White Heteroskedasticity Test:			
F-statistic	26.64474	Probability	0.000000
Obs*R-squared	34.80976	Probability	0.000002

[?] - Viết phương trình hồi quy phụ ?

- Thực hiện kiểm định đối với kết quả này.

### 8.3 Lưu phần dư và giá trị ước lượng

Có thể thực hiện hồi quy phụ phần dư theo các biến độc lập hoặc theo ước lượng biến phụ thuộc. Để thực hiện chủ động các hồi quy phụ này, cần lưu lại các giá trị cần thiết dưới dạng các biến số mới.

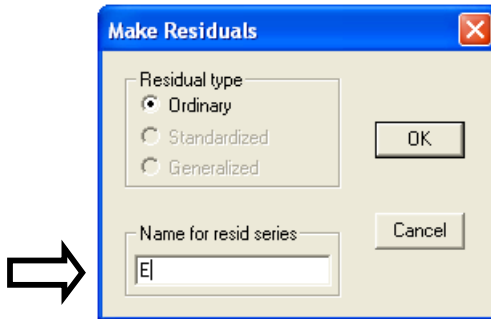
#### Lưu phần dư và giá trị ước lượng

Với bảng kết quả hồi quy mô hình (8.1), cửa sổ Equation:

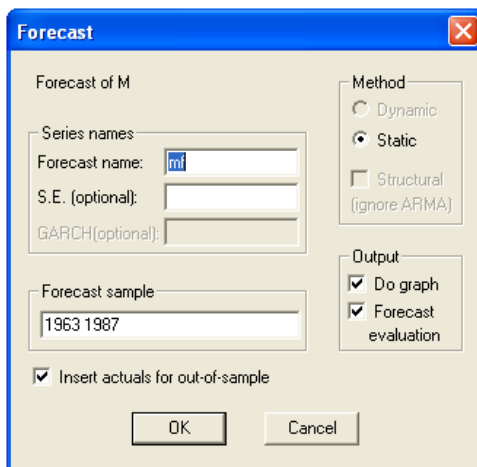
□ [Equation] **View** → **Estimation Output**

□ [Equation] **Procs** → **Make Residual Series...**

Cửa sổ Make Residuals, tên ngầm định cho phần dư là resid01, đổi thành **E** <Ok>, đóng cửa sổ Series E.



□ [Equation] → **Forecast** → Cửa sổ Forecast, trong ô Forecast name là tên của giá trị ước lượng biến phụ thuộc  $\hat{IM}$ , được ngầm định là **IMF**, sử dụng tên này.



Sau khi đặt chuỗi giá trị phần dư và giá trị ước lượng, trong cửa sổ [Workfile] xuất hiện hai biến mới là **E** và **IMF**.

## 8.4 Kiểm định bằng hồi quy phụ

Phương sai sai số được thay thế bởi  $e_i^2$  thay đổi theo các giả thiết khác nhau. Các kiểm định hồi quy phụ thực hiện kiểm định theo từng biến độc lập có trong mô hình gốc

### Hồi quy phụ với biến **EX**, **EX<sup>2</sup>**, **GDP**, **GDP<sup>2</sup>**

- LS E<sup>2</sup> C EX** ↙ Thấy PSSS thay đổi [?]
- LS E<sup>2</sup> C EX<sup>2</sup>** ↙ Thấy PSSS không đổi, nhưng nếu mức ý nghĩa là 10% thì lại có thể coi là thay đổi
- LS E<sup>2</sup> C GDP** ↙ Thấy PSSS thay đổi
- LS E<sup>2</sup> C GDP<sup>2</sup>** ↙ Thấy PSSS không đổi

Kết quả hồi quy phụ riêng biệt này không cùng nhận định với kiểm định White không tích chéo ở trên.



## Kiểm định Park theo biến EX, GDP

- LS LOG(E<sup>2</sup>) C LOG(EX)** ↙ Thấy PSSS thay đổi
  - LS LOG(E<sup>2</sup>) C LOG(GDP)** ↙ Thấy PSSS thay đổi, nhưng nếu mức ý nghĩa 10% thì có thể coi là thay đổi
- Qua các kiểm định trên, có thể thấy phương sai sai số mô hình (8.1) thay đổi, và sự thay đổi đó có thể liên quan đến biến EX hoặc GDP.

## 8.5 Ước lượng lại sai số chuẩn của các hệ số

Mô hình (8.1) có phương sai sai số thay đổi, ước lượng các hệ số vẫn không chệch, nhưng các sai số chuẩn của các ước lượng  $Se(\hat{\beta}_j)$  tính theo phương pháp OLS là không chính xác. Có thể tính lại các sai số chuẩn này theo thủ tục điều chỉnh của White

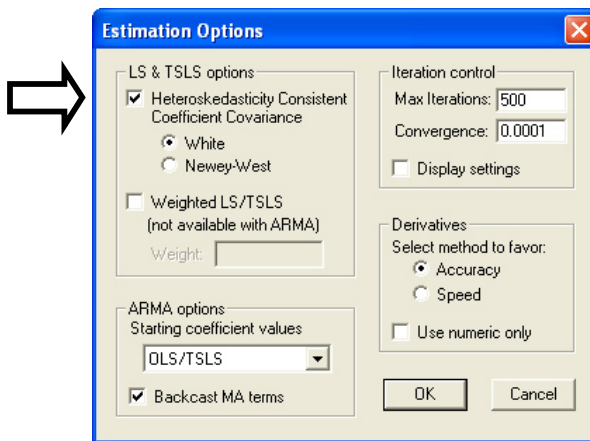
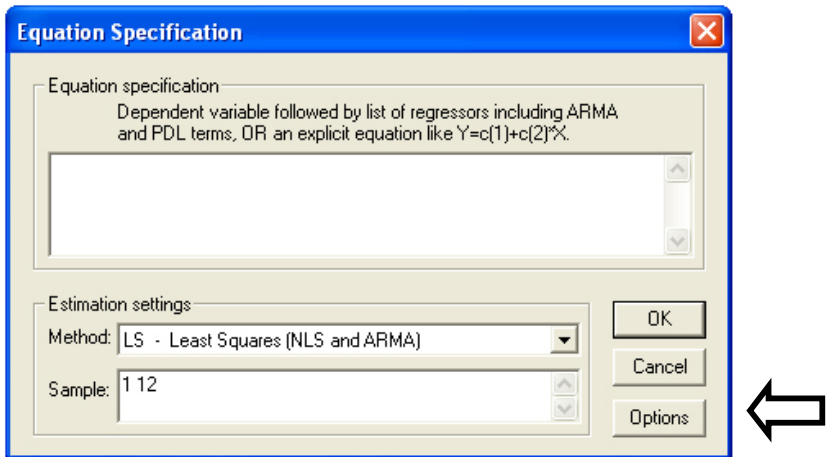
Mô hình có PSSS thay đổi

- LS IM C EX GDP** ↙

Dependent Variable: IM				
Included observations: 45 after adjusting endpoints				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	1879.134	<b>7997.118</b>	0.234976	0.8154
GDP	0.050309	<b>0.005070</b>	9.922462	0.0000
EX	0.813882	<b>0.030802</b>	26.42261	0.0000
R-squared	0.990809	Mean dependent var	277422.0	

Tại cửa sổ Equation

- [Equation] **Estimate** → Cửa sổ **Equation Specification** → **Option**
- Cửa sổ [Estimation Options] → Chọn  **Heteroskedasticity**
- Consistent Coefficient Covariance** → Chọn  **White**



Dependent Variable: IM  
White Heteroskedasticity-Consistent Standard Errors & Covariance

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	1879.134	<b>6354.506</b>	0.295717	0.7689
GDP	0.050309	<b>0.008827</b>	5.699615	0.0000
EX	0.813882	<b>0.047778</b>	17.03454	0.0000
R-squared	0.990809	Mean dependent var	277422.0	

So sánh hai kết quả, sai số chuẩn các ước lượng đã tính lại, các đại lượng khác không thay đổi.

