المحاضرة الرابعة:

مدخل إلى نماذج بانل

المحاضرة الرابعة: مدخل إلى نماذج بانل

مقدمة

تمثل نماذج بانل (Panel Data) إحدى القفزات النوعية الحديثة في حقل القياس الاقتصادي، نظرا لإعطائها الفرصة للباحث أن يأخذ بعين الاعتبار البعد الزمني والبعد الفردي (المكاني) في الدراسات الاقتصادية التطبيقية، وقد تم استخدام هذا المصطلح في الأبحاث الدولية حوالي 25 مرة خلال الفترة الممتدة من سنة 1975 إلى غاية 1995 حسب ما أشار إليه الباحث Patrick Sevestre[1]، فمن بين المزايا التي تتمتع بها هذه النماذج دورها في تحسين كفاءة المقدرات من خلال رفع درجة الحرية والحد من مشكلة التعدد الخطي بين المتغيرات التفسيرية[2]،كما تعمل على كشف العلاقات الديناميكية والتخلص من مشكلة غياب بعض المشاهدات حول بعض المتغيرات المدروسة التي عادة ما تولد مقدرات متحيزة، إضافة إلى هذا فإن هذا النوع من النماذج سيمكن الباحثين من الحصول على تنبؤات جيدة والتحكم في عدم تجانس التباين الذي قد يظهر في حالة البيانات المقطعية أو حالة البيانات الزمنية...الخ

سنقدم في الفصل الأول من هذه المطبوعة نظرة شاملة عن نماذج بانل الساكنة والديناميكية وطرق تقديرهما.

أولا: نماذج بانل الساكنة

نخصص الجزء الأول من هذا الفصل لتناول البنية الرياضية لهذا النوع من النماذج، اختبار التجانس للباحث Hsiao، والأشكال الشهيرة لنماذج Panel وطرق تقديرها.

1. الشكل الهيكلي لبيانات بانل Panel Data

تمثل نماذج بانيل إحدى الطرق المستخدمة في القياس الاقتصادي التي تمكن الباحث من النظر إلى البيانات من خلال بعدين[3]: البعد الزمني والبعد المقطعي، والأمثلة على ذلك عديدة كتقدير العلاقة القياسية بين الصادرات والانفتاح التجاري على مدى 20 سنة (البعد الزمني) و 15 دولة (الأفراد)، أو دراسة العلاقة بين دخل الأسر واستهلاكها على مدى عدة سنوات...

فنجد البعد الفردي قد يكون ممثلا في: مؤسسات، عائلات، بلديات، ولايات، دول،...الخ، أما البعد الزمني فيكون في شكل سنوي، نصف سنوي، فصلي، شهري...الخ.

---

(1) Patrick Sevestre (2002), **économétrie des données de panel**, Paris : Dunod, P 01.

(2) لمزيد من التفصيل يمكنك الرجوع إلى:

Cheng Hsiao (2014), **Analysis of panel data**, UK : Cambridge university press, third edition, PP 04-05

(3) Régis Bourbonnais (2015), **économétrie : cours et exercices corrigés**, Paris : Dunod, 9e édition, P345.

ولتوضيح هيكل بيانات بانل سنأخذ المثال التالي والمتعلق بمحددات معدل الجريمة كمتغيرة تابعة (عدد الحالات لكل ألف شخص)، حيث سنقترح متغيرتين مفسرتين وهما معدل البطالة (%) وحجم السكان (بالآلاف) خلال سنتين وهما 1992 و1998 لثمان مدن، كما هو مبين في الجدول التالي:

جدول رقم (1): مثال حول بيانات بانل

city	year	pop	unem	crime
1	92	85,1	7,8	75,42648
1	98	85,1	3,1	70,01455
2	92	42,6	7,6	93,61669
2	98	42,6	5,5	90,34896
3	92	134,5	8,3	85,18089
3	98	134,5	7,7	76,95070
4	92	168,2	11,6	89,43108
4	98	168,2	5,2	83,96927
5	92	34,7	12,9	107,90584
5	98	34,7	7,7	104,56469
6	92	59,8	14,0	137,31123
6	98	59,8	4,8	112,26517
7	92	18,6	9,3	71,97062
7	98	18,6	5,8	86,28617
8	92	95,3	8,0	97,52621
8	98	95,3	4,6	76,55205

المصدر: https://real-statistics.com/panel-data-models/panel-data-two-time-periods/ **تاريخ التحميل: 2023/12/27**

من خلال الجدول أعلاه نلاحظ أن عدد المشاهدات الإجمالي هو 16 ، لأن عدد الأفراد يساوي N=8 ، أما عدد السنوات فهو T=2 ، أي أن العدد الإجمالي هو : N×T ، كما يمكن الترميز للمدن بالرمز i=1,2,.....N ، أما السنوات t=1,2,.....T .

ويمكن تقسيم بيانات بانل من حيث توفر أو عدم توفر المشاهدات إلى:

- **بيانات بانل المتوازنة Balanced Panel Data**: جميع الأفراد لهم نفس عدد الفترات الزمنية؛
- **بيانات بانل غير المتوازنة Unbalaced Panel Data:** كل فرد له فترات زمنية خاصة به.

يمكن تقديم هذا المثال الذي يوضح هذين الصنفين:

جدول رقم (2): بيانات بانل المتوازنة وغير المتوازنة.

