مشاكل القياس الاقتصادي

تصنف مشاكل الانحدار تحت خمسة عناوين رئيسية هي[38]:

- **عدم التجانس (hetroskedasticity or unequal variance)**: أي ان تكون قيمة تباين الخطأ العشوائي غير ثابتة.
- **الإرتباط الذاتي (autocorrelation)**: أي أن التباين المشترك بين الأخطاء العشوائية غير مساوي للصفر، ويكون إرتباطها مع بعضها البعض مختلفا عن الصفر.
- **التداخل الخطي (multicolinearity)**: أي عدم إستقلال المتغيرات المستقلة عن بعضها بعضا، وتنشأ هذه الحالة عندما تعتمد بعض أو كل المتغيرات المستقلة على بعضها بعضا بشكل خطي (linear dependence)، وتسمى هذه المشكلة في بعض الأحيان الترابط الخطي المتعدد.
- **خطأ التوصيف (specification error)**: يحدث هذا الخطأ عندما لا يتم تحديد النموذج بطريقة سليمة، نتيجة لحذف متغير أو متغيرات ضرورية، أو أن يشمل النموذج متغير أو متغيرات غير ضرورية، أو أن يأخذ شكلا غير مناسب، رياضيا أو اقتصاديا.
- **الأنية (simultaneity)**: أي أن يتأثر النموذج الموصوف بنموذج أخر، فلا يعرف مسار السببية بين المتغير (المتغيرات) في الطرف الأيمن والمتغير (الظاهرة) في الطرف الأيسر من الاقتران، أو أن توجد أثار أنية من أي نوع، أو أن تتأثر قيمة المتغير المستقل بالخطأ العشوائي، مما يترتب على ذلك انتهاكا لفرضية ثبات ذات المتغير.

سوف نقوم فيما يلي بتحليل، أسباب وعرض الاختبارات الكشف والحلول المقترحة لمشكلة عدم تجانس التباين، الارتباط الخطي بين الأخطاء، والتعدد الخطي.

[38]– عبدالرزاق بني هاني، **الاقتصاد القياسي نظرية الانحدار البسيط والمتعدد**، الطبعة الأولى، دار وائل، الأردن،2014 ،ص136.

مشكل الارتباط الذاتي بين الأخطاء

أولا: مفهوم الارتباط الخطي بين الأخطاء

تحدث مشكلة الارتباط الذاتي بين الأخطاء عندما ترتبط البواقي العشوائية مع بعضها البعض أي أن قيمة المتغير العشوائي ε_t في السنة t يرتبط مع قيمته السابقة ε_{t-1} أو مع قيمته اللاحقة ε_{t+1} وأن هذا الارتباط يحدث بسبب العلاقات الاقتصادية بين الظاهرة المدروسة[39]، مما يخرق أحد الفروض المحيطة بالمتغير العشوائي:

استقلالية الأخطاء إذا كان $t \neq t'$ فإن $E(\varepsilon_t, \varepsilon_{t'}) = 0$ أي أن التباين المشترك ما بين المتغيرات العشوائية يجب أن تساوي صفرا.

ثانيا: أسباب ظهور الارتباط الذاتي بين الأخطاء

يظهر الارتباط الذاتي بين الأخطاء للأسباب التالية[40]:

1- الأثار الممتدة لبيانات السلاسل الزمنية : إن بعض العوامل العشوائية الطارئة وغير المتكررة قد ينتج عنها ترابط في قيم العنصر العشوائي ε_t لأكثر من فترة زمنية واحدة، فالحروب والفيضانات والزلازل تمتد بأثارها على فعالية الاقتصاد لعدة سنوات متتالية، مما يتسبب في حصول إرتباط ذاتي بين قيم ε_t المتلاحقة، حيث أن القيم الحالية تتأثر بالقيم الأخرى للفترات السابقة.

2- حذف بعض المتغيرات المستقلة من العلاقة المدروسة لسبب أو لأخر: مثل عدم توفر البيانات المناسبة عنها أو لغرض تبسيط هيكل النموذج، وقد يكون من بين هذه المتغيرات المحذوفة متغير أو أكثر مرتبط ذاتيا، الأمر الذي يؤدي إلى جعل العنصر العشوائي يتضمن تلك المتغيرات المرتبطه، ومن ثم فإن ε_t لا يعكس الخطأ العشوائي في النموذج فحسب، إنما يعكس أيضا المتغيرات المحذوفة.

3- معالجة البيانات: تجرى على البيانات المنشورة أحيانا عمليات تشذيب وقد يتم تقدير قيم بعض المشاهدات اعتمادا على قيم مشاهدات أخرى، ذلك أن عمليات التشذيب والتقدير تعتمد في العادة على أخذ معدلات قيم المشاهدات المتتالية، مما يخلق علاقة ما بين أخطاء تلك المشاهدات وبالتالي التأثير على طبيعة توزيعها.

---

[39]- عدنان داود العذاري، صادق على الجبوري، **الاقتصاد القياسي نظرية وحلول تطبيق باستخدام Minitab, Relase 14**، الطبعة الأولى، دار جرير، الأردن، 2010، ص 99.

[40]- حسين علي بخيت، سحر فتح الله، مرجع سبق ذكره، ص189.

4- الصياغة غير الدقيقة للنموذج: بمعنى أن شكل العلاقة الدالية المستخدمة لا يتطابق مع الشكل الحقيقي، فإذا افترضنا علاقة خطية بين المتغيرين y و x، في حين أن العلاقة الحقيقية غير خطية فإنه بمكن أن ينتج عن ذلك ترابط ذاتي في العنصر العشوائي.

ثالثا: النتائج المترتبة على وجود الارتباط الذاتي

تبقى مقدرات طريقة المربعات الصغرى العادية (OLS) تتسم بالخطية وعدم التحيز، إلا أنها تفقد خاصية أفضل وأصغر تباين (التباين بوجود الارتباط الذاتي يفوق التباين في غياب الارتباط الذاتي)، كما يؤثر الارتباط الذاتي على نتائج تحليل الانحدار فتعطي الاختبارات t ، F نتائج أقل دقة من تلك في حالة عدم وجوده، حيث تستند تلك الاختبارات على التباين والخطاء المعياري وعدم دقة التنبؤات المتحصلة عليها بطريقة OLS، مما يتطلب استخدام طرق أخرى للتقدير كطريقة المربعات الصغرى المعممة (GLS) Generalized Least Squares Method لأنها تعطي أفضل تقدير خطي غير متحيز في حالة وجود الارتباط الذاتي شريطة أن تكون قيمة (ρ) معلومة[41].