## 8.6 Khắc phục phương sai sai số thay đổi

Mô hình (8.1)  $IM_i = \beta_1 + \beta_2 GDP_i + \beta_3 EX_i + u_i$  có phương sai sai số thay đổi, qua các hồi quy phụ, nhận thấy có thể cho rằng sự thay đổi là theo biến GDP, EX, khắc phục dựa trên các giả thiết này.

### Khắc phục theo biến GDP

Chia (8.1) cho  $GDP_i$ :

$$\frac{IM_i}{GDP_i} = \beta_1 \frac{1}{GDP_i} + \beta_2 + \beta_3 \frac{EX_i}{GDP_i} + \frac{u_i}{EX_i} \quad (8.2)$$

□ LS IM/GDP 1/GDP C EX/GDP ↴

Dependent Variable: IM/GDP				
Included observations: 45 after adjusting endpoints				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
1/GDP	116.6000	85.39516	1.365417	0.1794
C	0.032509	0.033145	0.980800	0.3323
EX/GDP	0.843903	0.043022	19.61549	0.0000
R-squared	0.901771	Mean dependent var	0.557541	
Durbin-Watson stat	0.954404	Prob(F-statistic)	0.000000	

Để xem mô hình mới (8.2) đã khắc phục được hiện tượng phương sai sai số thay đổi hay chưa, sử dụng kiểm định White  
 Kiểm định White không có tích chéo và có tích chéo

White Heteroskedasticity Test: No-cross term			
F-statistic	1.733907	Probability	0.161560
Obs*R-squared	6.649603	Probability	0.155605
White Heteroskedasticity Test: Cross term			
F-statistic	1.960378	Probability	0.106333
Obs*R-squared	9.038278	Probability	0.107547

- [?] - Mô hình (8.2) có phương sai sai số đồng đều hay thay đổi?  
 - Với mô hình này, ước lượng điểm mức thay đổi của IM khi GDP, khi EX tăng 1 đơn vị? So sánh với kết quả từ mô hình (8.1)

### Khắc phục theo biến EX

Chia (8.1) cho  $EX_i$ :

$$\frac{IM_i}{EX_i} = \beta_1 \frac{1}{EX_i} + \beta_2 \frac{GDP_i}{EX_i} + \beta_3 + \frac{u_i}{EX_i} \quad (8.3)$$

□ LS IM/EX 1/EX GDP/EX C ↵

- [?] – Kiểm định hiện tượng PSSS thay đổi trong mô hình (8.3), việc chia cho biến EX có khắc phục được hiện tượng này không?  
 - So sánh với cách chia cho GDP, cách chia cho EX có khắc phục được hiện tượng PSSS thay đổi triệt để hơn không?  
 - Với mô hình này, ước lượng điểm mức thay đổi của IM khi GDP, khi EX tăng 1 đơn vị? So sánh với kết quả từ mô hình (8.1) và (8.2)

### Đổi dạng mô hình

Đổi mô hình (8.1) sang dạng hàm logarit,

□ LS LOG(IM) C LOG(GDP) LOG(EX) ↵

Sử dụng kiểm định White cho biết có còn hiện tượng phương sai sai số thay đổi hay không?

## § 9 MÔ HÌNH VỚI SỐ LIỆU CHUỖI THỜI GIAN

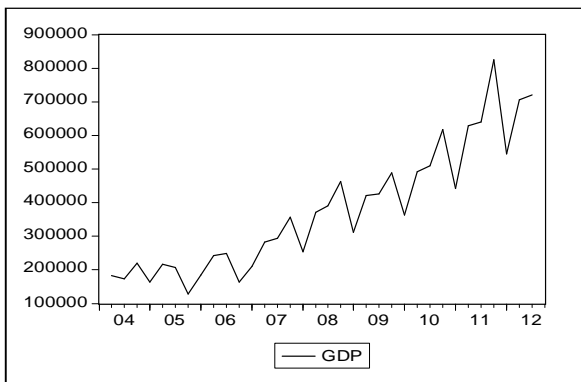
Sử dụng bộ số liệu **VNQ\_GDP** trong thư mục DATA2012.

Bộ số liệu gồm một số biến kinh tế vĩ mô của Việt Nam, từ Quý 1 năm 2004 đến Quý 3 năm 2012 (tính theo giá hiện hành, nguồn: Tổng cục thống kê). Cột frequency bắt đầu từ **2004:1** và kết thúc tại **2012:3**.

Tập trung phân tích, dự báo cho chuỗi GDP của Việt Nam.

Vẽ đồ thị của GDP theo quan sát:

Chọn **GDP** → Cửa sổ [Series] **View** → **Graph** → **Line**



### 9.1 Biến xu thế thời gian

Biến xu thế thời gian nhận giá trị lần lượt 1, 2, 3,... (hoặc 0, 1, 2,...) đến hết bộ số liệu. Biến xu thế tăng một đơn vị tương ứng với sau mỗi khoảng thời gian, nếu với bộ số liệu này là sau mỗi quý.

**GENR** **T = @TREND** ↵

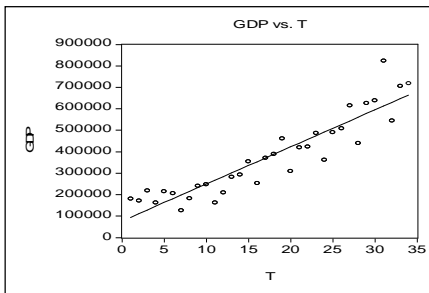
Mở cửa sổ biến T, sẽ thấy chuỗi giá trị bắt đầu từ 0.

Vẽ đồ thị của GDP theo biến T.

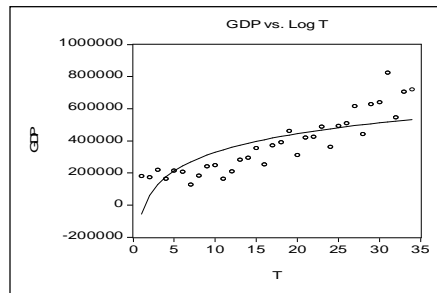
Lưu ý: Vẽ đồ thị với dạng Logarit không thực hiện được khi  $T = 0$ , do đó phải lấy mẫu từ quan sát 2004 Quý 2, tương ứng khi  $T = 1$ .

□ **SMPL 2004:2 2012:3** ↵

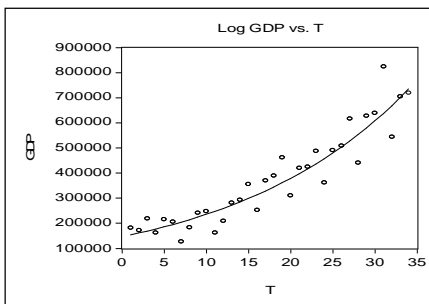
□ Chọn **T** trước, rồi **GDP** → Cửa sổ [Group] **View** → **Graph** → **Scatter** → **Scatter with regression**



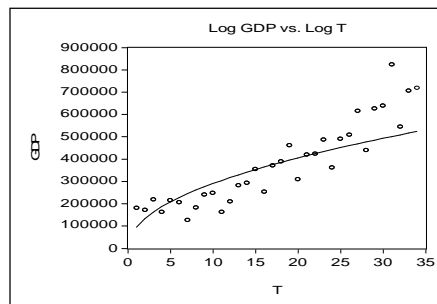
GDP theo T



GDP theo  $\ln T$  ( $T > 0$ )



$\ln$ GDP theo T



$\ln$ GDP theo  $\ln T$  ( $T > 0$ )

[?] - Dạng nào là phù hợp với xu thế của GDP? Với các dạng đó có phải loại trừ trường hợp  $T = 0$  hay không?