بيانات بانل متوازنة

	A	B	C	D	E
1	i	t	GDP	Consumption	Price
2	1	1980	9,14206153	4,70048037	2,34535718
3	1	1981	9,24657586	4,85203026	2,36755789
4	1	1982	9,29807649	4,89034913	2,52852473
5	1	1983	9,36383369	4,96284463	2,69311929
6	1	1984	9,44208681	5,06259503	2,70394847
7	1	1985	9,50286072	5,14749448	2,7263877
8	1	1986	9,55215545	5,18178355	2,7292636
9	1	1987	9,60326045	5,25749537	2,73954887
10	1	1988	9,6718707	5,35185813	2,76579067
11	1	1989	9,74020362	5,36597602	2,9385268
12	1	1990	9,81743935	5,42934563	2,93015855
13	2	1980	9,15207546	3,95124372	2,74373261
14	2	1981	9,27977315	4,11087386	2,85481366
15	2	1982	9,29155163	4,15888308	2,97297529
16	2	1983	9,36280392	4,24849524	2,91086042
17	2	1984	9,43755562	4,30406509	2,72110462
18	2	1985	9,49829741	4,38202663	2,83238948
19	2	1986	9,53481232	4,4543473	3,00848067
20	2	1987	9,61132881	4,49980967	3,06625131
21	2	1988	9,67325644	4,58496748	3,18265987
22	2	1989	9,71595243	4,67282883	3,25064965
23	2	1990	9,72579526	4,76217393	3,33581228
24	3	1980	9,10052551	4,60517019	2,80486833
25	3	1981	9,19136141	4,68213123	2,78833903
26	3	1982	9,26198361	4,75359019	2,859798
27	3	1983	9,3064684	4,83628191	2,94090115
28	3	1984	9,37822488	4,84418709	3,0435553
29	3	1985	9,41743584	4,93447393	3,08017911
30	3	1986	9,4526589	4,99043259	3,12348764
31	3	1987	9,50315923	5,07517382	3,15734504
32	3	1988	9,58231764	5,18178355	3,18263084
33	3	1989	9,66205273	5,23644196	3,21215328
34	3	1990	9,72752573	5,26785816	3,23254991

بيانات بانل غير متوازنة

	A	B	C	D	E
1	i	t	GDP	Consumption	Price
2	1	1980	9,14206153	NaN	2,34535718
3	1	1981	9,24657586	NaN	2,36755789
4	1	1982	9,29807649	4,89034913	2,52852473
5	1	1983	9,36383369	4,96284463	2,69311929
6	1	1984	9,44208681	5,06259503	2,70394847
7	1	1985	9,50286072	5,14749448	2,7263877
8	1	1986	9,55215545	5,18178355	2,7292636
9	1	1987	9,60326045	5,25749537	2,73954887
10	1	1988	9,6718707	5,35185813	2,76579067
11	1	1989	9,74020362	5,36597602	NaN
12	1	1990	9,81743935	5,42934563	NaN
13	2	1980	9,15207546	3,95124372	2,74373261
14	2	1981	9,27977315	4,11087386	2,85481366
15	2	1982	9,29155163	4,15888308	2,97297529
16	2	1983	9,36280392	4,24849524	2,91086042
17	2	1984	9,43755562	4,30406509	2,72110462
18	2	1985	9,49829741	4,38202663	2,83238948
19	2	1986	9,53481232	4,4543473	3,00848067
20	2	1987	9,61132881	4,49980967	3,06625131
21	2	1988	9,67325644	4,58496748	3,18265987
22	2	1989	9,71595243	4,67282883	3,25064965
23	2	1990	9,72579526	4,76217393	3,33581228
24	3	1980	NaN	4,60517019	2,80486833
25	3	1981	NaN	4,68213123	2,78833903
26	3	1982	NaN	4,75359019	2,859798
27	3	1983	9,3064684	4,83628191	2,94090115
28	3	1984	9,37822488	4,84418709	3,0435553
29	3	1985	9,41743584	4,93447393	3,08017911
30	3	1986	9,4526589	4,99043259	3,12348764
31	3	1987	9,50315923	5,07517382	3,15734504
32	3	1988	9,58231764	5,18178355	3,18263084
33	3	1989	9,66205273	5,23644196	3,21215328
34	3	1990	9,72752573	5,26785816	3,23254991

Christophe Hurlin (2018) ; **Advanced Econometrics II**, School of Economics and Management - University of Geneva, University of Orléans ; February 2018 ; P17-19.

من خلال بيانات بانل المتوازنة نلاحظ أن $N=3$ و $T=11$، أما بيانات بانل غير المتوازنة فنلاحظ أنه بالنسبة للفرد الأول عدد الفترات الزمنية يساوي $T_1=7$، أما الفرد الثاني فـ $T_2=11$، أما الفرد الثالث فـ $T_3=8$، كما نلاحظ أن العدد الإجمالي للمشاهدات في بيانات بانل المتوازن هو $N\times T=33$، أما في بيانات بانل غير المتوازن فعدد المشاهدات هو: $T_1+T_2+T_3=26$.

كما يمكن تقسيم بيانات بانل تبعا لحجم الفترات الزمنية T وعدد الأفراد N ، حيث نجد[1]:

➢ **بيانات بانل القصيرة (Short Panel)**: يكون عدد الفترات الزمنية أقل بكثير جدا عن عدد الأفراد $(T\ small\ and\ N \to \infty)(T << N)$ ، عادة ما تظهر هذه البيانات من خلال المسوحات التي تجرى على عدد كبيرا جدا من الأفراد خلال فترات زمنية قصيرة؛

[1] Colin Cameron (2007) ; **Panel data methods for microeconometrics using Stata**, Univ. of California ; October 25, 2007 2018 ; P3-5 https://www.stata.com/meeting/wcsug07/cameronwcsug.pdf (01/01/2024).

- **بيانات بانل الطويلة (Long Panel)**: يكون البعد الزمني كبير جدا مع عدد أفراد قليل، أو عدد أفراد كبير $(T \to \infty\ and\ small\ N\ or\ N \to \infty)$.