رابعا: اختبارات الكشف عن الارتباط الذاتي بين الأخطاء

1- اختبار Durbin-Watson:

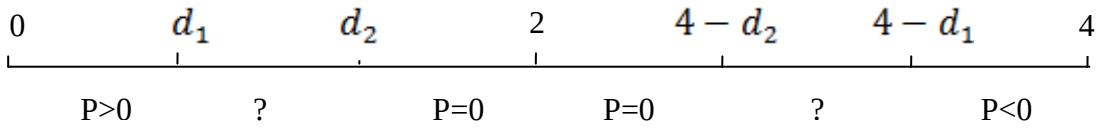
هذا الاختبار يخص الارتباط الذاتي من الدرجة الأولى، و تأخذ الفرضية الشكل التالي:

$$\begin{cases} H_0: \rho = 0 \\ H_1: \rho \neq 0 \end{cases}$$

نقوم بحساب الإحصائية DW التي تعطى بالعلاقة التالية: $DW = \frac{\sum_{t=2}^{n}(\varepsilon_t - \varepsilon_{t-1})}{\sum_{t=1}^{n}\varepsilon_t^2}$.

بعد حساب DW نقارنها مع القيمتين المجدولتين، d_1 التي تمثل الحد الأدنى لانعدام الارتباط الذاتي، و d_2 التي تمثل الحد الأقصى لانعدام الارتباط الذاتي، و ذلك حسب عدد المشاهدات n ، و عدد المتغيرات التفسيرية في النموذج عند درجة معنوية 5%.

و يتم قبول و رفض الفرضيتين حسب المخطط التالي:



قيمة DW الوسطية هي 2 و عندها ينعدم الارتباط الذاتي، أي: P=0.

[41]- حسين علي بخيت، سحر فتح الله، مرجع سبق ذكره، ص 198.

ويتم قبول و رفض H_0 حسب الحالات التالية:

$0 < DW < d_1$ وجود ارتباط ذاتي موجب.

$d_1 < DW < d_2$ مجال غير محسوم، هناك شك في وجود أو عدم وجود ارتباط ذاتي.

$d_2 < DW < 4 - d_2$ عدم وجود ارتباط ذاتي.

$4 - d_2 < DW < 4 - d_1$ مجال غير محسوم، هناك شك في وجود أو عدم وجود ارتباط ذاتي.

$4 - d_1 < DW < 4$ وجود ارتباط ذاتي سالب.

سلبيات اختبار دربن واتسون:

لا يمكن استعمال هذا الاختبار إلا في ظل الشروط التالية:

- يجب أن يكون النموذج متضمنا للمعلم الثابت b_0 .
- النموذج المقدر لا يتضمن متغيرات تابعة ذات فترات إبطاء كمتغيرات مستقلة.
- لا يختبر دربن واتسون إلا الارتباط الذاتي من الدرجة الأولى.
- لا يمكن اتخاذ القرار بوجود أو عدم وجود ارتباط ذاتي بين الأخطاء في وقوع إحصائية دربن واتسون $\widehat{DW}$ في منطقة الشك.

2- اختبار Breusch-Godfrey

يرتكز هذا الاختبار على مضاعف لاغرنج والذي يسمح بإختبار وجود ارتباط ذاتي من درجة أكبر من الواحد. نموذج الانحدار الذاتي للأخطاء من الدرجة p يكتب على الشكل التالي[42]:

$$\varepsilon_t = \rho_t \varepsilon_{t-1} + \rho_2 \varepsilon_{t-2} + \cdots + \rho_p \varepsilon_{t-p} + \mu_t$$

ليكن النموذج العام حيث أن الأخطاء مرتبطة ذاتيا:

$$Y_t = B_0 + B_1 X_{t1} + \cdots + B_k X_{tk} + \rho_t \varepsilon_{t-1} + \rho_2 \varepsilon_{t-2} + \cdots + \rho_p \varepsilon_{t-p} + \mu_t$$

هناك ثلاث خطوات لإجراء هذا الإختبار:

- تقدير النموذج العام بطريقة المربعات الصغرى ثم حساب البواقي $\hat{\varepsilon}_t$
- تقدير المعادلة الوسيطية التالية:

$$\hat{\varepsilon}_t = B_0 + B_1 X_{t1} + \cdots + B_k X_{tk} + \rho_t \hat{\varepsilon}_{t-1} + \rho_2 \hat{\varepsilon}_{t-2} + \cdots + \rho_p \hat{\varepsilon}_{t-p} + \mu_t$$ -

[42]- شيخي محمد، مرجع سبق ذكره، ص 100.

ثم حساب معامل التحديد الخاص بهذه المعادلة R^2. نذكر أن بإستعمال هذه المعادلة، سنفقد p مشاهدة.

- فرضية استقلالية الأخطاء H_0 التي ينبغي إختبارها هي:

$$H_0 = \rho_1 = \rho_2 = \cdots = \rho_P = 0$$

الإحصائية $LM = (n-p) \times^2$ تتبع χ^2 بدرجة حرية p . إذا كان $(n-) \times R^2$ أكبر من $\chi^2(p)$ (القيمة الحرجة لتوزيع χ^2 بنسبة معنوية α)، فإننا نرفض H_0 فرضية استقلالية الأخطاء.

خامسا: الحلول المقترحة للارتباط الذاتي

إن حل كل حالة يعتمد على مصدر الارتباط الذاتي[43]:

فإذا كان المصدر هو حذف متغيرات أساسية من النموذج، يكون الحل بإعادة إدخال هذه المتغيرات في مجموعة المتغيرات المستقلة. وتكون أبسط طريقة لإكتشاف ما إذا كان سبب الارتباط الذاتي هو حذف متغيرات من النموذج بواسطة تقدير بواقي المربعات الصغرى $\hat{\varepsilon}_t$ ضد المتغيرات التي تكون مسبقا لها علاقة بالمتغيرات المستقلة للظاهرة الموجودة تحت الدراسة.

عندما نكتشف ظاهرة الارتباط الذاتي (بإجراء الاختبارات المناسبة) يمكن أن نصحح النموذج وذلك بالحصول على مقدر ρ بواسطة تطبيق المربعات الصغرى العادية لمجموعة البيانات المحولة.