Lấy lại mẫu từ quan sát đầu tiên

□ **SMPL 2004:1 2012:3** ↵

## 9.2 Hồi quy và dự báo

Ước lượng mô hình với số liệu từ 2004 Q1 đến 2012 Q3, sử dụng kết quả đó để dự báo GDP cho Quý 4 2012, Quý 1 2013. Biến xu thế bắt đầu từ 0, thời gian 2012Q4 có  $T = 35$ ; thời gian 2013Q1 có  $T = 36$ .

Giá trị dự báo (forecast) kí hiệu là **GDPF**.

Thay vì dùng biến  $T$  sau khi phải đặt, có thể dùng trực tiếp biến dạng @TREND. (Nhắc lại, lấy mẫu từ 2004:1 đến 2012:3)

Số liệu thời gian, dùng chỉ số  $t$  thay cho  $i$ .

### Mô hình tuyến tính – tuyến tính

$$GDP_t = \beta_1 + \beta_2 t + u_t \quad (9.1)$$

Công thức dự báo:  $GDPF_t = \hat{\beta}_1 + \hat{\beta}_2 t$

**LS GDP C @TREND** ↴

Dependent Variable: GDP				
Sample: 2004:1 2012:3		Included observations: 35		
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	83290.55	24259.66	3.433295	0.0016
@TREND	16989.49	1226.864	13.84790	0.0000
R-squared	0.853180	Mean dependent var	372111.8	
Durbin-Watson stat	1.961474	Prob(F-statistic)	0.000000	

[?] - GDP có xu thế tăng dần không? Sau mỗi quý trung bình GDP tăng bao nhiêu?

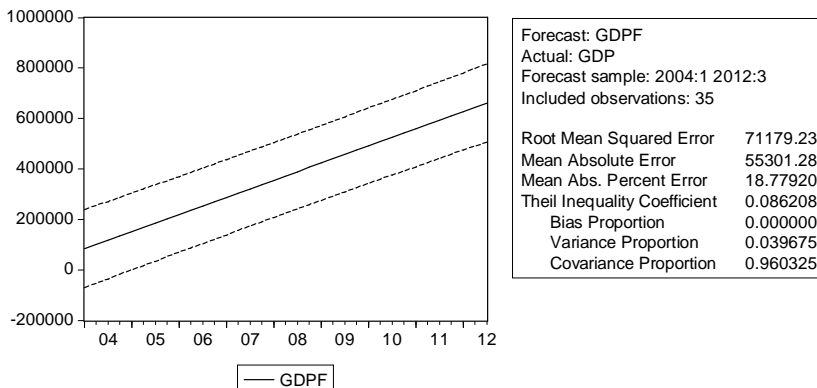
Dự báo cho GDP

2012Q4:  $t = 35 \Rightarrow GDPF_{35} = 83290,55 + 16989,49 \times 35 = 677922,7$

2013Q1:  $t = 36 \Rightarrow GDPF_{36} = 694912,2$

Đánh giá chất lượng của dự báo:

Cửa sổ [Equation] → **Forecast** → OK



Các đại lượng đánh giá chất lượng dự báo của mô hình là  
 RMSE = 71179,23; MAE = 55301,28; MAPE = 18,7792%

### Mô hình logarit – tuyến tính

$$\ln GDP_t = \beta_1 + \beta_2 t + u_t \quad (9.2)$$

Công thức dự báo:  $GDPF_t = e^{\hat{\beta}_1 + \hat{\beta}_2 t}$

□ **LS LOG(GDP) C @TREND** ↵

Dependent Variable: LOG(GDP)		Included observations: 35		
Sample: 2004:1 2012:3				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	11.88810	0.059552	199.6270	0.0000
@TREND	0.047718	0.003012	15.84459	0.0000
R-squared	0.883824	Mean dependent var	12.69931	
Durbin-Watson stat	2.180578	Prob(F-statistic)	0.000000	

[?] - Giải thích ý nghĩa hệ số góc?

Dự báo cho GDP

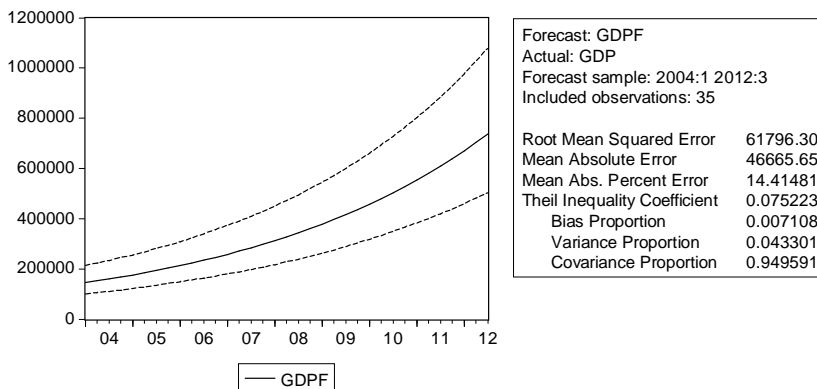
$$2012Q4: t = 35 \Rightarrow GDPF_{35} = e^{11,8881+0,047718 \times 35} = 773151,3.$$

$$2013Q1: t = 36 \Rightarrow GDPF_{36} = 810938,9.$$

Kết quả này khác biệt khá lớn so với dự báo theo mô hình tuyến tính.



Đánh giá chất lượng dự báo: Lưu ý: đánh giá cho dự báo chuỗi GDP, do đó phải chọn  $\odot$  GDP chứ không chọn LOG(GDP)



RMSE, MAE, MAPE của mô hình này đều nhỏ hơn mô hình tuyến tính, dự báo bằng kết quả này tốt hơn.

### 9.3 Dự báo tự động

Eviews có thể dự báo tự động chuỗi số theo xu thế thời gian. Ví dụ để dự báo cho GDP đến hết năm 2013 theo mô hình trong mục 9.2, cần theo các thao tác sau

#### Bước 1. Mở rộng khoảng số liệu

Cửa sổ [Workfile] → **Procs** → **Change workfile range** → Cửa sổ [Changing workfile range] → End date: **2013:4** <OK>

Trong cửa sổ Workfile có Range: 2004:1 2013:4

#### Bước 2. Hồi quy mô hình tuyến tính (9.1) và dự báo

**LS GDP C @TREND** ↵

Cửa sổ [Equation] → **Forecast** → Cửa sổ [Forecast] → Forecast name: **GDPF** → Forecast sample: 2004:1 **2013:4**

### Bước 3. Xem giá trị dự báo

□ Cửa sổ [Workfile] Chọn biến **GDPI** → Mở biến với cửa sổ [Series]  
 Các giá trị dự báo ứng với các quan sát mới mở rộng, hai giá trị ứng với 2012Q4 và 2013Q1 giống như đã tính trong mục 9.2.

Thực hiện tương tự, thực hiện dự báo bằng mô hình (9.2) lnGDP phụ thuộc xu thế, có kết quả dự báo cho GDP. Kết quả này có chênh lệch so với tính thủ công ở mục 9.2 do có sự làm tròn số.

So sánh hai kết quả

	GDP theo T	LOG(GDP) theo T
2012:4	677922.6	773158.7
2013:1	694912.1	810947.0
2013:2	711901.6	850582.3
2013:3	728891.1	892154.7
2013:4	745880.5	935758.9

## 9.4 Phân tích mùa vụ

Có bốn mùa ứng với bốn quý, do đó có thể đặt 4 biến giả, nhưng trong mô hình chỉ sử dụng tối đa 3 biến giả. Đặt các biến giả mùa vụ:

□ **GENR S1 = @SEAS(1)**

□ **GENR S2 = @SEAS(2)**

□ **GENR S3 = @SEAS(3)**

□ **GENR S4 = @SEAS(4)**

Mở 4 biến S1, S2, S3, S4 trong cùng cửa sổ Group, để thấy giá trị của các biến bằng 1 lần lượt theo các quý.