إن الاختلاف أو التغير في بيانات بانل يمكن أن يأخذ ثلاث زاويا، الزاوية الأولى تتعلق بداخل الأفراد Within حيث يكون التغير حسب الزمن لكل فرد، الزاوية الثانية تأخذ بعين الاعتبار الاختلاف بين الأفراد between وهنا نقوم بتثبيت الزمن، أما الزاوية الثالثة وهو التغير الكلي أو الشامل Overall والذي يأخذ عامل الزمن وعامل الأفراد بعين الاعتبار. ولتوضيح هذا الأمر يمكن أخذ المثال التالي:

- مثال: لتكن لديك بيانات بانل للمتغيرة x_{it} حيث أن عدد الأفراد يساوي 3 وعدد الفترات الزمنية يساوي 3، المطلوب: أوجد التغيرات داخل الأفراد وما بين الأفراد والتغير الكلي أو الشامل للمتغيرة x_{it} ؟

ID	الزمن	المتغيرة	المتوسط الفردي Individual Mean	المتوسط الكلي Overall Mean	الانحرافات الكلية Overall Deviation	الانحرافات بين الأفراد Between Deviation	الانحرافات داخل الأفراد Within Deviation
i	t	x_{it}	$\overline{x_{i.}}$	$\bar{\bar{x}}$	$x_{it} - \bar{\bar{x}}$	$\overline{x_{i.}} - \bar{\bar{x}}$	$x_{it} - \overline{x_{i}}$
1	2019	9	10	20	-11	-10	-1
1	2020	10	10	20	-10	-10	0
1	2021	11	10	20	-9	-10	1
2	2019	20	20	20	0	0	0
2	2020	20	20	20	0	0	0
2	2021	20	20	20	0	0	0
3	2019	25	30	20	5	10	-5
3	2020	30	30	20	10	10	0
3	2021	35	30	20	15	10	5

Source : Rizaudin Sahlan (2016) ; **Within and Between Variation in Panel Data with Stata (Panel) ;** Universiti Utara Malaysia ; http://rizaudinsahlan.blogspot.com/2016/06/within-and-between-variation-in-panel.html (02/01/2024).

حيث أن:

- **المتوسط الفردي** يتم حسابه كما يلي: $\overline{x_{i.}} = \frac{1}{T}\sum_{t=1}^{T} x_{it}$ ؛
- **المتوسط الكلي** يتم حسابه كما يلي: $\bar{\bar{x}} = \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T} x_{it}$ ؛
- **التباين الكلي** يتم حسابه كما يلي: $\boldsymbol{S_O^2} = \frac{1}{NT-1}\sum_{i=1}^{N}\sum_{t=1}^{T}(x_{it} - \bar{\bar{x}})^2$ ؛
- **التباين ما بين الأفراد** يتم حسابه كما يلي: $\boldsymbol{S_B^2} = \frac{1}{N(T-1)} T\sum_{i=1}^{N}(\overline{x_{i.}} - \bar{\bar{x}})^2$ ؛
- **التباين داخل الأفراد** يتم حسابه كما يلي: $\boldsymbol{S_W^2} = \frac{1}{NT-1}\sum_{i=1}^{N}\sum_{t=1}^{T}(x_{it} - \overline{x_{\iota}})^2$ ؛
- $\sum_{i=1}^{N}\sum_{t=1}^{T}(x_{it} - \bar{\bar{x}})^2 = T\sum_{i=1}^{N}(\overline{x_{i.}} - \bar{\bar{x}})^2 + \sum_{i=1}^{N}\sum_{t=1}^{T}(x_{it} - \overline{x_{\iota}})^2$؛

اعتمادا على برنامج STATA وبواسطة الأمر xtsum يمكننا استخراج المتوسط الكلي، الانحراف المعياري الكلي، الانحراف المعياري ما بين الأفراد؛ الانحراف المعياري داخل الأفراد.

```
. xtsum x
```

Variable		Mean	Std. Dev.	Min	Max	Observations
x	overall	20	9.027735	9	35	N = 9
	between		10	10	30	n = 3
	within		2.54951	15	25	T = 3

المحاضرة الخامسة:

الهيكل الرياضي لنماذج بانل

المحاضرة الخامسة: الهيكل الرياضي لنماذج بانل

2. الهيكل الرياضي لنماذج بانل:

لفهم الطبيعة الرياضية لهذا النوع من النماذج يمكننا أخذ المثال التالي الذي يتعلق بدراسة التكاليف في ست شركات طيران خلال الفترة (1970-1984)[1]، فنموذج بانيل في هذه الحالة يحتوي على 90 مشاهدة، إن الصيغة الرياضية المقترحة لهذا النموذج هي:

$$C_{it} = \beta_{1i} + \beta_{2i}Q_{it} + \beta_{3i}PF_{it} + \beta_{4i}LF_{it} + \varepsilon_{it} \text{ .... } (1)$$
$$i = 1,2, \ldots ,6; t = 1,2, \ldots ,15$$

حيث أن:

C_{it} : التكاليف الكلية للشركةiخلال اللحظة الزمنيةt.

Q_{it}: المخرجات التي يتم قياساها بالمداخيل المتحصل عليها من طرف الركاب للشركةiخلال اللحظة الزمنيةt.

PF_{it}: أسعار الوقود الخاصة بالشركةiخلال اللحظة الزمنيةt.

LF_{it} : معامل الحمولة للشركةiخلال اللحظة الزمنيةt، والذي يُقَاس بمتوسط قدرة استخدام الأسطول.

إذا كان النموذج (1) يخضع للفرضيات الكلاسيكية حول الأخطاء العشوائية ونقصد بذلك على وجه الخصوص مسألة التجانس، الاستقلالية الزمنية، الاستقلالية بين فرد وفرد آخر$\left(cov\left(\varepsilon_{it}, \varepsilon_{jt} = 0,\ \ i \neq j\right)\right)$، يمكننا تطبيق طريقة المربعات الصغرى العادية.