إن تحويل البيانات الأصلية يعتمد على الشكل أو البناء الهيكلي الأصلي للارتباط الذاتي. حيث إذا كان الارتباط الذاتي الذاتي من المرتبة الأولى AR(1)، يكون التحويل بواسطة طرح النموذج المؤخر بفترة واحدة ومضروبا بمقدر المعلمة ρ أي $\hat{\rho}$ من النموذج الأصلي أي:

$$Y_t = \beta_1 + \beta_2 X_{2t} + \cdots\cdots + \beta_k X_{kt} + \varepsilon_t$$

$$\varepsilon_t = \rho \varepsilon_{t-1} + \mu_t$$

$$Y_{t-1} = \beta_1 + \beta_2 X_{2t-1} + \cdots\cdots + \beta_k X_{kt-1} + \varepsilon_{t-1}$$

ثم نضرب المعادلة الأخيرة ب $\hat{\rho}$ ثم نطرحها من المعادلة الممثلة ل Y_t لنجد:

$$Y_t - \hat{\rho} Y_{t-1} = (\beta_1 - \hat{\rho}\beta_1) + \beta_2 (X_{2t} - \hat{\rho} X_{2t-1}) + \cdots\cdots + \beta_k (X_{kt} - \hat{\rho} X_{kt-1}) + \mu_t$$

$$Y_t^* = \beta_1^* + \beta_2 X_{2t}^* + \cdots\cdots + \beta_k X_{kt}^* + \mu_t$$

[43]– تومي صالح، **مدخل لنظرية القياس الاقتصادي**، الجزء الثاني، الطبعة الثانية، ديوان المطبوعات الجامعية، الجزائر،2010، ص69–ص71.

حيث أن:

$$\beta_1^* = \beta_1(1-\hat{\rho})$$
$$X_{jt}^* = X_{jt} - \hat{\rho}X_{jt-1} \qquad j = 2,3,\cdots\cdots,k$$

ومنه إذا عرفنا قيمة $\hat{\rho}$ نستطيع تطبيق قانون المربعات الصغرى العادية على النموذج المحول. كما نلاحظ أن النموذج المحول يحتوى على (n−1) ملاحظات بسبب سقوط الملاحظة الأولى.

ولتحاشي هذا الضياع في الملاحظات فإن بعض الإحصائيين أمثال Prais-Winsten، يقترحون كتابة الشكل المحول للملاحظات الأولى كما يلي:

$$Y_1^* = \sqrt{1-\hat{\rho}^2}.Y_1$$
$$X_1^* = \sqrt{1-\hat{\rho}^2}.X_1$$

ويلاحظ بعض الكتاب أنه إذا كانت إحصائية DW أقل من معامل التحديد المضاعف أي $DW < R^2$، فإنه يمكن إستعمال الفروقات الأولى في ملاحظات متغيرات النموذج. حيث نقدر الفرق الأول $(Y_t - Y_{t-1})$ في مجموعة الفروقات للمتغيرات المستقلة $(X_t - X_{t-1})$، وذلك تحت الفرضية القائلة بأن الفروقات الأولى للأخطاء $(\varepsilon_t - \varepsilon_{t-1})$ تكون مستقلة.

تمرين حول مشكل الارتباط الذاتي للأخطاء

لتكن لدينا البيانات التالية الموضحة في الجدول التالي:

السنوات	1981	1982	1983	1984	1985	1986	1987	1988	1989	1990
Y	87.4	97.9	96.7	98.2	99.8	100.5	103.2	107.8	96.6	88.9
X1	98.6	101.2	102.4	100.9	102.3	101.5	101.6	101.6	99.8	100.3
X2	99.1	99.1	98.9	110.8	108.2	105.6	109.8	108.7	100.6	81
X3	108.5	110.1	110.4	104.3	107.2	105.8	107.8	103.4	102.7	104.1
السنوات	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000
Y	75.1	76.9	84.6	90.6	103.1	105.1	96.4	104.4	110.7	127.1
X1	97.6	97.2	97.3	96	99.2	100.3	100.3	104.1	105.3	107.6
X2	68.6	70.9	81.4	102.3	105	110.5	92.5	89.3	93	106.6
X3	99.2	99.7	102	94.3	97.7	101.1	102.3	104.4	108.5	111.3

المطلوب:

1- تقدير معالم النموذج ؟

2- اختبار وجود الارتباط الذاتي بين الأخطاء :أ- اختبار دربن واتسون ؟

ب- اختبار Breusch –Godfrey ؟

3- صحح النموذج من مشكل الارتباط الذاتي للأخطاء إن وجد ؟

الحل:

1- تقدير معالم النموذج

Dependent Variable: Y
Method: Least Squares
Date: 01/25/20 Time: 10:39
Sample: 1 20
Included observations: 20

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-242.7146	26.93755	-9.010268	0.0000
X1	3.887240	0.402380	9.660612	0.0000
X2	0.404491	0.061666	6.559355	0.0000
X3	-0.869804	0.241450	-3.602423	0.0024

R-squared	0.938269	Mean dependent var	97.55000
Adjusted R-squared	0.926694	S.D. dependent var	11.83076
S.E. of regression	3.203188	Akaike info criterion	5.343027
Sum squared resid	164.1666	Schwarz criterion	5.542173
Log likelihood	-49.43027	Hannan-Quinn criter.	5.381902
F-statistic	81.06246	Durbin-Watson stat	1.060806
Prob(F-statistic)	0.000000		

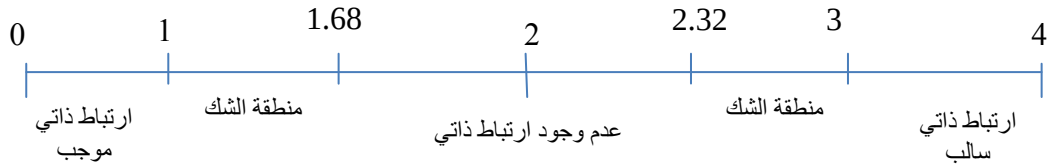
من الجدول نلاحظ أن معالم النموذج (معاملات المتغيرات المفسرة) لها معنوية إحصائية خلافا للصفر($t_{16}^{0.05} = 2.12$)،و قيمة فيشر المحسوبة أكبر من (F(3, 16, 0.05)=3.24).

2- اختبار وجود الارتباط الذاتي للأخطاء:

أ- اختبار دربن واتسون

من الجدول نلاحظ أن إحصائية دربن واتسون المحسوبة DW= 1.0608

لدينا أن K=3 و N=20 من جدول دربن واتسون نجد أن d1= 1.00 و d2=1.68



وبالتالي فإن DW=1.0608 محصورة بين d1= 1.00 و d2=1.68أي نحن في منطقة الشك لا يمكن اتخاذ قرار بوجود أو عدم مشكل الارتباط الذاتي بين الأخطاء.

ب-اختبار Breusch –Godfrey

Breusch-Godfrey Serial Correlation LM Test:

F-statistic	2.782095	Prob. F(2,14)	0.0961
Obs*R-squared	5.688137	Prob. Chi-Square(2)	0.0582

Test Equation:
Dependent Variable: RESID
Method: Least Squares
Date: 01/25/20 Time: 10:43
Sample: 1 20
Included observations: 20
Presample missing value lagged residuals set to zero.