Trong ví dụ này lấy Quý 1 làm gốc, so sánh các quý khác với Quý 1.

### Mô hình chỉ có yếu tố mùa vụ

$$GDP_t = \beta_1 + \beta_2 S1_t + \beta_3 S3_t + \beta_4 S4_t + u_t \quad (9.3)$$

□ **LS GDP C S2 S3 S4 ↵**

(Có thể dùng trực tiếp @SEAS(2) @SEAS(3) @SEAS(4) )

Dependent Variable: GDP				
Sample: 2004:1 2012:3			Included observations: 35	
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	290237.9	63520.47	4.569203	0.0001
S2	103356.3	89831.50	1.150558	0.2587
S3	110635.0	89831.50	1.231584	0.2274
S4	117458.2	92596.19	1.268500	0.2141
R-squared	0.067941	Mean dependent var	372111.8	
Durbin-Watson stat	0.161748	Prob(F-statistic)	0.528872	

[?] - Theo các kiểm định T, các hệ số góc có ý nghĩa không?

- Hàm hồi quy có phù hợp không? Có thể nói các kiểm định T không ý nghĩa là do có đa cộng tuyến hay không?

- Theo kết quả này, GDP có thay đổi theo các quý không?

### Mô hình yếu tố mùa vụ và xu thế

$$GDP_t = \beta_1 + \beta_2 S2_t + \beta_3 S3_t + \beta_4 S4_t + \beta_5 t + u_t \quad (9.4)$$

□ **LS GDP C S2 S3 S4 @TREND ↵**

Dependent Variable: GDP				
Sample: 2004:1 2012:3			Included observations: 35	
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	21513.77	27491.23	0.782568	0.4400
S2	86561.08	30298.43	2.856949	0.0077
S3	77044.49	30355.88	2.538042	0.0166
S4	100663.0	31229.74	3.223305	0.0030
@TREND	16795.26	1077.742	15.58374	0.0000
R-squared	0.897521	Mean dependent var	372111.8	
Durbin-Watson stat	1.360504	Prob(F-statistic)	0.000000	

[?] - Theo các kiểm định T, các hệ số góc có ý nghĩa không?

- Theo kết quả này, GDP có khác nhau giữa các quý không? Nhìn chung, GDP quý nào lớn nhất, quý nào thấp nhất

### Dự báo với mô hình có mùa vụ và xu thế

Dự báo cho bốn quý năm 2013, ứng với  $t = 36, 37, 38, 39$ , kết hợp với bốn hệ số mùa vụ.

Công thức dự báo:  $GDPF_t = \hat{\beta}_1 + \hat{\beta}$  (mùa tương ứng)  $+ \hat{\beta}_5 t$

Quý	t	$GDPF_t =$	Kết quả
1	36	21513.77 +16795.26*36 =	626143
2	37	21513.77 + 86561.08 +16795.26*37 =	729499
3	38	21513.77 + 77044.49 +16795.26*38 =	736778
4	39	21513.77 + 100663 +16795.26*39 =	777191

[?] - Đánh giá chất lượng dự báo, các đại lượng RMSE, MAE, MAPE bằng bao nhiêu?

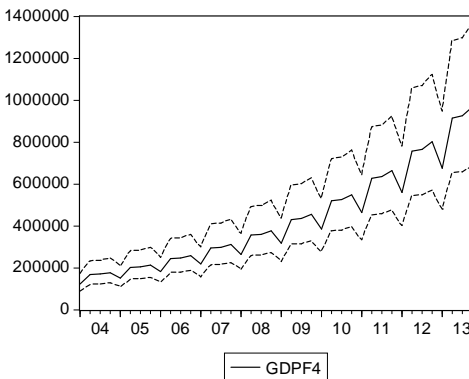
Có thể sử dụng mô hình dạng **lnGDP** phụ thuộc **T** và các biến mùa vụ **S2, S3, S4** để dự báo tự động mô hình:

$$\ln GDP_t = \beta_1 + \beta_2 S2_t + \beta_3 S3_t + \beta_4 S4_t + \beta_5 t + u_t \quad (9.5)$$

Hay  $GDP_t = e^{\beta_1 + \beta_2 S2_t + \beta_3 S3_t + \beta_4 S4_t + \beta_5 t + u_t}$

Kết quả dự báo tự động như sau:

2013:1	675912.5
2013:2	914502.9
2013:3	925073.3
2013:4	967256.5



Forecast: GDPF4	
Actual: GDP	
Forecast sample: 2004:1 2013:4	
Included observations: 35	
Root Mean Squared Error	44224.19
Mean Absolute Error	28391.88
Mean Abs. Percent Error	9.144191
Theil Inequality Coefficient	0.053595
Bias Proportion	0.006183
Variance Proportion	0.032673
Covariance Proportion	0.961143

## 9.4 Kiểm định tính dừng

Một chuỗi thời gian được gọi là dừng (*stationary*) nếu nó có trung bình không đổi, phương sai không đổi, mức độ tương quan với quá khứ là không đổi khi khoảng cách cố định. Trên đồ thị nhận thấy biến GDP có trung bình tăng theo thời gian, phương sai thay đổi theo thời gian, do đó bằng trực giác có thể cho rằng nó không dừng. Bên cạnh đó có thể thực hiện kiểm định Dickey-Fuller, còn gọi là kiểm định nghiệm đơn vị (*unit root test*).

### Kiểm định DF có hệ số chặn

Kiểm định DF cho biến GDP, hồi quy phụ có hệ số chặn là:

$$\Delta GDP_t = \alpha_0 + \alpha_1 GDP_{t-1} + v_t \quad (9.6)$$

Với  $\Delta GDP_t = GDP_t - GDP_{t-1}$  là sai phân bậc nhất của GDP, trong Eviews kí hiệu là D(GDP).

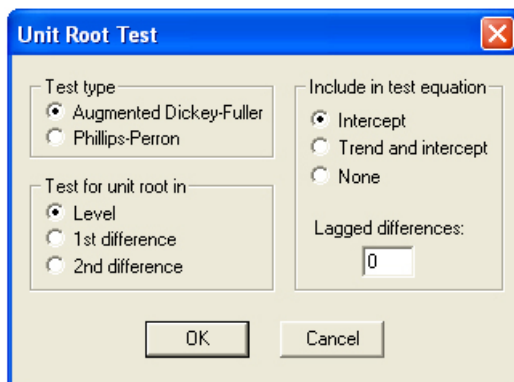
Kiểm định cặp giả thuyết:

$H_0: \alpha_1 = 0$  : Chuỗi là Không dừng

$H_1: \alpha_1 \neq 0$  : Chuỗi là dừng

$\tau_{qs} = \frac{\hat{\alpha}_1}{Se(\hat{\alpha}_1)}$ , nếu  $|\tau_{qs}| > |\tau_\alpha|$  thì bác bỏ  $H_0$ , chuỗi là dừng.

- [Workfile] Mở biến **GDP** thành cửa sổ riêng
- [Series: GDP] **View** → **Unit Root Test...**
- [Unit Root Test] Test type:  **Augmented Dickey-Fuller** → Test for unit root in:  **Level** → Include in test equation:  **Intercept** → Lag differences: **0** (lưu ý máy ngầm định là 1 nhưng gõ lại là 0)



ADF Test Statistic	-1.325049	1% Critical Value*	-3.6353
		5% Critical Value	-2.9499
		10% Critical Value	-2.6133

\*MacKinnon critical values for rejection of hypothesis of a unit root.