أما إذا كانت $cov\left(\varepsilon_{it}, \varepsilon_{jt}\right) = \sigma_{ij}^2 \neq 0; i \neq j$ فإن مقدرات المربعات الصغرى العادية تفقد خاصية أفضل مقدر خطي غير متحيز، في هذه الحالة نقوم بتطبيق طريقة (Seemingly Unrelated Regressions, SUR)[2].

إن نظرة الباحث للنموذج رقم (1) تختلف حسب أربع حالات[3]:

- الحالة الأولى: إذا كان هنالك تجانس تام بين شركات الطيران، أي أن الحد الثابت هو نفسه بالنسبة لجميع الأفراد $(\beta_{1i} = \beta_1)$، كما أن معاملات المتغيرات التفسيرية هي نفسها، في هذه الحالة يكون النموذج

---

[1]Damodar N. Gujarati and Dawn C. Porter (2009), **Basic Econometrics**, New York: MC-Graw-Hill, Fifth edition , PP 593-595.

[2] لمزيد من التفصيل، يمكنك الرجوع إلى:

Régis Bourbonnais, Op.cit., PP347-348.

[3]Ibid, PP 348-349.

مكتوب وفق معادلة واحدة، والتي نقوم بتقديرها باستخدام طريقة المربعات الصغرى العادية أو طريقة المربعات الصغرى المعممة؛

- الحالة الثانية: في حالة عدم وجود تجانس تام بين شركات الطيران، أي أن قيم المعلمات تختلف حسب كل شركة، لا يمكننا أن نستخدم نماذج Panel، بل نقوم بتقدير كل معادلة على حدة؛
- الحالة الثالثة: عدم تجانس معاملات المتغيرات التفسيرية أي تختلف حسب شركات الطيران، وتجانس الحدود الثابتة أي تساويها بالرغم من الاختلاف بين الشركات، في هذه الحالة نقوم بتقدير كل معادلة انحدار على حدة، ويتم هنا رفض هيكل بانيل، مثل الحالة السابقة؛
- الحالة الرابعة: عدم تجانس الحدود الثابتة وتجانس معاملات المتغيرات التفسيرية، يطلق على هذا النوع من النماذج نموذج تأثيرات الأفراد.

لمعرفة أي حالة يمكننا استخدامها في دراستنا التطبيقية نلجأ إلى اختبار Hsiao (1986)، وهو ما سنتناوله في النقطة الموالية.

3. اختبار التجانس للباحث Hsiao :

لنفترض أنه لدينا النموذج الخطي التالي:

$$y_{it} = \beta_{0i} + \beta_i' x_{it} + \varepsilon_{it} \text{ .... } (2)$$
$$i = 1,2,\dots,N;\ t = 1,2,\dots,T$$

حيث أن:

y_{it} : تمثل المتغيرة الداخلية المشاهدة للفردiخلال اللحظة الزمنيةt؛

x_{it} : شعاع يحتوي على k متغيرة خارجية $x_{it}' = (x_{1it} \quad x_{2it} \quad \cdots \quad x_{kit})$؛

