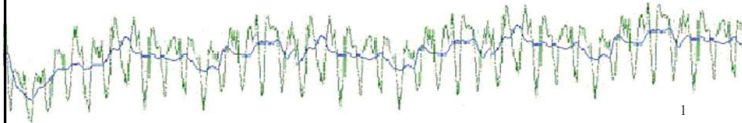


## CHƯƠNG 4. MÔ HÌNH HỒI QUY DỮ LIỆU BẢNG

Kinh tế lượng ứng dụng  
Khoa Kinh tế quốc tế Đại học Ngân Hàng Tp.HCM

Đỗ Hoàng Oanh



1

## Tài liệu tham khảo cho nội dung

- **Phạm Thị Tuyết Trinh, Kinh tế lượng ứng dụng trong kinh tế và tài chính, Khoa Kinh tế quốc tế, 2016**  
**Đặt sách: Khoa Kinh tế quốc tế**
- **Brooks, Chris, Introductory econometrics for finance, 2<sup>nd</sup> edition, Cambridge University Press, 2008.**
- **Enders, Walter, Applied Econometric Time Series, 3<sup>rd</sup> edition, Wiley, 2010.**
- **Nguyễn Quang Dong, Kinh tế lượng, NXB Đại học Kinh tế quốc dân, 2012.**



Tadyathā: Om Bhaiṣajye Bhaiṣajye Bhaiṣajya Samudgate Svāhā

2

## I. Phân loại và tổ chức dữ liệu bảng

- Dữ liệu bảng (panel data/ longitudinal data) là dữ liệu kết hợp của dữ liệu chuỗi thời gian và dữ liệu chéo. Một bảng (panel) bao gồm các quan sát của nhiều đơn vị chéo/ cá thể/ hoặc đối tượng (cross-sectional units/ individuals/ objects), ký hiệu là  $i$  ( $i = 1, 2, 3, \dots, N$ ), được lặp lại tại các giai đoạn khác nhau, ký hiệu là  $t$  ( $t = 1, 2, 3, \dots, T$ ). Biến số của dữ liệu bảng được ký hiệu là  $Y_{it}$ .



Tadyathā: Om Bhaiṣajye Bhaiṣajye Bhaiṣajya Samudgate Svāhā

3

## Phân loại và tổ chức dữ liệu bảng (tt)

- **Dữ liệu bảng cân đối và không cân đối**
  - Dữ liệu bảng cân đối (balanced panel data) có tất cả các đơn vị chéo hiện diện đầy đủ trong tất cả các khoảng thời gian. Nói cách khác, số lượng đơn vị chéo  $i$  tại từng quan sát thời gian  $t$  là bằng nhau, hoặc số quan sát thời gian  $t$  của từng đơn vị chéo  $i$  là bằng nhau.



Tadyathā: Om Bhaiṣajye Bhaiṣajye Bhaiṣajya Samudgate Svāhā

4

## Phân loại và tổ chức dữ liệu bảng (tt)

### • Dữ liệu bảng rời và bảng gộp

- Dữ liệu bảng tổ chức rời (Unstacked panel data):

FIRM	YEAR	INV	K	V	FIRM	YEAR	INV	K	V
1	1935	317.6	2.8	3078.5	2	1935	209.9	53.8	1362.4
1	1936	391.8	52.6	4661.7	2	1936	355.3	50.5	1807.1
1	1937	410.6	156.9	5387.1	2	1937	469.9	118.1	2676.3
1	1938	257.7	209.2	2792.2	2	1941	472.8	261.4	2380.5
1	1939	330.8	203.4	4313.2	2	1952	645.5	444.2	2159.4
1	1940	461.2	207.2	4643.9	2	1953	641	623.6	2031.3
					2	1954	459.3	669.7	2115.5