Augmented Dickey-Fuller Test Equation

Dependent Variable: D(GDP) Sample(adjusted): 2004:2 2012:3

Included observations: 34 after adjusting endpoints

Variable	Coefficient	Std. Error	t-Statistic	Prob.
GDP(-1)	-0.129245	0.097540	-1.325049	0.1945
C	63921.36	39356.54	1.624161	0.1142

Theo kết quả hồi quy phụ này  $\tau_{qs} = -1.325$ ;

Các giá trị tới hạn:  $\tau_{0.01} = -3.6353$ ;  $\tau_{0.05} = -2.9499$ ;  $\tau_{0.1} = -2.6133$

Vậy  $|\tau_{qs}| < |\tau_{\alpha}|$  với cả ba mức ý nghĩa 10%, 5%, 1, chưa bác bỏ  $H_0$ , chuỗi là không dừng.

Trong kết quả thấy hệ số chặn không có ý nghĩa thống kê, có thể thực hiện kiểm định không có hệ số chặn (chọn None trong Include test equation), kết luận cũng không thay đổi. Thông thường hệ số chặn được cho trong hồi quy phụ để kiểm định.

### Kiểm định DF có hệ số chặn và xu thế thời gian

Kiểm định DF cho biến GDP, hồi quy phụ có hệ số chặn và xu thế thời gian là:

$$\Delta GDP_t = \alpha_0 + \alpha_1 GDP_{t-1} + \alpha_2 t + v_t \quad (9.7)$$

Kiểm định cặp giả thuyết:

$H_0: \alpha_1 = 0$  : Chuỗi Không dừng xu thế

$H_1: \alpha_1 \neq 0$  : Chuỗi là dừng xu thế

Lưu ý: Chuỗi dừng xu thế tức là dao động quanh một xu thế thời gian, trung bình của chuỗi vẫn tăng (hoặc giảm) theo thời gian. Do đó chuỗi Dừng xu thế là chuỗi Không dừng.

Khi cho yếu tố xu thế thời gian vào là ta đã ngầm định chuỗi là không dừng, chỉ xem nó có dao động ổn định quanh xu thế thời gian hay không mà thôi.

[Series: GDP] **View** → **Unit Root Test...** →

[Unit Root Test] Chọn  **Level** →  **Trend and Intercept** → Lag differences: **0**

ADF Test Statistic	-5.543229	1% Critical Value*	-4.2505
		5% Critical Value	-3.5468
		10% Critical Value	-3.2056

\*MacKinnon critical values for rejection of hypothesis of a unit root.

Augmented Dickey-Fuller Test Equation

Dependent Variable: D(GDP) Sample(adjusted): 2004:2 2012:3

Included observations: 34 after adjusting endpoints

Variable	Coefficient	Std. Error	t-Statistic	Prob.
GDP(-1)	-0.997592	0.179966	-5.543229	0.0000
C	76507.13	29152.04	2.624417	0.0134
@TREND(2004:1)	17236.92	3273.962	5.264851	0.0000

Theo kết quả này  $\tau_{qs} = -5.543$ , có  $|\tau_{qs}| > |\tau_{\alpha}|$  với mức 1%, bác bỏ  $H_0$ , chuỗi là dừng xu thế, hay GDP có thể coi là dao động ổn định quanh xu thế. Xu thế mang dấu dương và có ý nghĩa thống kê.

### Kiểm định tính dừng của sai phân

Một chuỗi không dừng, có thể kiểm định tính dừng của sai phân chuỗi đó. Sai phân là mức thay đổi tuyệt đối của chuỗi ở kì sau so với kì trước.

Sai phân bậc nhất của GDP:  $\Delta GDP_t = GDP_t - GDP_{t-1}$

Để kiểm định sai phân bậc nhất của GDP có dừng không, xét sai phân là chuỗi mới, hồi quy phụ kiểm định có hệ số chặn là:

$$\Delta(\Delta GDP)_t = \alpha_0 + \alpha_1 \Delta GDP_{t-1} + v_t \quad (9.8)$$

Kiểm định cặp giả thuyết:

$H_0: \alpha_1 = 0$  : Chuỗi sai phân không dừng

$H_1: \alpha_1 \neq 0$  : Chuỗi sai phân là dừng

Lưu ý: Kiểm định tính dừng của chuỗi sai phân chỉ thực hiện khi chuỗi gốc là không dừng. Nếu chuỗi gốc đã dừng thì không thực hiện kiểm định này. Kiểm định này cho biết chuỗi đã là không dừng thì sự thay đổi của nó có ổn định không.

[Series: GDP] **View** → **Unit Root Test...** →

[Unit Root Test] Chọn  **1st difference** →  **Intercept** → Lag differences: **0**

ADF Test Statistic	-11.67530	1% Critical Value*	-3.6422
		5% Critical Value	-2.9527
		10% Critical Value	-2.6148

\*MacKinnon critical values for rejection of hypothesis of a unit root.

Augmented Dickey-Fuller Test Equation

Dependent Variable: D(GDP,2) Sample(adjusted): 2004:3 2012:3

Included observations: 33 after adjusting endpoints



Variable	Coefficient	Std. Error	t-Statistic	Prob.
D(GDP(-1))	-1.628305	0.139466	-11.67530	0.0000
C	27153.74	14518.96	1.870226	0.0709

Có thể thấy chuỗi sai phân là dừng.

Vậy chuỗi GDP không dừng nhưng sự thay đổi của nó được coi là dừng. Trong các phân tích, hồi quy nên sử dụng chuỗi sai phân của GDP.

Trường hợp với chuỗi sai phân của GDP nhưng vẫn cho biến xu thế thời gian:

ADF Test Statistic	-11.64497	1% Critical Value*	-4.2605
		5% Critical Value	-3.5514
		10% Critical Value	-3.2081

\*MacKinnon critical values for rejection of hypothesis of a unit root.

Dependent Variable: D(GDP,2)

Included observations: 33 after adjusting endpoints

Variable	Coefficient	Std. Error	t-Statistic	Prob.
D(GDP(-1))	-1.635665	0.140461	-11.64497	0.0000
C	4716.644	30784.63	0.153214	0.8793
@TREND(2004:1)	1253.565	1514.413	0.827757	0.4143

Ta thấy dù kết luận chuỗi sai phân là dừng, nhưng biến xu thế không có ý nghĩa thống kê, P-value là 0.4143 nên kiểm định này không nên sử dụng.

[?] - Với chuỗi CO, kiểm định tính dừng có hệ số chặn? Với mức ý nghĩa bao nhiêu thì chuỗi được coi là dừng?

- Kiểm định tính dừng xu thế của CO, tính dừng của sai phân chuỗi CO?

[?] - Với chuỗi GI, IND kiểm định tính dừng có hệ số chặn? ?

- Kiểm định tính dừng xu thế, tính dừng của sai phân của hai chuỗi GI, IND?

## § 10 TỰ TƯƠNG QUAN VÀ MÔ HÌNH CÓ BIẾN TRỄ

Sử dụng bộ số liệu **VNQ\_GDP** trong thư mục DATA2012.

Xét mô hình Giá trị sản xuất ngành Xây dựng (CO) phụ thuộc Đầu tư xây dựng cơ bản (GI), Giá trị sản xuất công nghiệp (IND)

### 10.1 Hiện tượng tự tương quan

Hồi quy mô hình  $CO_t = \beta_1 + \beta_2 GI_t + \beta_3 IND_t + u_t$  (10.1)

□ LS CO C GI IND ↴

Dependent Variable: CO				
Sample: 2004:1 2012:3				
Included observations: 35				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-2257.031	2807.746	-0.803859	0.4274
GI	0.445348	0.172523	2.581390	0.0146
IND	0.106931	0.039731	2.691362	0.0112
R-squared	0.770290	Mean dependent var	23883.86	
Durbin-Watson stat	2.848833	Prob(F-statistic)	0.000000	

Theo kết quả hồi quy này, giá trị  $d$  của kiểm định Durbin-Watson bằng 2,848833; với giá trị  $d_L = 1.343$  và  $d_U = 1.584$

[?] - Kiểm định DW cho biết điều gì về mô hình?