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-8.632449	24.80630	-0.347994	0.7330
X1	0.101819	0.370069	0.275135	0.7872
X2	-0.025021	0.056883	-0.439867	0.6667
X3	0.008226	0.231046	0.035601	0.9721
RESID(-1)	0.653275	0.282721	2.310670	0.0366
RESID(-2)	-0.383934	0.290205	-1.322976	0.2070

R-squared	0.284407	Mean dependent var	-2.35E-15
Adjusted R-squared	0.028838	S.D. dependent var	2.939447
S.E. of regression	2.896753	Akaike info criterion	5.208383
Sum squared resid	117.4765	Schwarz criterion	5.507103
Log likelihood	-46.08383	Hannan-Quinn criter.	5.266697
F-statistic	1.112838	Durbin-Watson stat	2.011463
Prob(F-statistic)	0.397368		

من خلال الجدول نلاحظ أن $R^2=0.289762$ و لدينا $n=20$:

نحسب إحصائية مضاعف لاغرنج

$LM=n.R^2=20X0.284407=5.688137$

نستخرج إحصائية كي دو بدرجة حرية 2 من جدول كي دو

$$\aleph^2_{0.05}(2) = 5.99$$

Breusch-Godfrey Serial Correlation LM Test:

F-statistic	2.782095	Prob. F(2,14)	0.0961
Obs*R-squared	5.688137	Prob. Chi-Square(2)	0.0582

المقارنة :

نلاحظ أن

$\aleph^2_{0.05}(2) = 5.99$ أكبرمن $LM=5.795238$ وبالتالي نقبل الفرضية Ho أي عدم وجود مشكل الإرتباط ذاتي للإخطاء

3- في هذه الحالة لا يوجد مشكل الارتباط الذاتي بين الأخطاء لكن سوف نوضح كيفية تصحيح النموذج مشكل الارتباط الذاتي للأخطاء في حالة وجوده وذلك بتحويل المتغيرات:

نحسب raul :

$$\hat{\rho}_1 = 1 - \frac{DW}{2} = 1 - \frac{1.060}{2} = 0.470$$

نحول المتغيرات:

$$dy = (y_t - \hat{\rho}_1 * y_{t-1}) = (y_t - 0.47 * y_{t-1})$$
$$dx_1 = (x_{1,t} - \hat{\rho}_1 * x_{1,t-1}) = (x_{1,t} - 0.47 * x_{1,t-1})$$
$$dx_2 = (x_{2,t} - \hat{\rho}_1 * x_{2,t-1}) = (x_{2,t} - 0.47 * x_{2,t-1})$$
$$dx_3 = (x_{3,t} - \hat{\rho}_1 * x_{3,t-1}) = (x_{3,t} - 0.47 * x_{3,t-1})$$

نحسب قيمة المشاهدة الأولى للمتغيرات

$$Y^*_{1981} = \sqrt{1-\hat{\rho}^2}.Y_{1981} = \sqrt{1-(0.47)^2}.(87.4) = 68.093$$
$$x^*_{1,1981} = \sqrt{1-\hat{\rho}^2}.x_{1,1981} = \sqrt{1-(0.47)^2}.(98.6) = 76.819$$
$$x^*_{2,1981} = \sqrt{1-\hat{\rho}^2}.x_{2,1981} = \sqrt{1-(0.47)^2}.(99.1) = 77.208$$
$$x^*_{3,1981} = \sqrt{1-\hat{\rho}^2}.x_{3,1981} = \sqrt{1-(0.47)^2}.(108.5) = 84.532$$

ثم نقدر النموذج الجديد

Dependent Variable: DY
Method: Least Squares
Date: 01/25/20 Time: 11:38
Sample: 1 20
Included observations: 20

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-1.478704	17.84695	-0.082855	0.9350
DX1	0.795992	1.058123	0.752269	0.4628
DX2	0.484086	0.159842	3.028522	0.0080
DX3	-0.247851	0.753566	-0.328905	0.7465

R-squared	0.621326	Mean dependent var	53.72300
Adjusted R-squared	0.550324	S.D. dependent var	9.386202
S.E. of regression	6.294187	Akaike info criterion	6.693987
Sum squared resid	633.8686	Schwarz criterion	6.893133
Log likelihood	-62.93987	Hannan-Quinn criter.	6.732862
F-statistic	8.750889	Durbin-Watson stat	1.096600
Prob(F-statistic)	0.001149		

ثم حساب الثابت:

$$\hat{b}_0^{\,*} = {\hat{b}_0}/{(1-\hat{\rho}_1)} = {-1.478704}/{(1-0.47)} = -2,790007$$

مشكل التعدد الخطي (التداخل الخطي)

أولا: مفهوم التداخل الخطي المتعدد

التداخل الخطي مصطلح معرب من الأصل الإنجليزي (Multicolinearty) وهو مصطلح مركب من ثلاثة مقاطع

الأول: Multi ومعناه تعدد أو متعدد.

الثاني: Co وتعني التلازم أو التبادل أو الاشتراك أو التداخل.

الثالث: Linearity وتعني الخطية.

وقد اختلفت الترجمات العربية له، فمنهم من يسميه (الارتباط الخطي المتعدد). وهو مفهوم قد يخلق خلطا له مع الارتباط الخطي المتعدد (Multiple Correlation) ولهذا سنستبعده من الاستخدام. ومنهم من يترجمه ب(تعدد العلاقات الخطية) وهو يفقد الوضوح والتحديد. ومنهم من يسميه (عدم استقلالية المتغيرات مسبقا عن بعضها). وأخرون يترجمونه إلى (الامتداد الخطي المتعدد). ويترجمه مؤلف أخر ب (الاشتراك الخطي) أو (تعدد العلاقات الخطية) وغيرهما من الترجمات.

أما أقرب ترجمة لهما ضمن المفهوم العام فهي المصطلحات الأتية[44]:

التزامن– التلازم الخطي – الخطية المتلازمة للمتغيرات المفسرة – (الترابط الخطي) – التشابك الخطي المتعدد أو التداخل الخطي المتعدد.

ويعتبر النرويجي (راكنرفريش) أول من اكتشف وجود هذه الظاهرة الإحصائية وذلك في معرض تحليله لبيانات السلاسل الزمنية. فقد اتضح له وجود تداخل أو ترابط خطي متبادل بين المتغيرات الاقتصادية التفسيرية بسبب التطور المتناغم أو المتوافق لهذه المتغيرات عبر الزمن أي أنها تزداد وتنخفض بوتيرة واحدة أو متقاربة عبر الزمن مما يجعل أحدهما غير ذو مقدرة تفسيرية في النموذج ويجعل بالتالي معلمة او أكثر من المعلمات الإنحدارية أو المقطعية مساوي للصفر أو مقاربة له، أو تتساوى المعلمتان الانحداريتان مما يجعل معامل الارتباط الجزئي بين المتغيرين المستقلين مساويا للواحد عدد صحيح.

والمشكلة هنا هي بين المتغيرات التفسيرية ذاتها وليس بين المتغير التفسيري والمتغير التابع.

وهذا يعني أن المشكلة تظهر فقط عندما يكون هناك أكثر من متغير تفسيري واحد مثل X و Z و W.