FIRM	YEAR	INV	K	V	FIRM	YEAR	INV	K	V
3	1935	33.1	97.8	1170.6	4	1935	40.29	10.5	417.5
3	1936	45	104.4	2015.8	4	1936	72.76	10.2	837.8
3	1937	77.2	118	2803.3	4	1937	66.26	34.7	883.9
3	1938	44.6	156.2	2039.7	4	1938	51.6	51.8	437.9
3	1939	48.1	172.6	2256.2	4	1939	52.41	64.3	679.7
3	1940	74.4	186.6	2132.2	4	1940	69.41	67.1	727.8
3	1941	113	220.9	1834.1	4	1941	68.35	75.2	643.6
3	1942	91.9	287.8	1588					
3	1943	61.3	319.9	1749.4					



Tadyathā: Om Bhaishajye Bhaishajye Bhaishajya Samudgate Svaha

## Phân loại và tổ chức dữ liệu bảng (tt)

- Dữ liệu bảng tổ chức gộp (Stacked panel data):

FIRM	YEAR	INV	K	V
1	1935	317.6	2.8	3078.5
1	1936	391.8	52.6	4661.7
1	1937	410.6	156.9	5387.1
1	1938	257.7	209.2	2792.2
1	1939	330.8	203.4	4313.2
1	1940	461.2	207.2	4643.9
2	1935	209.9	53.8	1362.4
2	1936	355.3	50.5	1807.1
2	1937	469.9	118.1	2676.3
2	1941	472.8	261.4	2380.5
2	1952	645.5	444.2	2159.4
2	1953	641	623.6	2031.3
2	1954	459.3	669.7	2115.5
3	1935	33.1	97.8	1170.6
3	1936	45	104.4	2015.8
3	1937	77.2	118	2803.3
3	1938	44.6	156.2	2039.7
3	1939	48.1	172.6	2256.2
3	1940	74.4	186.6	2132.2
3	1941	113	220.9	1834.1
3	1942	91.9	287.8	1588



Tadyathā: Om Bhaishajye Bhaishajye Bhaishajya Samudgate Svaha

## Phân loại và tổ chức dữ liệu bảng (tt)

### ■ Dữ liệu bảng cân đối và không cân đối

- Dữ liệu bảng không cân đối (unbalanced panel data) khi một hoặc nhiều đơn vị chéo trong bộ dữ liệu có số quan sát thời gian khác nhau, nói cách khác, từng thời gian  $t$  có số đơn vị chéo  $i$  không như nhau. Tổng số quan sát của dữ liệu bảng do đó khác  $N \times T$ .



Tadyathā: Om Bhaishajye Bhaishajye Bhaishajya Samudgate Svaha

## II. Ưu điểm dữ liệu bảng

■ Dữ liệu bảng cung cấp dữ liệu của nhiều đơn vị chéo theo thời gian, nên cũng chứa những đặc điểm không đồng nhất giữa chúng. Dữ liệu bảng cho phép giải thích sự khác biệt hay sự không đồng nhất (heterogeneity) của các đơn vị chéo. Các đơn vị chéo khác nhau thường không đồng nhất với nhau.



Tadyathā: Om Bhaishajye Bhaishajye Bhaishajya Samudgate Svaha

## Ưu điểm dữ liệu bảng (tt)

- Cung cấp số lượng quan sát lớn.
- Làm giảm bớt hiện tượng đa cộng tuyến thường gặp trong các mô hình dữ liệu chuỗi thời gian nhiều biến giải thích



Tadyathā: Om Bhaisajye Bhaisajye Bhaisajya Samudgate Svāhā

9

## Ưu điểm dữ liệu bảng (tt)

- Sự kết hợp của dữ liệu chuỗi thời gian và dữ liệu chéo, dữ liệu bảng cho phép vừa phân tích được tính động theo thời gian vừa phân tích được sự khác nhau giữa các đơn vị chéo nhờ thành phần chéo trong dữ liệu



Tadyathā: Om Bhaisajye Bhaisajye Bhaisajya Samudgate Svāhā

10

## Ưu điểm dữ liệu bảng (tt)

- Các mô hình dữ liệu bảng cho phép xây dựng và kiểm định những mô hình hành vi phức tạp hơn so với dữ liệu chéo hay dữ liệu chuỗi thời gian thuần túy, như mô hình hiệu quả về mặt kỹ thuật



Tadyathā: Om Bhaisajye Bhaisajye Bhaisajya Samudgate Svāhā

11

## Ưu điểm dữ liệu bảng (tt)

- Những thiên lệch do tổng hợp số liệu (data aggregation) về các công ty hoặc các cá nhân sẽ giảm đi trong dữ liệu bảng. Do đó dữ liệu bảng sẽ tạo ra những biến chính xác hơn so với số liệu thu thập và đo lường ở góc độ vĩ mô.