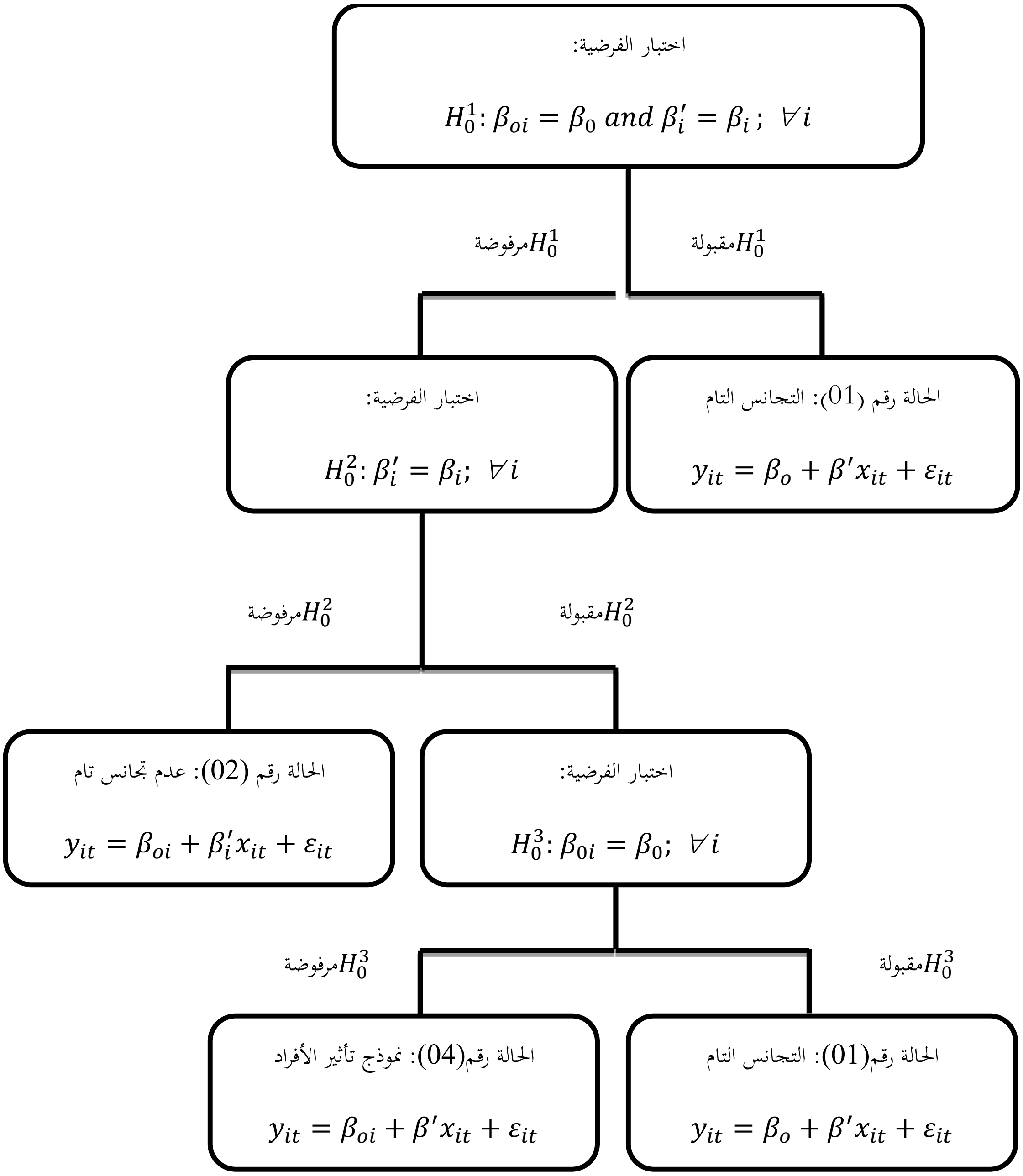
β_{oi}: الحد الثابت الخاص بالفرد i؛

β_i' : شعاع يتكون من k معلمة خاصة بالمتغيرات الخارجية $\beta_i' = (\beta_1 \quad \beta_2 \quad \cdots \quad \beta_k)$ ؛

ε_{it}: شعاع البواقي.

يكون تطبيق الخوارزمية التالية الملخصة في الشكل رقم (2):

شكل رقم (1): مراحل اختبار التجانس للباحث Hsiao



Source : Régis Bourbonnais, Op.cit., P349.

يتم اختبار الفرضيات: H_0^1، H_0^2 ، H_0^3 من خلال تقدير النماذج المقيدة والحرة ثم إجراء اختبار فيشر لاتخاذ القرار المناسب.

المحاضرة السادسة:

الأشكال الشهيرة لنماذج بانل وطرق تقديرها والمفاضلة بينها.

المحاضرة السادسة: الأشكال الشهيرة لنماذج بانل وطرق تقديرها والمفاضلة بينها

4. الأشكال الشهيرة لنماذج Panel وطرق تقديرها:

عند تقديرنا للنموذج (3-94) ككل دون الأخذ بعين الاعتبار الاختلاف بين شركات الطيران نكون في حالة نموذج الانحدار التجميعي(Pooled Regression Model, PM) ، أما إذا كان الحد الثابت β_{1i} يختلف من مجموعة إلى مجموعة أخرى أو من شركة طيران إلى شركة أخرى نكون في هذه الحالة نتعامل مع نموذج التأثيرات الثابتـــــــة (Fixed Effects Model, FEM)، أما إذا اختلت إحدى الفرضيات الكلاسيكية التي يقوم عليها نموذج التأثيرات الثابتة؛ فإننا نواجه نموذجا آخر يدعى بنموذج التأثيرات العشوائية (Random Effects Model , REM)، هذا وتجدر الإشارة إلى أنه في حالة غياب بعض المشاهدات حول بعض المتغيرات المدروسة خلال بعض الفترات الزمنية نكون في حالة نماذج بانيل غير المتوازنة (Unbalanced Panel Data)، أما إذا توفرت جميع المشاهدات فيطلق على هذا النوع تسمية نماذج بانيل المتوازنة (Balanced Panel Data)، سنتطرق إلى هذه النماذج بشيء من الإيجاز.

- **نموذج الانحدار التجميعي:**

إن نموذج الانحدار التجميعي يأخذ الصيغة التالية[1]:

$$y_{it} = \beta_0 + \beta' x_{it} + \varepsilon_{it} \text{ …….. } (1)$$
$$i = 1,2, \dots , N; t = 1,2, \dots , T$$

مع توفر الفرضيات التالية:

$$\begin{cases} E(\varepsilon_{it}/x_{i1}\, , x_{i2}, \cdots , x_{iT}) = 0 \\ Var(\varepsilon_{it}/x_{i1}, x_{i2}, \cdots , x_{iT}) = \sigma_\varepsilon^2 \\ Cov(\varepsilon_{it}, \varepsilon_{js}/x_{i1}, \cdots , x_{iT}) = 0, if\colon i \neq j; t \neq s \end{cases}$$

يتم في هذه الحالة تطبيق طريقة المربعات الصغرى العادية على النموذج (1) ، إلا أن هذه الحالة نادرة الوقوع، فالحد الثابت في طبيعته يختلف بين الأفراد أو من مجموعة لأخرى، وقد يكون في بعض الحالات متغيراً عشوائياً، لذلك فطريقة المربعات الصغرى العادية تفقد خواصها الشهيرة.

---

[1] لمزيد من التفصيل، يمكنك الرجوع إلى:

William H. Greene (2012), **Econometric Analysis**, England: Pearson Education Limited, Seventh Edition, PP389-399.

- **نموذج التأثيرات الثابتة:**

يأخذ هذا النموذج الصيغة الرياضية التالية:

$$y_{it} = \beta_{0i} + \beta' x_{it} + \varepsilon_{it} \ \ldots\ldots..\ (2)$$
$$i = 1, \cdots, N; t = 1, \ldots, T;\ \beta' = (\beta_1 \quad \beta_2 \quad \cdots \quad \beta_k)$$

إن طريقة تقدير النموذج (2) تعتمد على هيكل الأخطاء العشوائية، فإذا كانت الأخطاء العشوائية متماثلة ومتجانسة، وغير مرتبطة فيما بينها خلال البعد الزمني والفردي، الذي يُعَبَّرُ عنه إحصائيا كما يلي:

$$\begin{cases} cov(\varepsilon_{it}, \varepsilon_{it'}) = 0, t \neq t' \\ cov(\varepsilon_{it}, \varepsilon_{jt}) = 0, i \neq j \end{cases} \ \ldots\ldots.\ (3)$$