ولهذا فإن هذه المشكلة غير موجودة في النموذج الخطي البسيط بل في النموذج الخطي العام والمتعدد.

[44]– وليد إسماعيل السيفو، فيصل مفتاح شلوف، صائب جواد إبراهيم جواد، مرجع سبق ذكره، ص 93.

ثانيا: أسباب ظهور التداخل الخطي المتعدد

هناك عدة أسباب لتطور مشكلة التلازم الخطي أو التعدد الخطي بين المتغيرات المستقلة وخاصة في السلاسل الزمنية وأهمها:

1- اشتراك بعض المتغيرات المستقلة في التحرك باتجاه زمني واحد ودون تخلف زمني.

2- التغير المتداخل لعدم جمع بيانات كافية ومن عينات أكثر.

3- التحرك بإتجاه واحد أو متعاكس بمعدل متزامن أو واحد ولنفس الفترة الزمنية.

4- استخدام المتغيرات المتخلفة زمنيا مثل استخدام دخل سنة (t) المعتمد على دخل سنة (t-1).

5- عدم إمكانية التحكم ببيانات المشاهدات لأنها لا تخضع للسيطرة والتجارب.

ثالثا: مؤشرات وجود التداخل الخطي

هناك بعض المؤشرات لوجود تداخل خطي بين المتغيرات المستقلة ومن أهمها الأتي:

1- ظهور قيمة واحدة أو أكثر من قيم المعلمات الانحدارية والثوابت مساوية للصفر.

2- تضخم قيمة الخطأ المعياري للتقدير وقد يصل إلى ما لا نهاية.

3- انعدام المعنوية الإحصائية لإحدى المعلمات المقدرة او أكثر من واحدة رغم أن (R^2) يكون عاليا، إضافة إلى أنها غير متحيزة.

4- ظهور ارتباط جزئي تام بين متغيرين تفسيريين أو أكثر، أو بينه وبين المتغير التابع.

5- تكون قيم بعض المعلمات الانحدارية والثوابت غير صفرية لكنها متساوية القيم مما يجعل أحد المتغيرية التفسيريين فائضا وليس بذي أهمية في النموذج لعدم إمكان عزل تأثير أحدهما على الأخر، لهذا فإن القدرة التفسيرية للنموذج والمتغير تنخفض.

6- عدم قبول الحاسب الإلكتروني للبيانات أو عدم احتساب قيم الثوابت.

رابعا: اختبارات الكشف عن التداخل الخطي المتعدد

إن مشكلة التداخل الخطي وتأثيره تظهر في درجة الترابط المتبادل (R_{xz}) بين المتغيرات المفسرة، وكذلك معامل الارتباط المتعدد الكلي للنموذج (Ryxz). ويعتقد الكثيرون أن المعامل الأخير يمكن أن يستخدم لاختبار وجود تداخل خطي بين المتغيرات المفسرة.

غير ان الممارسة العملية قد أثبتت أن استخدام أي من هذه المتغيرات لوحده كمعيار لوجود التداخل الخطي المتعدد بين المتغيرات المفسرة هو غير كاف للأسباب الأتية:

1- أن وجود قيمة عالية للخطأ المعياري للتقدير لا تظهر دائما مع التداخل الخطي، بدليل أن هناك نماذج اقتصادية مأخوذة من عينات كثيرة مثل دالة (كوب- دوغلاس) تظهر فيها أخطاء معيارية كبيرة.

2- أن R^2 يمكن أن يكون عاليا لكن النتائج لا تكون دقيقة ولا معنوية وقد تحصل على إشارات خاطئة أو خطأ معياري عال.

مع ذلك فإن استخدام مثل هذه المؤشرات يمكن أن ينفع في الإحساس بهذه المشكلة والكشف عن وجوده. ولهذا فإننا نستخدم هذه الطرق إضافة إلى طرق أخرى مبتكرة لكتاب قياسيين معتمدين.

وفيما يأتي عرض لأهم الاختبارات الخاصة بالكشف عن هذه العلاقة.

أولا: طريقة (فريش) Frich's Confluence Analysis

تقوم طريقة فريش على الأتي:

1- إجراء انحدار للمتغير التابع على كل متغير تفسيري على حدى. بهذه الطريقة نحصل على الانحدار الأساسي ونفحصها في ضوء معلومات مسبقة Apriori ومعايير إحصائية.

2- نختار النتائج الانحدارية التي تعطي نتائج أفضل في ضوء المعايير أعلاه.

3- ندخل بالتدريج متغير تفسيري بعد أخر على أفضل انحدار في ضوء ما جاء في (2) ونفحص أثره على المعاملات الارتباطية الفردية وعلى الأخطاء المعيارية وعلى (R^2).

4- وعلى أساس ذلك يمكننا تصنيف المتغيرات التفسيرية على أنها متغيرات تحديدية أو فائضة في النموذج، أي ليست لها أهمية قياسية واقتصادية وعلى الأسس الأتية:

أ- إذا ما حسن المتغير التفسيري المضاف (R^2) دون أن يؤثر في دقة المعلمات بحيث يحدث فيها أخطاء غير مقبولة في ضوء المعلومات النظرية الاقتصادية، عند ذلك سيكون المتغير هذا مفيد في التحليل القياسي.

ب- عندما لا يغير التغير التفسيري المضاف القدرة التفسرية للنموذج و(R^2) كما لا يؤثر في قيمة المعلمات إيجابيا سيعتبر هذا المتغير زائدا او فائضا ولا يدخل ضمن المتغيرات التفسيرية للنموذج.

ج- إذا ما أثر إدخال المتغير التفسيري الجديد على قيمة معلمات النموذج وحول إشاراتها لتكون مقبولة اقتصاديا عند ذلك سيكون المتغير تحديديا.

د- إذا ما أثر إدخال المتغير التفسيري على قيمة معلمات النموذج، وحول إشارتها بإتجاه يحولها إلى قيم غير مقبولة اقتصاديا عند ذلك سيكون لدينا مشكلة تداخل خطي متعدد ذات اهمية جدية.

ولهذا ولأجل ألا نهمل او نرمي خارجا متغير تفسيري تحديدي علينا أن نلجأ إلى حلول لهذه المشكلة بعد تحديدها، بخلافه فإن تأثير هذا المتغير سيمتص من قبل المتغيرات التفسيرية والخطأ العشوائي.

ثانيا: طريقة فارارا- كلاوبير للإختبار Glauber Test- The Farrar

وهذه الطريقة مؤلفة من ثلاثة اختبارات وهي:

1- اختبار مربع كاي Square-Chi للكشف عن وجود درجة شدة التداخل الخطي المتعدد من خلال اختبار الاستقلالية الإحصائية للمتغيرات المستخدمة في العينة.