- Hãy nêu 1 cách để khắc phục hiện tượng tự tương quan bậc 1 dựa trên thống kê DW?

### 10.2 Kiểm định Breusch-Godfrey

Kiểm định tự tương quan bậc 1 của mô hình (10.1) bằng kiểm định Breusch-Godfrey, hồi quy phụ là:

$$e_t = (\alpha_1 + \alpha_2 GI_t + \alpha_3 IND_t) + \rho_1 e_{t-1} + v_t \quad (10.1a)$$

$$e_t = (\alpha_1 + \alpha_2 GI_t + \alpha_3 IND_t) + v_t \quad (10.1b)$$

$H_0: \rho_1 = 0$ : Mô hình không có tự tương quan bậc 1

$H_1: \rho_1 \neq 0$ : Mô hình có tự tương quan bậc 1

$$F_{qs} = \frac{R^2_{(10.1a)} - R^2_{(10.1b)}}{1 - R^2_{(10.1a)}} \times \frac{n_{(10.1a)} - k_{(10.1a)}}{1} \quad \text{và} \quad \chi^2_{qs} = n_{(10.1a)} R^2_{(10.1a)}$$

Mô hình (10.1a) có trễ bậc 1, số quan sát bớt đi 1. Tuy nhiên chương trình Eviews tự động thay giá trị thiếu đó bằng 0, do đó số quan sát vẫn được giữ nguyên.

□ [Equation] **View** → **Residual Tests** → **Residual Correlation LM Test...**

⇒ Cửa sổ [Lag specification] xác định bậc của tự tương quan, ngầm định là 2, cần kiểm định tự tương quan bậc 1

□ [Lag specification] Lags to include: **1** <Ok>



Breusch-Godfrey Serial Correlation LM Test:				
F-statistic	8.078363	Probability	0.007854	
Obs*R-squared	7.235275	Probability	0.007148	
Test Equation:				
Dependent Variable: RESID				
Presample missing value lagged residuals set to zero.				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	1329.848	2583.488	0.514749	0.6104
GI	-0.101550	0.160154	-0.634074	0.5307
IND	0.012601	0.036226	0.347840	0.7303
RESID(-1)	-0.475989	0.167469	-2.842246	0.0079

[?] - Bậc tự do của F-statistic và Obs\*R-squared là bao nhiêu?

- Kiểm định và kết luận về hiện tượng tự tương quan bậc 1?
- Có thể lấy ước lượng hệ số tự tương quan bậc 1 bằng bao nhiêu thông qua kiểm định này?

Tương tự, thực hiện kiểm định tự tương quan đến bậc 2.

- [?] - Với kiểm định tự tương quan bậc 2, hồi quy phụ thể nào?
- Thực hiện kiểm định cụ thể, có tự tương quan bậc 1, bậc 2 hay không?

### 10.3 Kiểm định bằng hồi quy phụ

Kiểm định tự tương quan dựa trên việc xem xét phần dư  $e_t$  có phụ thuộc vào các trễ của nó không.

#### Kiểm định tự tương quan bậc 1

Không có hệ số chặn 
$$e_t = \alpha_1 e_{t-1} + v_t \quad (10.1c)$$

Có hệ số chặn 
$$e_t = \alpha_0 + \alpha_1 e_{t-1} + v_t \quad (10.1d)$$

[Equation] **Procs** → **Make Residual Series...**

[Make Residuals] → Name of resid series: **E**

Hồi quy không có hệ số chặn

**LS E E(-1)** ↵

Dependent Variable: E				
Included observations: 34 after adjusting endpoints				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
E(-1)	-0.442081	0.157660	-2.804010	0.0084
R-squared	0.192361	Mean dependent var	55.83272	

Hồi quy phụ này không có hệ số chặn nên không có kiểm định F

- [?] - Phần dư có phụ thuộc trễ của nó không?
- Có thể lấy hệ số tự tương quan bậc 1 bằng bao nhiêu?

Hồi quy có hệ số chặn

□ **LS E C E(-1)** ↵

[?] - Hệ số chặn có ý nghĩa thống kê không?

- Kết luận gì về tự tương quan bậc 1 của mô hình gốc?

- Các ước lượng hệ số và sai số chuẩn của nó trong mô hình (10.1) có phải là ước lượng không chệch, ước lượng hiệu quả?

**10.4 Ước lượng lại sai số chuẩn**

Mô hình  $CO_t = \beta_1 + \beta_2 GI_t + \beta_3 IND_t + u_t$  có tự tương quan bậc 1 âm, tính lại các sai số chuẩn của hệ số  $Se(\hat{\beta}_j)$  bằng thủ tục Newey-West.

□ **LS CO C GI IND** ↵

□ [Equation] → **Estimate** → Cửa sổ [Equation Specification]

□ [Equation Specification] → **Option** → Cửa sổ [Estimation Option]

- [Estimation Option] → Chọn  **Heteroskedasticity**
- Consistent Coefficient Covariance** → Chọn  **Newey-West**
- <OK>

Dependent Variable: CO				
Sample: 2004:1 2012:3		Included observations: 35		
Newey-West HAC Standard Errors & Covariance (lag truncation=3)				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-2257.031	<b>1408.430</b>	-1.602516	0.1189
GI	0.445348	<b>0.154274</b>	2.886732	0.0069
IND	0.106931	<b>0.042730</b>	2.502485	0.0176
R-squared	0.770290	Mean dependent var	23883.86	
Durbin-Watson stat	2.848833	Prob(F-statistic)	0.000000	

So với bảng kết quả ở mục 10.1, các sai số chuẩn đã được tính toán lại.

## 10.5 Khắc phục tự tương quan

### Sử dụng phương trình sai phân tổng quát

Phương trình sai phân tổng quát của mô hình (10.1) là

$$CO_t - \rho CO_{t-1} = \beta_1(1 - \rho) + \beta_2(GI_t - \rho GI_{t-1}) + \beta_3(IND_t - \rho IND_{t-1}) + u_t - \rho u_{t-1} \quad (10.2)$$

Qua thống kê DW và hồi quy phụ (10.1c), (10.1d), ước lượng hệ số tương quan bậc nhất  $\hat{\rho} \cong -0,44$ , thay vào phương trình sai phân tổng quát được:

$$CO_t + 0,44CO_{t-1} = \beta_1(1 + 0,44) + \beta_2(GI_t + 0,44GI_{t-1}) + \beta_3(IND_t + 0,44IND_{t-1}) + v_t$$

□LS (CO+0.44\*CO(-1)) C (GI+0.44\*GI(-1)) (IND+0.44\*IND(-1))

Dependent Variable: CO+0.44*CO(-1)				
Included observations: 34 after adjusting endpoints				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-1620.712	2693.379	-0.601739	0.5517
GI+0.44*GI(-1)	0.418240	0.147785	2.830067	0.0081
IND+0.44*IND(-1)	0.104762	0.032270	3.246449	0.0028
R-squared	0.876555	Mean dependent var		34663.45
Durbin-Watson stat	2.084425	Prob(F-statistic)		0.000000

Mô hình (10.2) có ước lượng hệ số chặn là  $-1620.7$  do đó ước lượng hệ số chặn của mô hình (10.1) sẽ là  $\hat{\beta}_1 = -1620.7/1.44 = -1125.5$ , ước lượng hệ số góc của mô hình gốc  $\hat{\beta}_2 = 0.41824$ ;  $\hat{\beta}_3 = 0.104762$ .