في ظل هذه الفرضيات، ولغرض تقدير معلمات النموذج، والسماح للمعلمة: β_{0i}بالتغير بين مجموعات الأفراد نستخدم متغيرات وهمية لكي نتجنب مشكلة التعدد الخطي التام، كما نستعمل طريقة المربعات الصغرى للمتغيرات الوهمية (Least Squares Dummy Variable Model, LSDV)،ولتطبيق هذه الطريقة نقوم ببناء N متغيرة ثنائية (وهمية) وفق الشكل الرياضي التالي[1]:

$$\begin{cases} D_i = 0, i \ \text{من الفردأجل} \\ D_i \neq 0, \text{غير ذلك} \end{cases}$$

يصبح النموذج (2) كما يلي:

$$\mathrm{y_{it}} = \beta_0 + \beta_{01}\mathrm{D}_1 + \beta_{01}\mathrm{D}_1 + \cdots + \beta_{0N}\mathrm{D_N} + \beta'\mathrm{x_{it}} + \varepsilon_{it} \ \ldots\ldots.\ (4)$$

من الناحية العملية نقوم بتقدير النموذج أعلاه بدون حد ثابتβ_0 لتفادي مشكلة التعدد الخطي التام، ليصبح النموذج كما يلي:

$$\mathrm{y_{it}} = \beta_1\mathrm{D}_1 + \beta_2\mathrm{D}_2 + \cdots + \beta_N\mathrm{D_N} + \beta'\mathrm{x_{it}} + \varepsilon_{it} \ \ldots\ldots.\ (5)$$

نقوم بتقدير معالم النموذج(5) بواسطة طريقة المربعات الصغرى العادية أو طريقة المربعات الصغرى المعممة، إذا كانت الأخطاء غير متجانسة أو مرتبطة فيما بينها.

بعد ذلك نقوم بحساب المعاملات: β_{0i} للنموذج الأصلي رقم (4)وفق العلاقة التالية:

$$\beta_{0i} = \beta_0 + \beta_i \ \ldots\ldots..\ (6)$$

مع العلم أن: الحد الثابت يتم الحصول عليه عن طريق حساب متوسطات المعاملات β_i المقدرة.

(1) Régis Bourbonnais, Op-cit, P356.

يتم الحصول على نفس النتائج عند استخدام مقدرات داخل الأفراد Within التي تقوم على كتابة المتغيرات التفسيرية والتابعة في الشكل المركز بالنسبة لمتوسطاتها، ثم نقوم بتطبيق طريقة المربعات الصغرى العادية أو طريقة المربعات الصغرى المعممة للنموذج التالي:

$$y_{it} - \bar{y}_i = \beta'(x_{it} - \bar{x}_i) + \varepsilon_{it}; i = 1,2,\dots,N; t = 1,2,\cdots,T$$

بعد تقدير معاملات β'، يتم الحصول على المعاملات الثابتة للأفراد β_{0i} كما يلي[1]:

$$\begin{cases} \hat{\beta}_{01} = \bar{y}_1 - \hat{\beta}'\bar{x}_1 \\ \hat{\beta}_{02} = \bar{y}_2 - \hat{\beta}'\bar{x}_1 \\ \vdots \\ \hat{\beta}_{0N} = \bar{y}_N - \hat{\beta}'\bar{x}_N \end{cases} \dots\dots (7)$$

- **نموذج التأثيرات العشوائية :**

في نموذج التأثيرات الثابتة يكون: $\varepsilon_{it} \sim N(0;\ \sigma^2)$، ولكي تكون معلمات النموذج تتمتع بخاصة عدم التحيز ، لا بد أن يكون[2] تباين الخطأ ثابتاً لجميع المشاهدات، مع عدم وجود ارتباط ذاتي بين الأخطاء العشوائية خلال الزمن، حيث يصبح لدينا في هذا النوع من النماذج الحد الثابت β_{0i} عبارة عن متغير عشوائي وفق العلاقة التالية:

$$\beta_{0i} = \mu + v_i \dots\dots (8)$$

بتعويض المعادلة(8) في النموذج رقم(1):

$$y_{it} = \mu + \beta' x_{it} + \varepsilon_{it} + v_i \dots\dots.. (9)$$

يطلق على نموذج التأثيرات العشوائية تسمية أخرى "نموذج مركبات الخطأ" (Error Components Model)، ويعود ذلك لاحتواء النموذج على حدين للخطأ العشوائي وهما: ε_{it} وv_i.

لنبحث عن المميزات الإحصائية لمجموع ε_{it} و v_i، بعد وضع :

$$w_{it} = \varepsilon_{it} + v_i \dots\dots.. (10)$$

$$\begin{cases} E(w_{it}) = E(\varepsilon_{it} + v_i) = 0 + 0 = 0 \\ Var(w_{it}) = var(\varepsilon_{it}) + var(v_i) = \sigma_\varepsilon^2 + \sigma_v^2 \end{cases} \dots\dots. (11)$$

إن طريقة المربعات الصغرى العادية لا تعطي مقدرات كفؤة، مما يؤثر في اختبار المعلمات، لأن:

---

[1] لم نتعرض إلى الطرق بشكل معمق، وبالأخص من الزاويتين الإحصائية والرياضية، لمزيد من البحث والاطلاع، أنظر:
- Jeffrey M. Wooldridge (2001), **Econometric Analysis of cross section and panel data**, England: Massachusetts Institute of Technology press Cambridge, PP247-332.
- Hsiao, Op-cit, PP31-77.

[2] زكرياء يحي الجمال (2012) ، اختيار النموذج في نماذج البيانات الطولية الثابتة والعشوائية، **المجلة العراقية للعلوم الإحصائية**، المجلد 12 ، العدد 21 ، جامعة الموصل، العراق، ص ص272- 274.

$$cov(w_{it}, w_{is}) = \sigma_v^2 \neq 0; t \neq s$$

للتخلص من هذا الإشكال يتم استخدام طريقة المربعات الصغرى المعممة (Generalized Least Squares, GLS) الذي يمثل المتوسط المرجح لمقدرات ما بين الأفراد (Between) وداخل الأفراد (Within).