2- اختبار (F) لتحديد المتغيرات المرتبطة خطيا.

3- اختبار إيجاد نموذج التداخل الخطي المتعدد وتحديد المتغيرات المسؤولة عن التداخل الخطي المتعدد.

4- اختبار (t) لتحديد المتغيرات المسببة للتداخل الخطي.

ويحدد فارار الأتي:

إذا كان التداخل الخطي المتعدد تاما فإن المعلمات ستكون غير محددة. ويمكن قياس الارتباط بين المتغيرات التفسيرية من خلال معاملات الارتباط الجزئية.

خامسا: حلول التداخل الخطي المتعدد

رغم ان التداخل الخطي المتعدد يعتبر عند البعض مشكلة خطيرة إلا أنها عند الأخرين لا تؤلف مشكلة ويمكن التعايش معها وخاصة عندما لا يؤثر بشكل خطير على المعلمات المقدرة.

ومع ذلك يفضل الكثير من الكتاب إيجاد حلول لهذه المشكلة والتي يمكن ان تتمثل بالأتي:

1- حذف المتغير التفسيري المسبب لهذه الظاهرة إن كان تأثيره في القدرة التفسيرية للنموذج ليست بذات أهمية وخاصة تلك التي تكون معلماتها مقاربة للصفر أو مساوية له حتى لا يقع في مطب التوصيف بعدم إدخال المتغيرات المطلوبة وأن تصاب المعلمات المقدرة الأخرى بالتحييز. والمعيار هو هل سيساوي تخفيض التباين بالمقارنة مع زيادة تحييز المعلمات؟ وقد يفضل بعض الكتاب معيار (OLS) ليس أصغر تباين بل (وسط مربع الخطأ) أي S بدلا من σ.

2- قد تصيب هذه الظاهرة بعض المعلمات ولا تصيب الأخرى بالتحيز. فيمكن في هذه الحالة استخدام غير المصابة لأغراض التحليل والتنبؤ ووضع السياسة. وأن استخدمت للتنبؤ يفضل أن يتواصل النموذج من الماضي إلى المستقبل.

3- قد يكون صغر العينة سببا في ظهور هذه المشكلة عند ذلك يصار إلى زيادة حجم العينة بالحصول على بيانات أكثر لا تتصف بالتداخل الخطي.

4- استخدام معلومات معطاة مسبقا عن المعلمات مثل الميل الحدي للإستهلاك أي استخدام معلومات خارجية تدمج في النموذج.

5- تحويل بعض المتغيرات بقسمتها على (k) مثلا أو (z) أو ضرب المتغيرات والمعلمات في (yw) وقسمتها على (z).

6- استخدام طرق مختلفة لتكوين معلومات مقطعية خارجية مثل، طريقة المربعات الصغرى المقيدة، وتعديل (دربن) للمربعات الصغرى العامة، وطريقة التقدير المختلط التي اقترحها (ثايل وكولد بركر).

7- إحلال المتغيرات المتخلفة زمنيا بدل الأنية.

8- إدخال معاملات إضافية للنموذج كقلبها إلى معادلة أنية.

تمرين حول مشكل التعدد الخطي

لتكن لدينا البيانات الموضحة في الجدول التالي:

annee	1964	1965	1966	1967	1968	1969	1970	1971
y	28.4	32	37.7	40.6	47.7	52.9	58.5	64
X1	635.7	688.1	753	796.3	868.5	935.5	982.4	1063.4
X2	92.9	94.5	97.2	100	104.2	109.8	116.3	121.3
annee	1972	1973	1974	1975	1976	1977	1978	1979
y	75.9	94.4	131.9	126.9	155.4	185.8	217.5	260.9
X1	1171.1	1306.6	1412.9	1528.8	1702.2	1899.5	2127.6	2368.5
X2	125.3	133.1	147.7	161.2	170.5	181.5	195.4	217.4

المطلوب :

1- اختبر وجود مشكل التعدد الخطي بتطبيق الاختبارات التالية:

أ- اختبار كلاين klien ؟

ب-اختبار فارار كلاوبير Fara klauber ؟

الحل:

أ- اختبار كلاين klien:

1- تقدير النموذج

Dependent Variable: Y
Method: Least Squares
Date: 05/12/14 Time: 19:36
Sample: 1964 1979
Included observations: 16

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-101.4885	33.08031	-3.067943	0.0090
X1	0.078534	0.055958	1.403450	0.1839
X2	0.758554	0.761246	0.996464	0.3372

R-squared	0.987366	Mean dependent var	100.6563
Adjusted R-squared	0.985422	S.D. dependent var	71.92798
S.E. of regression	8.684477	Akaike info criterion	7.328312
Sum squared resid	980.4618	Schwarz criterion	7.473172
Log likelihood	-55.62650	F-statistic	507.9814
Durbin-Watson stat	1.186381	Prob(F-statistic)	0.000000

2- - حساب معاملات الارتباط البسيط بين المتغيرات:

Correlation Matrix			
	X1	X2	
X1	1.000000	0.997165	
X2	0.997165	1.000000	

من مصفوفة الارتباط نجد أن: $r^2_{x1\,x2} = (0.997165)^2 = 0.994338$

نلاحظ أن معامل الارتباط بين المتغيرات المفسرة أكبر من معامل التحديد $R^2 = 0.987366$ وهذا يدل على وجود مشكل التعدد الخطي

ب-اختبار فارار كلاوبير Fara klauber:

1-حساب محدد مصفوفة الارتباط:

$$D = \begin{vmatrix} 1 & r_{x1\,x2} \\ r_{x1\,x2} & 1 \end{vmatrix} = \begin{vmatrix} 1 & 0.997165 \\ 0.997165 & 1 \end{vmatrix} = 1 - 0.994328 = 0.005662$$

2-حساب إحصائية كي دو:

$$\aleph^2 = -\left[n - 1 - \frac{1}{6}(2K + 7)\right] \times LnD$$

$$\aleph^2 = -\left[16 - 1 - \frac{1}{6}(2x2 + 7)\right] \times (-2.247) = 34.0795$$

3- إيجاد قيمة كي دو المجدولة بدرجة حرية $\frac{1}{2}K(K+1) = \frac{1}{2}2(2+1) = 3$

$$\aleph^2(3) = 7.815$$

4- المقارنة: نلاحظ أن قيمة كي دو المحسوبة 34.0795 أكبر من كي دو المحسوبة $\aleph^2(3) = 7.815$ وهذا يدل على وجود مشكلة التعدد الخطي.