Tadyathā: Om Bhaisajye Bhaisajye Bhaisajya Samudgate Svāhā

12

### III. Mô hình OLS Pooled Panel

Mô hình vốn đầu tư Grunfeld có dạng:

$$Y_{it} = \beta_1 + \beta_2 X_{2it} + \beta_3 X_{3it} + u_{it}$$

Trong đó Y là tổng đầu tư thực (I);

$X_2$  là giá trị thực công ty (CAP);

$X_3$  là giá trị tài sản thực (PL).



Tadyathā: Om Bhaisajye Bhaisajye Bhaisajya Samudgate Svāhā

13

### Giả thuyết 1: $\beta_1$ , $\beta_2$ , và $\beta_3$ là không đổi theo đơn vị chéo và theo thời gian

- Nếu giả định rằng  $\beta_1$ ,  $\beta_2$ , và  $\beta_3$  là không đổi theo thời gian và đơn vị chéo, tức là giữa các công ty không có sự khác biệt về cách thức quản lý, triết lý quản lý, (...và nhiều yếu tố khác không thể quan sát được), và  $u_{it}$  đại diện cho tất cả sự khác biệt theo thời gian và từng công ty, khi đó phương pháp ước lượng bình phương bé nhất dữ liệu gộp (pooled OLS) sẽ cho ước lượng hiệu quả.



Tadyathā: Om Bhaisajye Bhaisajye Bhaisajya Samudgate Svāhā

14

### Giả thuyết 1: $\beta_1$ , $\beta_2$ , và $\beta_3$ là không đổi theo đơn vị chéo và theo thời gian

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-63.30414	29.61420	-2.137628	0.0357
X2	0.110096	0.013730	8.018809	0.0000
X3	0.303393	0.049296	6.154553	0.0000

R-squared	0.756528	Mean dependent var	290.9154
Adjusted R-squared	0.750204	S.D. dependent var	284.8528
S.E. of regression	142.3682	Akaike info criterion	12.79149
Sum squared resid	1560690	Schwarz criterion	12.88081
Log likelihood	-508.8596	Hannan-Quinn criter.	12.82730
F-statistic	119.6292	Durbin-Watson stat	0.355645
Prob(F-statistic)	0.000000		

**$R^2 > DW$   
Spurious  
regression**



Tadyathā: Om Bhaisajye Bhaisajye Bhaisajya Samudgate Svāhā

15

### Giả thuyết 2: $\beta_1$ khác nhau theo các đơn vị chéo và không đổi theo thời gian; $\beta_2$ , và $\beta_3$ không đổi theo đơn vị chéo và thời gian

$$Y_{it} = \alpha_1 + \alpha_2 D_{2i} + \alpha_3 D_{3i} + \alpha_4 D_{4i} + \beta_2 X_{2it} + \beta_3 X_{3it} + u_{it} \quad (6.3)$$

Trong đó  $D_{2i}$ ,  $D_{3i}$ ,  $D_{4i}$  là các biến giả có qui ước nhận giá trị như sau:

- $D_{2i} = 1$  nếu quan sát thuộc về GM, 0 nếu ngược lại (không thuộc GM).
- $D_{3i} = 1$  nếu quan sát thuộc về US, 0 nếu ngược lại (không thuộc US).
- $D_{4i} = 1$  nếu quan sát thuộc về WEST, 0 nếu ngược lại (không thuộc WEST).