[?] - Dùng kiểm định Durbin-Watson chứng tỏ mô hình (10.2) không còn tự tương quan.

- Tại sao ước lượng hệ số chặn  $\beta_1$  của mô hình (10.1) lại phải lấy ước lượng hệ số chặn của mô hình (10.2) chia cho 1.44?

- Khi đó sai số chuẩn  $Se(\hat{\beta}_1)$  sẽ bằng bao nhiêu?

Dùng kiểm định Breusch-Godfrey kiểm định mô hình (10.2), được:

Breusch-Godfrey Serial Correlation LM Test:			
F-statistic	0.109745	Probability	0.742737
Obs*R-squared	0.123925	Probability	0.724816

[?] - Cho biết mô hình (10.2) có tự tương quan bậc 1 hay không?

### Sử dụng phương pháp ước lượng $\rho$ nhiều bước – phương pháp lặp AR(1)

Kí hiệu tự tương quan hoặc tự hồi quy là AR (Autoregression), để thực hiện khắc phục tự tương quan bậc 1, thêm AR(1) vào sau phương trình

□ **LS CO C GI IND AR(1)** ↴

Dependent Variable: CO				
Included observations: 34 after adjusting endpoints				
Convergence achieved after 8 iterations				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-1115.426	1886.725	-0.591197	0.5588
GI	0.417874	0.152362	2.742644	0.0102
IND	0.104767	0.033086	3.166531	0.0035
AR(1)	-0.455065	0.166619	-2.731175	0.0105
R-squared	0.807906	Mean dependent var	24427.62	
Durbin-Watson stat	2.057495	Prob(F-statistic)	0.000000	
Inverted AR Roots	-.46			
Breusch-Godfrey Serial Correlation LM Test:				
F-statistic	0.316799	Probability	0.577864	
Obs*R-squared	0.367405	Probability	0.544422	

Phương pháp dừng lại sau 8 bước lặp, ước lượng hệ số tự tương quan bậc nhất là  $-0,455$ . Kiểm định B-G:

Breusch-Godfrey Serial Correlation LM Test: AR(1)			
F-statistic	0.316799	Probability	0.577864
Obs*R-squared	0.367405	Probability	0.544422

[?] - Bằng kiểm định DW và kiểm định BG cho biết mục đích khác phục tự tương quan đã đạt được chưa?

- Viết phương trình sai phân tổng quát ứng với bước cuối cùng của mô hình. Khi đó ước lượng hệ số chặn của phương trình sai phân tổng quát đó bằng bao nhiêu?

## 10.6 Mô hình có biến trễ

Bộ số liệu sử dụng theo quý chưa hiệu chỉnh mùa vụ, do đó việc đưa biến trễ có thể có lỗi về mô hình như đa cộng tuyến cao, dao động của các biến liên tục đổi chiều. Những thực hành sau có tính tham khảo về phương pháp, để phân tích về kinh tế cần có nhiều bước hiệu chỉnh.

### Trễ của biến độc lập

$$CO_t = \beta_1 + \beta_2 GI_t + \beta_3 IND_t + \beta_4 GI_{t-1} + \beta_5 IND_{t-1} + u_t \quad (10.3)$$

#### □ LS CO C GI IND GI(-1) IND(-1)

Dependent Variable: CO				
Sample(adjusted): 2004:2 2012:3				
Included observations: 34 after adjusting endpoints				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-1163.076	2856.413	-0.407181	0.6869
GI	0.465494	0.167217	2.783766	0.0094
IND	0.211045	0.061237	3.446372	0.0018
GI(-1)	0.091127	0.168552	0.540649	0.5929
IND(-1)	-0.142473	0.059598	-2.390577	0.0235
R-squared	0.799888	Mean dependent var	24427.62	
Durbin-Watson stat	2.783134	Prob(F-statistic)	0.000000	

[?] - Giải thích ý nghĩa ước lượng hệ số của hai biến trễ

- Theo kết quả này, CO phụ thuộc biến trễ nào?

### Kiểm định B-G cho thấy có tự tương quan bậc 1

Breusch-Godfrey Serial Correlation LM Test:			
F-statistic	9.260874	Probability	0.005047
Obs*R-squared	8.450411	Probability	0.003650



Việc cho thêm các trễ bậc 2, bậc 3... gây ra đa cộng tuyến cao.

### Trễ của biến phụ thuộc (mô hình tự hồi quy)

$$CO_t = \beta_1 + \beta_2 GI_t + \beta_3 IND_t + \beta_4 CO_{t-1} + u_t \quad (10.4)$$

□ LS CO C GI IND CO(-1)

Dependent Variable: CO				
Sample(adjusted): 2004:2 2012:3				
Included observations: 34 after adjusting endpoints				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-1042.395	2749.436	-0.379130	0.7073
GI	0.374147	0.164674	2.272045	0.0304
IND	0.167408	0.043855	3.817320	0.0006
CO(-1)	-0.297965	0.115738	-2.574475	0.0152
R-squared	0.802710	Mean dependent var	24427.62	
Durbin-Watson stat	2.283338	Prob(F-statistic)	0.000000	

### Kiểm định B-G

Breusch-Godfrey Serial Correlation LM Test:			
F-statistic	3.653637	Probability	0.065875
Obs*R-squared	3.804282	Probability	0.051122

[?] - Giải thích ý nghĩa ước lượng hệ số của biến trễ

- Theo kết quả này, CO có phụ thuộc trễ của chính nó không?

- Mô hình có tự tương quan bậc 1 không? Nếu mức ý nghĩa là 10% thì kết luận thế nào?

## 10.7 Mô hình với các chuỗi sai phân

Thực hiện kiểm định tính dừng với các chuỗi CO, GI, IND được kết quả chuỗi CO dừng ở mức ý nghĩa 5% nhưng không dừng với mức 1%, sai phân của nó dừng ở mọi mức ý nghĩa. Chuỗi GI, IND không

dùng nhưng sai phân dùng ở mọi mức ý nghĩa. Do đó hồi quy mô hình sai phân các chuỗi theo nhau.

Mô hình:

$$\Delta CO_t = \beta_1 + \beta_2 \Delta GI_t + \beta_3 \Delta IND_t + u_t \quad (10.5)$$

□ **LS D(CO) C D(GI) D(IND)**

Dependent Variable: D(CO)				
Sample(adjusted): 2004:2 2012:3				
Included observations: 34 after adjusting endpoints				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-1893.533	1819.165	-1.040880	0.3060
D(GI)	0.585601	0.171817	3.408277	0.0018
D(IND)	0.373346	0.079726	4.682884	0.0001
R-squared	0.590589	Mean dependent var	1164.353	
Durbin-Watson stat	3.493292	Prob(F-statistic)	0.000001	

[?] - Giải thích ý nghĩa ước lượng các hệ số gốc

- Kiểm định hiện tượng tự tương quan bằng DW và B-G?

Mô hình có hiện tượng tự tương quan, có thể thực hiện các thay đổi với mô hình để ước lượng chính xác và phù hợp hơn:

- Thực hiện ước lượng lại các sai số chuẩn trong điều kiện có tự tương quan?
- Thêm vào mô hình biến trễ của biến độc lập, trễ của biến phụ thuộc?
- Khắc phục theo phương pháp ước lượng  $\rho$  nhiều bước.

## BÀI TẬP TỔNG HỢP

**Bài tập 1.** Sử dụng bộ số liệu **HANOI\_HH\_2009** với các biến số về chi tiêu-thu nhập của một số hộ gia đình tại Hà Nội năm 2009 (nguồn: *Điều tra VHLSS*).