إن مقدر ما بين الأفراد نرمز له بالرمز: $\hat{\beta}_{Bet}$ الذي يمثل مقدر المربعات الصغرى العادية والمطبق على متوسطات المتغيرات التابعة والمتغيرات المفسرة:

$$\bar{y}_i = \mu + \beta' \bar{x}_i + \bar{w}_i \text{ .... (12)}$$

يعطى مقدر المربعات الصغرى المعممة كما يلي:

$$\hat{\beta}_{GLS} = \Delta \hat{\beta}_{Bet} - (1 - \Delta) \hat{\beta}_{LSDV} \text{ .... (13)}$$

حيث أن: Δ عبارة عن مصفوفة الأوزان، التي تمثل مقلوب مصفوفة التباين والتباين المشترك لـ $\hat{\beta}_{Be}$.

تتم المفاضلة بين نموذج الآثار العشوائية والنموذج التجميعي باستخدام اختبار Breusch and Pagan (1980) ويكون ذلك من خلال اختبار مضاعف لاقرنج Lagrange multiplier test ، كما يتم المفاضلة بين نموذج الآثار الثابتة ونموذج الآثار العشوائية عن طريق اختبار هوسمان.

المحاضرة السابعة:

اختبارات الاستقرارية في بيانات بانل

المحاضرة السابعة: اختبارات الاستقرارية في بيانات بانل

تم استخدام اختبارات جذر الوحدة في بيانات السلاسل الزمنية خلال سنوات السبعينيات، غير أن تطبيقها على نماذج Panel لم يكن إلا خلال سنوات التسعينيات وكان ذلك على يدي Levin and lin (1992)، أن اختبارات جذر الوحدة يمكن تصنيفها إلى اختبارات الجيل الأول التي درست مشكل الاستقرار في ظل الاستقلالية بين الأفراد ويكون ذلك في حالتي التجانس وعدم التجانس، أما اختبارات الجيل الثاني فتكون في ظل وجود ترابط بين الأفراد، سنحاول أن نتناول بعضا من هذه الاختبارات في النقاط التالية[1]:

- **اختبار Levin-Lin-Chu(LLC, 2002):**

إن النموذج الذي اعتمد عليه الباحثان مستوحى من نموذج Dickey and Fuller(1979, 1981)، مع إضافة البعد الفردي، يأخذ النموذج الصيغة الرياضية التالية[2]:

$$\Delta y_{i,t} = \rho_i y_{i,t-1} + \sum_{L=1}^{p_i} \phi_{il} \Delta y_{i,t-L} + \alpha_{mi} d_{mt} + \varepsilon_{it} \quad \ldots\ldots.. (1);$$
$$i=1,\ldots.N\ ;\ t=1,\ldots..T$$

إن الفرضيات محل الاختبار هي:

H_0: السلسلة غير مستقرة؛

H_1: السلسلة مستقرة.

- **المرحلة الأولى:**

نعمل على تحديد درجات التأخير للأفراد p_i للنموذج (1) وفقا للمعايير المشهورة (AIC , SC).

- **المرحلة الثانية:**

نقوم بإجراء انحدار $\Delta y_{i,t}$ على $\Delta y_{i,t-1}$ و d_{mt} للحصول على $\hat{e}_{it}$؛

نقوم بإجراء انحدار $y_{i,t-1}$ على $\Delta y_{i,t-1}$ و d_{mt} للحصول على $\hat{v}_{i,t-1}$.

- **المرحلة الثالثة:**

نقوم بجعل الأخطاء في الشكل المعياري المتحصل عليه سابقا، كما يلي:

---

[1] لمزيد من التفصيل، يمكنك الرجوع إلى:
Christophe Hurlin, Valérie Mignon (2005), Une synthèse des tests de racine unitaire sur données de panel, **économie et prévision**, 2005/3-4-5, N^0 196-170-171, République Française, PP 253-294

[2] Hsiao, Op-cit, PP386-390.

$$\tilde{e}_{it} = \frac{\hat{e}_{it}}{\hat{\sigma}_{\varepsilon i}} \; ; \; \tilde{v}_{i,t-1} = \frac{\hat{v}_{i,t-1}}{\hat{\sigma}_{\varepsilon i}}$$

حيث أن: $\hat{\sigma}_{\varepsilon}$يمثل الانحراف المعياري للأخطاء المحصل عليها من النموذج (1).

- **المرحلة الرابعة:**

نقوم باستخدام طريقة المربعات الصغرى العادية على تقدير النموذج التالي:

$$\tilde{e}_{it} = \rho \tilde{v}_{i,t-1} + \tilde{\varepsilon}_{it} \; \ldots\ldots \; (2)$$

تكـون فرضيـة العدم $\rho = 0$ ، لاختبـار فرضيـة العدم تم تعديل إحصائية t-student واقتراح جدول آخر من طرف الباحثين levin et alسنة 2002 [(1)] .

إن الشرط الضروري لتطبيق اختبارLLC هو أن تكون:$\sqrt{NT}/\mathrm{T} \longrightarrow 0$، أما الشرط الكافي فهو أن تكون:$\frac{NT}{T} \rightarrow 0$، وفقا للباحثين فإن الاختبار يكون صالحا إذا كانت قيمة N تتراوح ما بين 10 و 250 ، وعدد مشاهدات T ما بين 5 و 250 ، أما إذا كانت Tصغيرة فإن الاختبار يكون ضعيفا.

إن أهم انتقاد وجه لهذا الاختبار هو أن فرضية العدم تنصّ على عدم وجود جذر الوحدة لكل المقاطع العرضية، لكن من الناحية العملية توجد حالات وسطية، ونقصد بذلك أن بعض المقاطع العرضية تحتوي على جذر الوحدة، وأخرى لا.