مشكلة عدم تجانس التباين
The Hetroscedasticity Problm

أولا: مفهوم مشكل عدم تجانس التباين

أحد الفرضيات الأساسية التي يقوم عليها النموذج الخطي (العام والبسيط) هو ثبات التباين لحدود الخطأ (تجانس تباين الخطأ)، Homoscedasticity، لجميع المشاهدات (i) أي:

$$E(\varepsilon_i^2) = \sigma_i^2$$

وفي النموذج الخطي العام فإن الفرض المناضر:

$$E(\varepsilon_i \varepsilon_i') = \sigma_i^2 I_n$$

وفي واقع الأمر، فإننا كثيرا مانواجه حالات يتعسر فيها إستيفاء الشرط أعلاه، ومن ثم فإن التباين لا يكون ثابتا بل يختلف لكل مشاهدة من مشاهدات العينة، وتصبح لدينا قيم مختلفة وغير ثابتة لتباينات حدود الخطأ العشوائية، وعليه فإن القطر الرئيسي لمصفوفة التباين والتباين المشترك الخاصة بحدود الخطأ يحتوى على قيم مختلفة وغير ثابتة، أي أن[45]:

$$E(\varepsilon_i \varepsilon_i') \neq \sigma_i^2 I_n$$

حيث أن : $\sigma_1^2 \neq \sigma_2^2 \neq \cdots \neq \sigma_n^2$

ويحدث هذا أساسا في حالة الدراسات المعتمدة على البيانات الإحصائية المقطعية Cross-Section Data

ثانيا: أسباب مشكلة عدم تجانس حد الخطأ

هناك عدة أسباب لظهور هذه المشكلة ومن أهمها ما يأتي:

1- سلوكية وتصرف الأفراد التي تقل الأخطاء فيها بمرور الزمن، وعليه فإن تباين حد الخطأ () يتناقص هو الأخر خلال الفترة الزمنية.

2- مع تحسن أساليب جمع البيانات يقل حد الخطأ وذلك لأن جمع البيانات الدقيقة والواقعية تقلل من الأخطاء، مثال ذلك أن الأخطاء التي ترد في المستندات في المؤسسات الحكومية والتي تستخدم الحاسب الألي لتحليل البيانات تكون أقل من مثيلتها في المؤسسات التي لا تستخدم الحاسب الألى.

[45]- حسين علي بخيت، سحر فتح الله، مرجع سبق ذكره، ص 260.

3- عادة ما تحدث هذه المشكلة في البيانات المقطعية، مثلا عند اعتبار عينة ممثلة لإنفاق الأسر ذات الدخل المنخفض والمرتفع يلاحظ أن تباين الخطأ الخاص بإنفاق الأسر ذات الدخل المنخفض (على السلع الضرورية) تكون أقل من إنفاق الأسر ذات الدخل المرتفع[46].

ثالثا: اكتشاف ظاهرة عدم تجانس تباينات الأخطاء

لاكتشاف ظاهرة عدم تجانس تباينات الأخطاء يستعمل بحثو الاقتصاد القياسي عدة اختبارات مختلفة من بينها:

1- اختبار Goldfed-Quandt

يصلح هذا الاختبار للعينات الصغيرة والكبيرة، حيث يشترط أن يكون حجم العينة على الأقل ضعف عدد المعالم المطلوب تقديرها في النموذج. أستعمل هذا الاختبار لأول مرة من طرف الباحثين Goldfeld-Quandt. حيث إذا رأينا أن أحد المتغيرات المستقلة x_{ji} هو المتسبب في ظاهرة عدم تجانس تباين الأخطاء نعتبر في هذه الحالة، أن σ_i^2 مرتبطة ومتناسبة إيجابيا (طرديا) مع أحد المتغيرات x_{ji} . ثم تقوم بتقسيم الملاحظات إلى عينتين، العينة الأولى تناسب قيم x_{ji} الكبيرة، والعينة الثانية تخص قيم x_{ji} الصغيرة. ثم نوفق إنحدارين مختلفين (منفصلين) على عينتين مختلفين (لكن متساويتين في الحجم)، وذلك لهدف تطبيق اختبار مساواة تبايني أخطاء العينتين المنفصلتين. ويقترح هذين الباحثين حذف بعض الملاحظات المركزية من العينة الأصلية لكي نفرق مابين تباين أخطاء العينتين[47].

مثال:

لنعتبر النموذج الخطي البسيط: $y_i = a_0 + a_1 x_i + \varepsilon_i \ , i = 1,2,\dots\dots,n$

ولنفترض أن الأخطاء ε_i لها التوزيع الطبيعي وغير مرتبطة فيما بينها

أي أن $E(\varepsilon_i \varepsilon_j) = 0$ ، $i \neq$. فإذا كانت لدينا:

$$E(\varepsilon_i^2) = \sigma_i^2 = \sigma_\varepsilon^2 x_i^2$$

ويجرى اختبار (G-Q) وفقا للخطوات التالية:

1- نرتب الملاحظات y_i و x_i وفقا لقيم x_i (من القيمة الصغيرة إلى القيمة الكبيرة)

[46]- حمداوي الطاوس، **مدخل للاقتصاد القياسي**، الطبعة الأولى، دار هومه، الجزائر، 2016، ص101.

[47]- تومي صالح، **مدخل لنظرية القياس الاقتصادي**، الجزء الثاني، مرجع سبق ذكره، ص32.

2- نحذف (نسقط)[48] m ملاحظات مركزية ليكون الباقي هو (n-m) ملاحظات متبقية ثم نقسم الباقي (n-m) إلى عينتين متساويتين (n-m)/2.

تحتوي الأولى على (n-m)/2 من قيم x_i الصغيرة، وتحتوي الثانية على (n-m)/2 ملاحظات من قيم x_i الكبيرة.

3- نكون انحدارين منفصلين بواسطة المربعات الصغرى العادية شريطة أن يكون $\left(\frac{n-m}{2}\right) > k$ على الصيغة التالية:

a) $y_i = a_0 + a_1 x_i + \varepsilon_i \ , i = 1,2,\dots\dots,\frac{n-m}{2}$.

b) $y_i = a_0 + a_1 x_i + \varepsilon_i \ , i = \frac{n-m}{2} + 1, \frac{n-m}{2} + 2, \dots\dots, n$

نحصل من الانحدار الأول على مجموع مربعات البواقي RSS_a، من الانحدار الثاني على RSS_b.

4- نكون الاختبار الإحصائي التالي:

$$\frac{RSS_a}{\sigma_\varepsilon^2} \sim \chi^2_{(n-m-2k)/2}$$
$$\frac{RSS_b}{\sigma_\varepsilon^2} \sim \chi^2_{(n-m-2k)/2}$$

ومنه نجد التوزيع:

$$Q = \frac{RSS_a/v_1}{RSS_b/v_2}$$
$$Q = \frac{RSS_a}{RSS_b} \sim F_{v_1, v_2}$$

حيث أن : $v_1 = v_2 = (n-m-2k)/2$

فإذا كانت Q=1 يعنى هذا تساوي التباينين ومنه نقبل H_0 القائلة بتجانس تباينات الأخطاء، وكلما كان المقدر Q كبيرا، كلما كان عدم التجانس أقوى. وتعتمد قوة الاختبار كذلك على عدد الملاحظات المركزية والمحذوفة من الحجم الكلي للعينة، حيث يقترح Johnston أن يكون هذا العدد في حدود ثلث حجم العينة الأصلية.