- (a) Xem ý nghĩa các biến, biến nào là biến định tính, biến nào là định lượng. Biến định lượng nào phân phối chuẩn?
- (b) Hệ số tương quan giữa CONSUM (chi tiêu) và INCOME (Thu nhập) bằng bao nhiêu?

Hồi Chi tiêu theo Thu nhập có hệ số chặn (mô hình 1).

- (c) Khi đó ước lượng điểm tiêu dùng tự định và khuynh hướng tiêu dùng cận biên bằng bao nhiêu? Hệ số xác định bằng bao nhiêu, hàm hồi quy có phù hợp không?
- (d) Phần dư và giá trị ước lượng với quan sát đầu tiên bằng bao nhiêu?
- (e) Kiểm định giả thuyết khuynh hướng tiêu dùng cận biên bằng 0.5?
- (f) Kiểm định việc thêm cả hai biến SIZE và AGE vào mô hình?
- (g) Mô hình có phương sai sai số thay đổi hay không đổi? Nếu có hãy ước lượng lại sai số chuẩn của các hệ số.
- (h) Với kiểm định Ramsey thêm một biến, hàm hồi quy được xác định có đúng không?

Đổi mô hình sang dạng hàm Cobb-Douglas (mô hình 2)

- (i) Ước lượng điểm mức thay đổi tương đối của chi tiêu khi thu nhập tăng tương đối 1%
- (j) Mô hình 2 còn các khuyết tật mà mô hình 1 mắc phải hay không?

Với biến giả URBAN = 1 với khu vực thành thị, = 0 với khu vực nông thôn, thêm biến giả vào mô hình 1 theo các yêu cầu sau:

- (k) Hồi quy mô hình chi tiêu phụ thuộc thu nhập, biến giả tác động đến hệ số chặn. Khi đó hệ số chặn có thực sự khác nhau giữa hai khu vực không? Nếu có thì tiêu dùng tự định khu vực nào cao hơn, cao hơn bao nhiêu?
- (l) Hồi quy mô hình chi tiêu phụ thuộc thu nhập có biến giả tác động đến hệ số góc. Khi đó hệ số góc có thực sự khác nhau không? Khuynh hướng tiêu dùng cận biên khu vực nào cao hơn, cao hơn bao nhiêu?
- (m) Hồi quy mô hình chi tiêu phụ thuộc thu nhập có biến giả tác động đến cả hai hệ số. Khi đó các biến mới có ý nghĩa thống kê không? So sánh với trường hợp câu (k), (l), có thể giải thích thế nào về hiện tượng đó.
- (n) Kiểm định về các hiện tượng phương sai sai số thay đổi, dạng hàm trong mô hình các câu (k), (l), (m).
- (o) Hồi quy Chi tiêu bình quân đầu người theo Thu nhập bình quân đầu người, có hệ số chặn. Khi đó mô hình giải thích được bao nhiêu % sự biến động của biến phụ thuộc, mô hình có khuyết tật nào không?

**Bài tập 2.** Sử dụng bộ số liệu **HANOI\_SALES\_2009** với các biến số của một số doanh nghiệp mua bán xe máy ô tô trên tại Hà Nội năm 2009 (nguồn: *Điều tra doanh nghiệp*). Các doanh nghiệp này thuộc 4 loại: Nhà nước trung ương, Nhà nước địa phương, Tư nhân, Có vốn đầu tư nước ngoài, được mã hóa trong biến OWNERSHIP.

- (a) Xem ý nghĩa các biến, biến nào thể hiện kết quả hoạt động của doanh nghiệp, biến đó có phân phối chuẩn không?
- (b) Biến GO (Sản lượng) tương quan với biến nào chặt nhất, lỏng nhất? Hệ số tương quan tương ứng đó bằng bao nhiêu?

Hồi quy GO theo ASSET (được hiểu là vốn) và LABOR (lao động) có hệ số chặn (mô hình 1)

- (c) Giải thích ý nghĩa các ước lượng hệ số góc và hệ số xác định, hiệp phương sai ước lượng hai hệ số góc bằng bao nhiêu?
- (d) Kiểm định về các khuyết tật của mô hình
- (e) Khi thêm biến Tiền công (WAGE) vào mô hình, có nhận xét gì về mô hình mới, có nên dùng mô hình này không?
- (f) Hồi quy năng suất lao động theo vốn bình quân cho một lao động, có hệ số chặn, và phân tích kết quả

Hồi quy GO theo ASSET và LABOR dưới dạng hàm Cobb-Douglas

- (f) Phân tích ý nghĩa và đánh giá kết quả.
- (g) Theo ước lượng điểm, quá trình sản xuất có hiệu quả tăng, giảm, hay không đổi theo quy mô? Kiểm định giả thuyết quá trình sản xuất hiệu quả không đổi theo quy mô?

Đặt các biến giả để so sánh các loại doanh nghiệp, lấy doanh nghiệp tư nhân làm gốc. Với mô hình Cobb-Douglas, phân tích sự tác động của biến giả đến các hệ số góc.

- (h) So sánh với doanh nghiệp tư nhân, tác động của vốn đến sản lượng của doanh nghiệp nhà nước trung ương có khác biệt không?
- (i) So sánh với doanh nghiệp tư nhân, tác động của lao động đến sản lượng của doanh nghiệp có vốn nước ngoài có khác biệt không?

**Bài tập 3.** Sử dụng bộ số liệu **VNM\_EXIM** với số liệu theo tháng (nguồn: *Tổng cục thống kê*).

- (a) Đặt biến giả M6 bằng 1 nếu vào tháng 6, bằng 0 nếu vào tháng khác, và tương tự là các biến M7, M8 cho tháng 7, tháng 8.

- (b) Hồi quy EX (xuất khẩu) theo xu thế thời gian, M6, M7, M8. Xuất khẩu trung bình có khác biệt giữa tháng 6, tháng 7, tháng 8 với các tháng còn lại không? Nếu có thì cao hơn hay thấp hơn bao nhiêu?
- (c) Kiểm định tự tương quan bậc 1 của mô hình bằng kiểm định DW và kiểm định Breusch-Godfrey
- (d) Hồi quy EX theo GIC (Đầu tư từ trung ương), GIL (đầu tư địa phương) có hệ số chặn, được mô hình 1. Khi đó biến độc lập nào giải thích cho EX?
- (e) Mô hình 1 có tự tương quan bậc 1 không? Nếu có hãy ước lượng lại các sai số chuẩn của các hệ số.
- (f) Mô hình 1 còn khuyết tật nào nữa không?
- (g) Thêm biến trễ bậc 1 của GIL vào mô hình 1, biến mới thêm có ý nghĩa thống kê không? Mô hình còn tự tương quan bậc 1 không?
- (h) Thêm biến trễ bậc 1 của EX vào mô hình 1, mô hình mới có tốt hơn mô hình trước khi thêm hay không?
- (i) Hồi quy mô hình EX phụ thuộc GIL, ba biến giả M6, M7, M8 tác động đến hệ số góc. Khi đó hệ số góc của tháng nào là thực sự khác với các tháng còn lại? Kiểm định các khuyết tật của mô hình mới này.

**Bài tập 4.** Sử dụng bộ số liệu **VNY\_WB\_IMF** với số liệu theo năm (nguồn: *World Bank* và *IMF*). Với biến IM (nhập khẩu) là biến phụ thuộc, xây dựng mô hình kinh tế lượng với các biến độc lập là các biến phù hợp. Khi đó đánh giá kết quả ước lượng, nếu có các khuyết tật thì thực hiện các bước khắc phục để có được kết quả tốt nhất trong phạm vi có thể. Với kết quả đó phân tích về mối quan hệ, nêu các nhận xét kinh tế đánh giá về Nhập khẩu Việt Nam giai. Có thể thay đổi thời kì lấy mẫu từ 1992 để phù hợp với thực tế kinh tế Việt Nam.

**HẾT.**