➢ **اختبار Im, pesaran and shin test(2003)**:

إن اختبار IPS جاء ليكمل اختبار LLC، لأنه يسمح للمعاملات أن تكون غير متجانسة، فهو ليس مقيد، حيث أن فرضية العدم تنصّ على أن جميع الأفراد لديهم جذر وحدة:

$$H_0 : \rho_i = 0; \;\; \forall i$$

الفرضية البديلة تصاغ كما يلي: ليس كل الأفراد لديهم جذر الوحدة:

$$H_1 : \begin{cases} \rho_i < 0; for\ i = 1, \cdots, N_1 \\ \rho_i = 0; for\ i = N_1 + 1, \cdots, N \end{cases}$$

كما أن t_{ρ_i} تمثل إحصائية t-student الفردية من أجل اختبار فرضية العدم $\rho_i = 0$ ، بعد ذلك نقوم بحساب متوسطات اختبارات جذر الوحدة الفردية:

$$\bar{t} = \frac{1}{N} \sum_{i=1}^{N} t_{\rho i}$$

---

(1)Andrew levin, CheinFulin and Chia shangjameschu (2002), Unit root tests in panel data: asymptotic and finite sample properties; **Journal of econometrics**, Vol 108, no 24, North Holland P 14.

هذه الإحصائية تتبع التوزيع الطبيعي المعياري $Z \to N(0;1)$ ، كما أثبتت محاكاة Monto carlo أن اختبارIMP أفضل من LLC في العينات الصغيرة.

➢ **اختبار (Breitung(2000**:

إن المرحلة الأولى من هذا الاختبار هي نفسها التي اجتزناها عند تطبيق اختبار LLC، غير أنها لا تدمج المركبات المحددة.

نقوم بإجراء انحدار $\Delta y_{i,t}$ على $\Delta y_{i,t-1}$ للحصول على البواقي $\hat{e}_{it}$ ، أيضا نجري انحدار $y_{i,t-1}$ على $\Delta y_{i,t-1}$ لنحصل على $\hat{v}_{i,t-1}$ ، بعد ذلك نقوم بتطبيق التحول العمودي إلى الأمام Forward orthogonalization transformation على البواقي $\hat{e}_{it}$ للحصول على e_{it}^*.

أخيرا نقوم بإجراء الانحدار التجميعي التالي:

$$e_{it}^* = \rho v_{i,t-1}^* + \varepsilon_{it}^* \ldots\ldots.. (3)$$

يتم اختبار الفرضيات باستخدام التوزيع الطبيعي المعياري .

➢ **اختبار (Hadri(2000 :**

اعتمد الباحث Hadri في دراسته على اختبار Kwiatkowski-Phillips-Schmidt- Shin test (KPSS) المستعمل في أبحاث السلاسل الزمنية، إضافة إلى مضاعف لاغرنج LM، حيث يقوم هذا الاختبار على دراسة سلسلة البواقي الناتجة من انحدار y_{it} على الحد الثابت أو الحد الثابت مضافا إليه مركبة الاتجاه العام، كما أن الفرضيات المراد اختبارها هي[1]:

فرضية العدم: نموذج بانيل مستقر؛

الفرضية البديلة: نموذج بانيل غير مستقر.

إن النموذجين المقترحين من طرف Hadri هما:

$$y_{it} = r_{it} + \varepsilon_{it} \ \ldots\ldots\ldots\ldots.. (3\text{-}112)$$

$$y_{it} = r_{it} + \beta_i t + \varepsilon_{it} \ \ldots\ldots. (3\text{-}113);$$

[1] Robert Kunst (2011), **Summary based on Chapter 12 of Baltagi: Panel Unit Root Tests**, PhD-Course: Panel Data, University of Vienna, Department of Economics, pp01-09.

كما أن: $r_{it} = r_{i,t-1} + \mu_{it}$ وكل من: $\varepsilon_{it} \rightarrow IIN(0;\ \sigma_{\varepsilon}^2)$ و $\mu_{it} \rightarrow IIN\left(0; \sigma_{\mu}^2\right)$، وباستخدام التعويض الخلفي، فإن:

$$y_{it} = r_{i0} + \beta_i t + \sum_{s=1}^{t} \mu_{is} + \varepsilon_{it}$$

$$y_{it} = r_{i0} + \beta_i t + v_{it} \ldots\ldots (4)$$

حيث أن: $v_{it} = \sum_{s=1}^{t} \mu_{is} + \varepsilon_{it}$، كما أن فرضية الاستقرار تكون عندما : $\sigma_{\mu}^2 = 0$ (فرضية العدم)وإحصائية LM تعطى كما يلي[1]:

$$LM_1 = \frac{1}{N}\left(\left(\sum_{i=1}^{N} \frac{1}{T^2} \sum_{t=1}^{T} S_{it}^2\right) \Big/ \hat{\sigma}_{\varepsilon}^2\right) \ldots\ldots (5)$$

مع العلم أن: $S_{it} = \sum_{s=1}^{t} \hat{\varepsilon}_{it}$ تمثل المجموع الجزئي للبواقي المستخرجة من النموذج (4) ، كما أن $\hat{\sigma}_{\varepsilon}^2$ يمثل مقدر متسق لـ σ_{ε}^2، وفي ظل فرضية العدم فإن: $\hat{\sigma}_{\varepsilon}^2 = \frac{1}{NT} \sum_{i=1}^{N} \sum_{t=1}^{T} \hat{\varepsilon}_{it}^2$.

تم تعديل إحصائية LM_1من طرف(2000) Hadri عند الأخذ بعين الاعتبار مسألة عدم التجانس بين المقاطع العرضية، مع تقريب هذه الإحصائية إلى التوزيع الطبيعي.

---

[1]Badi H. Baltagi, Op-cit, PP 246-247.