Tadyathā: Om Bhaisajye Bhaisajye Bhaisajya Samudgate Svāhā

16

**Giả thuyết 2:  $\beta_1$  khác nhau theo các đơn vị chéo và không đổi theo thời gian;  $\beta_2$ , và  $\beta_3$  không đổi theo đơn vị chéo và thời gian (tt)**

$\beta_{1i}$  thay đổi qua các công ty, nhưng đối với mỗi công ty  $\beta_{1i}$  không đổi theo thời gian, tức có tác động cố định theo thời gian (fixed effects).  $\beta_2$ ,  $\beta_3$  không đổi theo công ty cũng như thời gian.



Tadyathā: Om Bhaisajye Bhaisajye Bhaisajya Samudgate Svāhā

17

**Giả thuyết 2:  $\beta_1$  khác nhau theo các đơn vị chéo và không đổi theo thời gian;  $\beta_2$ , và  $\beta_3$  không đổi theo đơn vị chéo và thời gian (tt)**

View Proc Object Print Name Freeze Estimate Forecast Stats Resids				
Dependent Variable: Y				
Method: Panel Least Squares				
Sample: 1935 1954				
Periods included: 20				
Cross-sections included: 4				
Total panel (balanced) observations: 80				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-245.7924	35.81112	-6.863577	0.0000
X2	0.107948	0.017509	6.165319	0.0000
X3	0.346162	0.026664	12.98212	0.0000
D2	161.5722	46.45639	3.477932	0.0008
D3	339.6328	23.98633	14.15943	0.0000
D4	186.5666	31.50681	5.921468	0.0000
R-squared	0.934563	Mean dependent var	290.9154	
Adjusted R-squared	0.930141	S.D. dependent var	284.8528	
S.E. of regression	75.28890	Akaike info criterion	11.55258	
Sum squared resid	419462.9	Schwarz criterion	11.73123	
Log likelihood	-456.1032	Hannan-Quinn criter.	11.62421	
F-statistic	211.3706	Durbin-Watson stat	0.925636	
Prob(F-statistic)	0.000000			



Tadyathā: Om Bhaisajye Bhaisajye Bhaisajya Samudgate Svāhā

18

**Giả thuyết 3:  $\beta_1$  khác nhau theo các đơn vị chéo và thời gian;  $\beta_2$ , và  $\beta_3$  không đổi theo thời gian và đơn vị chéo**

$$Y_{it} = \alpha_1 + \alpha_2 D_{GMi} + \alpha_3 D_{USi} + \alpha_4 D_{WESTi} + \lambda_0 + \lambda_1 D_{35} + \dots + \lambda_{19} D_{53} + \beta_2 X_{2i} + \beta_3 X_{3i} + u_{it}$$



Tadyathā: Om Bhaisajye Bhaisajye Bhaisajya Samudgate Svāhā

19

**Giả thuyết 4:  $\beta_1$ ,  $\beta_2$ , và  $\beta_3$  khác nhau theo các đơn vị chéo nhưng không thay đổi theo thời gian**

$$Y_{it} = \alpha_1 + \alpha_2 D_{2i} + \alpha_3 D_{3i} + \alpha_4 D_{4i} + \beta_2 X_{2it} + \beta_3 X_{3it} + \gamma_1 (D_{2i} X_{2it}) + \gamma_2 (D_{2i} X_{3it}) + \gamma_3 (D_{3i} X_{2it}) + \gamma_4 (D_{3i} X_{3it}) + \gamma_5 (D_{4i} X_{2it}) + \gamma_6 (D_{4i} X_{3it}) + u_{it}$$



Tadyathā: Om Bhaisajye Bhaisajye Bhaisajya Samudgate Svāhā

20

## IV. Vấn đề biến bị bỏ sót

$$Y_{it} = \beta X_{it} + \alpha_i + u_{it}$$

Trong đó,  $Y_{it}$ ,  $X_{it}$  có thể quan sát được,  $\alpha_i + u_{it}$  không thể quan sát được;  $\alpha_i$  là hằng số khác nhau theo các đơn vị chéo đại diện cho tác động của biến không quan sát được;  $u_{it}$  biến thiên theo thời gian và theo đơn vị chéo và không có hiện tượng tương quan chuỗi.



## Vấn đề biến bị bỏ sót (tt)

$$Y_{it} = \beta X_{it} + \alpha_i + u_{it}$$

Mối quan hệ giữa  $X_{it}$  và  $Y_{it}$  được phản ánh qua hệ số ước lượng  $\beta$ . Tuy nhiên, do có  $\alpha_i$  là hằng số của từng đơn vị chéo, nếu ước lượng mô hình bằng OLS,  $\alpha_i$  cũng sẽ xuất hiện trong sai số.



## Vấn đề biến bị bỏ sót (tt)

Trong trường hợp  $\alpha_i$  không tương quan với  $X_{it}$ ,  $\alpha_i$  chỉ là một yếu tố không thể quan sát được trong phần dư. Ước lượng OLS sẽ không còn hiệu quả, bởi vì sai số có hiện tượng tự tương quan



## Vấn đề biến bị bỏ sót (tt)

Trong trường hợp  $\alpha_i$  tương quan với  $X_{it}$ , Hệ số ước lượng  $\hat{\beta}$  khi đó không còn nhất quán và bị chệch, phương sai ước lượng  $var(\hat{\beta})$  cũng bị chệch so với mô hình vốn có. Do  $\alpha_i$  xuất hiện trong sai số sẽ dẫn đến vấn đề nghiêm trọng, làm  $X_{it}$  tương quan với  $u_{it}$ , trong ước lượng dữ liệu bảng, vấn đề này được gọi là vấn đề biến bị bỏ sót.



## Mô hình tác động cố định (FEM)

FEM giả định các hệ số hồi qui riêng phần giống nhau giữa các đơn vị chéo, nhưng các hệ số chặn hồi qui được phân biệt giữa các đơn vị chéo.

$$Y_{it} = \alpha_i + \beta_1 X_{1,it} + \beta_2 X_{2,it} + \dots + \beta_k X_{k,it} + u_{it}$$



Tadyathā: Om Bhaiṣajye Bhaiṣajye Bhaiṣajya Samudgate Svāhā

25

## Mô hình tác động cố định (FEM) (tt)

Cụ thể :

$$Y_{1t} = \alpha_1 + \beta_1 X_{1,1t} + \beta_2 X_{2,1t} + \dots + \beta_K X_{K,1t} + u_{1t}$$

$$Y_{2t} = \alpha_2 + \beta_1 X_{1,2t} + \beta_2 X_{2,2t} + \dots + \beta_K X_{K,2t} + u_{2t}$$

...

$$Y_{Nt} = \alpha_N + \beta_1 X_{1,Nt} + \beta_2 X_{2,Nt} + \dots + \beta_K X_{K,Nt} + u_{Nt}$$



Tadyathā: Om Bhaiṣajye Bhaiṣajye Bhaiṣajya Samudgate Svāhā

26

## Mô hình tác động cố định (FEM) (tt)

Tham số  $\beta_k$  chung cho tất cả các đơn vị chéo phản ánh tất cả các đơn vị chéo phản ánh các đơn vị chéo có tốc độ tăng giống nhau



Tadyathā: Om Bhaiṣajye Bhaiṣajye Bhaiṣajya Samudgate Svāhā

27

## Mô hình tác động cố định (FEM) (tt)

Tham số  $\alpha_i$  bao gồm hệ số chặn và biến bị bỏ sót của từng đơn vị chéo, được gọi là tham số đặc trưng của đối tượng (subject-specific parameters), đồng thời cũng được gọi là thành phần tác động cố định (fixed effect). Tác động cố định ở đây có nghĩa rằng  $\alpha_i$  không thay đổi theo thời gian. Sự xuất hiện của  $\alpha_i$  giúp phản ánh sự không đồng nhất giữa các đơn vị chéo do tác động của các biến không thể quan sát được, nhờ đó, FEM giải quyết được vấn đề biến bị bỏ sót.



Tadyathā: Om Bhaiṣajye Bhaiṣajye Bhaiṣajya Samudgate Svāhā

28



## Mô hình tác động cố định (FEM) (tt)

### ■ FEM có các giả định như sau:

- $E(u_{it}|X_i, \alpha_i) = 0$  [trung bình bằng 0]
- $\text{var}(u_{it}|X_i, \alpha_i) = \text{var}(u_{it}) = \sigma_u^2$  [phương sai không đổi cho tất cả  $t=1, \dots, T$ ]
- $\text{cov}(u_{it}, u_{is}|X_i, \alpha_i) = 0$  với  $t \neq s$  [các sai số ngẫu nhiên không tương quan với nhau]
- Với điều kiện của  $X_i$  và  $\alpha_i$ ,  $u_{it}$  thì độc lập và nhất quán. Do đó, sai số ngẫu nhiên theo phân phối chuẩn  $u_{it} \sim N(0; \sigma_u^2)$



Tadyathā: Om Bhaisajye Bhaisajye Bhaisajya Samudgate Svāhā

29

## Mô hình tác động cố định (FEM) (tt)

$H_0: \alpha_1 = \alpha_2 = \dots = \alpha_n = \alpha$ : không tồn tại tác động đặc trưng của từng đối tượng

$H_1$ : có ít nhất một trong các hệ số chặn  $\alpha_1, \alpha_2, \dots, \alpha_n$  sẽ khác nhau tùy thuộc vào đối tượng

$\text{Prob} < \alpha$ : bác bỏ giả thuyết  $H_0$



Tadyathā: Om Bhaisajye Bhaisajye Bhaisajya Samudgate Svāhā

30

## Mô hình tác động cố định (FEM) 1 chiều

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-73.84947	37.52291	-1.968117	0.0528
X2	0.107948	0.017509	6.165319	0.0000
X3	0.346162	0.029664	12.98212	0.0000

Effects Specification			
Cross-section fixed (dummy variables)			
R-squared	0.934563	Mean dependent var	290.9154
Adjusted R-squared	0.930141	S.D. dependent var	284.8528
S.E. of regression	75.28890	Akaike info criterion	11.55258
Sum squared resid	419462.9	Schwarz criterion	11.73123
Log likelihood	-456.1032	Hannan-Quinn criter.	11.62421
F-statistic	211.3706	Durbin-Watson stat	0.925636
Prob(F-statistic)	0.000000		



Tadyathā: Om Bhaisajye Bhaisajye Bhaisajya Samudgate Svāhā

## Mô hình tác động cố định (FEM) 2 chiều (tt)

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-129.0277	58.86415	-2.191967	0.0326
X2	0.129307	0.027424	4.715156	0.0000
X3	0.367249	0.041659	8.815577	0.0000

Effects Specification			
Cross-section fixed (dummy variables)			
Period fixed (dummy variables)			
R-squared	0.948925	Mean dependent var	290.9154
Adjusted R-squared	0.926638	S.D. dependent var	284.8528
S.E. of regression	77.15380	Akaike info criterion	11.77979
Sum squared resid	327399.0	Schwarz criterion	12.52417
Log likelihood	-446.1914	Hannan-Quinn criter.	12.07823
F-statistic	42.57690	Durbin-Watson stat	0.891196
Prob(F-statistic)	0.000000		

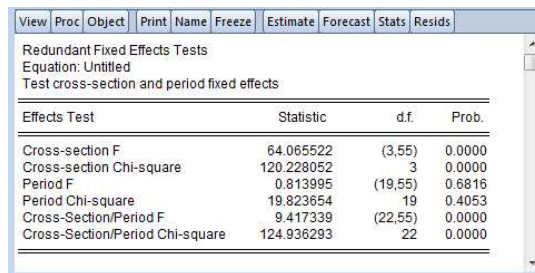


Tadyathā: Om Bhaisajye Bhaisajye Bhaisajya Samudgate Svāhā



## Mô hình tác động cố định (FEM) 2 chiều (tt)

Kiểm định tồn tại sự không đồng nhất giữa các đối tượng và thời gian



Effects Test	Statistic	d.f.	Prob.
Cross-section F	64.065522	(3,55)	0.0000
Cross-section Chi-square	120.228052	3	0.0000
Period F	0.813995	(19,55)	0.6816
Period Chi-square	19.823654	19	0.4053
Cross-Section/Period F	9.417339	(22,55)	0.0000
Cross-Section/Period Chi-square	124.936293	22	0.0000



Tadyathā: Om Bhaishajye Bhaishajye Bhaishajya Samudgate Svāhā

## Mô hình tác động ngẫu nhiên (REM)

Mô hình REM được trình bày như sau:

$$Y_{it} = \alpha + \beta_1 X_{1,it} + \beta_2 X_{2,it} + \dots + \beta_k X_{k,it} + \omega_{it}$$

$$\text{với } \omega_{it} = \varepsilon_i + v_{it}$$

Trong đó,  $\alpha$  là hệ số chặn chung của tất cả đơn vị chéo,  $\omega_{it}$  là sai số phức hợp (composite error term or error components),  $\varepsilon_i$  trong thành phần của  $\omega_{it}$  phản ánh tác động đặc trưng của từng đơn vị chéo và được gọi là thành phần tác động ngẫu nhiên (random effect);  $v_{it}$  là hạng nhiễu không tương quan lẫn nhau giữa các đối tượng (còn gọi là tương quan chéo, cross-correlation) và không tương quan chuỗi trong cùng đối tượng.



Tadyathā: Om Bhaishajye Bhaishajye Bhaishajya Samudgate Svāhā

34

## Mô hình tác động ngẫu nhiên (REM) (tt)

REM có thể xác định được:

- Các hệ số chặn khác nhau cho từng đơn vị chéo;
- Tác động chung (không thay đổi theo đơn vị chéo) của các biến giải thích.
- $\varepsilon_i$  đo lường độ lệch ngẫu nhiên (random deviation) giữa hệ số chặn của mỗi đối tượng và hệ số chặn chung  $\alpha$ .



Tadyathā: Om Bhaishajye Bhaishajye Bhaishajya Samudgate Svāhā

35

## Mô hình tác động ngẫu nhiên (REM) (tt)

Tuy nhiên, các hệ số chặn của từng đơn vị chéo được phát sinh từ: (i) một hệ số chặn chung  $\alpha$  không đổi theo đối tượng và thời gian; (ii) và một biến ngẫu nhiên  $\varepsilon_i$  (không tương quan với  $X_{it,k}$ ) là một thành phần của sai số thay đổi theo đối tượng nhưng không đổi theo thời gian (chính vì vậy mô hình còn được là mô hình các thành phần sai số)



Tadyathā: Om Bhaishajye Bhaishajye Bhaishajya Samudgate Svāhā

36

## Mô hình tác động ngẫu nhiên (REM) (tt)

### ■ REM có các giả định như sau:

- $E(\omega_{it}) = 0$  [trung bình bằng 0]
- $\text{var}(\omega_{it}) = \sigma_u^2 + \sigma_v^2$  [phương sai không đổi]
- $\text{cov}(\omega_{it}, \omega_{is}) = \sigma_v^2$  với  $t \neq s$  [các sai số ngẫu nhiên trong cùng một đơn vị chéo  $i$  ở các đơn vị thời gian tương quan với nhau]
- $\text{cov}(\omega_{it}, \omega_{js}) = 0$  với  $i \neq j$  [các sai số ngẫu nhiên của các đơn vị chéo khác nhau theo thời gian khác nhau không tương quan với nhau]



Tadyathā: Om Bhaisajye Bhaisajye Bhaisajya Samudgate Svāhā

37

## Mô hình tác động ngẫu nhiên (REM) (tt)

### ■ REM có các giả định như sau:

- $\text{cov}(\varepsilon_i, X_{it,k}) = 0$  [các phần dư  $\varepsilon_i$ , thành phần tác động ngẫu nhiên của từng đơn vị chéo không tương quan với các biến X's]. Lưu ý, đây là giả định khác với FEM.
- $\text{cov}(X_{it,k}, v_{it}) = 0$ ,  $\text{cov}(V_i, X_{3it}) = [v_{it} \text{ không tương quan với các biến X's}]$



Tadyathā: Om Bhaisajye Bhaisajye Bhaisajya Samudgate Svāhā

38

## Mô hình tác động REM 1 chiều (tt)

View	Proc	Object	Print	Name	Freeze	Estimate	Forecast	Stats	Resids
Dependent Variable: Y Method: Panel EGLS (Cross-section random effects) Sample: 1935 1954 Periods included: 20 Cross-sections included: 4 Total panel (balanced) observations: 80 Swamy and Arora estimator of component variances									
Variable	Coefficient	Std. Error	t-Statistic	Prob.					
C	-73.03531	84.23769	-0.867015	0.3886					
X2	0.107655	0.016875	6.379722	0.0000					
X3	0.345710	0.026636	12.97896	0.0000					
Effects Specification					S.D.	Rho			
Cross-section random					152.1582	0.8033			
Idiosyncratic random					75.28690	0.1967			
Weighted Statistics									
R-squared	0.804982	Mean dependent var	31.99227						
Adjusted R-squared	0.799896	S.D. dependent var	167.7316						
S.E. of regression	75.03139	Sum squared resid	433487.6						
F-statistic	158.8972	Durbin-Watson stat	0.903231						
Prob(F-statistic)	0.000000								
Unweighted Statistics									
R-squared	0.753716	Mean dependent var	290.9154						
Sum squared resid	1578718.	Durbin-Watson stat	0.348706						



Tadyathā: Om Bhaisajye Bhaisajye Bhaisajya Samudgate Svāhā

## Mô hình tác động REM 2 chiều (tt)

View	Proc	Object	Print	Name	Freeze	Estimate	Forecast	Stats	Resids
Dependent Variable: Y Method: Panel EGLS (Two-way random effects) Sample: 1935 1954 Periods included: 20 Cross-sections included: 4 Total panel (balanced) observations: 80 Swamy and Arora estimator of component variances									
Variable	Coefficient	Std. Error	t-Statistic	Prob.					
C	-72.99761	82.31654	-0.886792	0.3780					
X2	0.107642	0.016801	6.407000	0.0000					
X3	0.345687	0.026565	13.01274	0.0000					
Effects Specification					S.D.	Rho			
Cross-section random					152.1115	0.7954			
Period random					0.000000	0.0000			
Idiosyncratic random					77.15380	0.2046			
Weighted Statistics									
R-squared	0.804871	Mean dependent var	32.78472						
Adjusted R-squared	0.799803	S.D. dependent var	167.8286						
S.E. of regression	75.09221	Sum squared resid	434190.7						
F-statistic	158.9055	Durbin-Watson stat	0.902046						
Prob(F-statistic)	0.000000								
Unweighted Statistics									
R-squared	0.753722	Mean dependent var	290.9154						
Sum squared resid	1578678.	Durbin-Watson stat	0.348704						



Tadyathā: Om Bhaisajye Bhaisajye Bhaisajya Samudgate Svāhā

## Kiểm định Hausman test

Giả thuyết  $H_0: \text{cov}(\varepsilon_i, X_{it,k}) = 0$

Giả thuyết  $H_1: \text{cov}(\varepsilon_i, X_{it,k}) \neq 0$

Giả thuyết  $H_0: \beta_{\text{FEM}} = \beta_{\text{REM}}$

Giả thuyết  $H_1: \beta_{\text{FEM}} \neq \beta_{\text{REM}}$

$\text{Prob} < \alpha$  : giả thuyết  $H_0$  bị bác bỏ



Tadyathā: Om Bhaisajye Bhaisajye Bhaisajya Samudgate Svāhā

## Kiểm định Hausman test

Label	Location	S.D.	Rho
Cross-section random	152.1115	0.7954	
Period random	0.000000	0.0000	
Idiosyncratic random	77.15380	0.2046	

Weighted Statistics			
R-squared	0.804871	Mean dependent var	32.78472



Tadyathā: Om Bhaisajye Bhaisajye Bhaisajya Samudgate Svāhā

# Thank you for listening!