2- اختبار White

اقترح White (1980) اختبارا يعتمد على العلاقة بين مربعات البواقي وجميع المتغيرات المستقلة وكذا مربعاتها. يمكن إبراز خطوات هذا الاختبار كما يلي[49]:

[48]- حيث ان m عادة يساوي على الأقل ربع العينة (n/4). فمثلا العدد الأمثل ل m لما تكون n=30 هو m=8. ولما تكون n=60، فإن m=16، وهكذا.

أولا: تقدير النموذج العام $Y = X\beta + \varepsilon$ بطريقة المربعات الصغرى العادية ثم حساب مربعات البواقي $\hat{\varepsilon}_i^2$.

ثانيا: تقدير المعادلة الوسيطية التالية:

$$\hat{\varepsilon}_i^2 = \beta_0 + \beta_1 X_{i1} + \alpha_1 X_{i1}^2 + \cdots + \beta_k X_{ik} + \alpha_k X_{ik}^2 + \mu_i$$

ثم حساب معامل التحديد الخاص بهذه المعادلة R^2.

ثالثا: فرضية ثبات تباين الأخطاء H_0 التي ينبغي إختبارها هي:

$$H_0: \beta_0 = \alpha_1 = \beta_1 = \cdots = \alpha_k = \beta_k = 0$$

إحصائية مضاعف لاغرانج $LM = n \times R^2$ تتبع توزيع χ^2 بدرجة حرية $2k$.

إذا كان $n \times R^2$ أكبر من $\chi^2(2k)$ (القيمة الحرجة لتوزيع χ^2 بنسبة معنوية α)، فإننا نرفض H_0 أي إذا كان هناك على الأقل معامل واحد من معاملات المعادلة الوسيطية يختلف عن الصفر فإن تباين الأخطاء غير متجانس.

3- اختبار ثبات التباين الشرطي للأخطاء ARCH-LM

تسمح نماذج ARCH بنمذجة المتغيرات المالية التي تحتوي على تباين شرطي غير ثابت للأخطاء العشوائية حيث أن التطاير الشرطي الذي يعبر في الغالب عن المخاطرة غير ثابت. يعتمد إذن هذا الاختبار على مضاعف لاغرانج LM. خطوات هذا الاختبار كالتالي:

أولا: تقدير النموذج العام $Y = X\beta + \varepsilon$ بطريقة المربعات الصغرى العادية ثم حساب مربعات البواقي $\hat{\varepsilon}_i^2$.

ثانيا: تقدير المعادلة التالية:

$$\hat{\varepsilon}_i^2 = \theta_0 + \theta_1 \hat{\varepsilon}_{t-1}^2 + \cdots + \theta_q \hat{\varepsilon}_{t-q}^2 + \mu_t$$

ثم حساب معامل التحديد الخاص بهذه المعادلة R^2.نفقد في هذه الحالة q مشاهدة.

ثالثا: فرضية ثبات التباين الشرطي للأخطاء H_0 التي ينبغي اختبارها هي:

$$H_0: \theta_0 = \theta_1 = \cdots = \theta_q = 0$$

إحصائية مضاعف لاغرانج $LM = (n-q) \times R^2$ تتبع توزيع χ^2 بدرجة حرية q.

[49]- شيخي محمد، مرجع سبق ذكره، ص115.

إذا كان $(n-) \times R^2$ أكبر من $\chi^2(q)$ (القيمة الحرجة لتوزيع χ^2 بنسبة معنوية α)، فإننا نرفض H_0 أي إذا كان هناك على الأقل معامل واحد من معاملات معادلة ARCH يختلف معنويا عن الصفر فإن التباين الشرطي للأخطاء غير متجانس.

رابعا: معالجة مشكلة عدم التجانس

الحال الأولى: المربعات الصغرى الموزونة (weighted least squares)

تستخدم هذه الطريقة عندما يكون التباين σ_i^2 معلوما. لتوضيح ذلك، دعنا نفترض أننا نرغب بتقدير معادلة الانحدار البسيط التالية[50]:

$$Y_i = b_0 + b_1 X_i + \varepsilon_i$$

علمنا بإستخدام أحد الاختبارات أن هناك مشكلة " عدم التجانس " في البيانات المستخدمة للتقدير، وبنفس الوقت، تعرفنا على قيمة التباين (σ_i^2). في هذه الحال، يتم تحويل الدالة أعلاه كما يلي: بالقسمة على الانحراف المعياري نحصل على :

$$\frac{Y_i}{\sigma_i} = \frac{b_0}{\sigma_i} + \frac{b_1 X_i}{\sigma_i} + \frac{\varepsilon_i}{\sigma_i}$$

$$(Y_i/\sigma_i) = b_0(1/\sigma_i) + b_1 X_i(1/\sigma_i) + \nu_i$$

حيث ترمز (ν_i) للخطأ العشوائي الجديد. بتقدير معلمتي هذه الدالة نحصل على ما يسمى بمعادلة انحدار المربعات الصغرى الموزونة (weighted least squares) (WLS) وتكون صيغة المعادلة المقدر كما يلي:

$$(Y_i/\sigma_i) = (\hat{b}_0/\sigma_i) + b_1(X_i/\sigma_i)$$

نلاحظ أن الطرف الأيمن من المعادلة يتكون من حاصل قسمة $(\hat{b}_0)$ على (σ_i) وهو متغير مستقل، ويتم الحصول عليه بقسمة الواحد الصحيح على قيمة الانحراف المعياري. ويتكون الطرف الأيمن من حاصل قسمة (X_i) على (σ_i)، وهما متغيران مستقلان. وبذلك فإن المعادلة البسيطة الأصلية (ذات متغير مستقل واحد) قد تحولت إلى معادلة انحدار متعددة المتغيرات (من متغيرين) ولا يوجد لها نقطة تقاطع ويتم تفسير معامل التحديد (R^2) بطريقة تختلف عن الطريقة العادية.

مثال: إذا كانت لدينا نتائج تقدير دالة الانحدار التالية:

$$Y_i = 292.1798 + 10.4042 X_i$$

[50]– عبدالرزاق بني هاني، مرجع سبق ذكره ، ص150.

$$(33.6705) \quad (2.1235)$$

$$R^2 = 0.75 \qquad n = 10$$

بإستخدام المربعات الصغرى الموزونة بقسمة كل من (Y_i) و (X_i) على الانحراف المعياري (σ_i) وتقدير معلمتي دالتي الانحدار من البيانات المحولة نحصل على النتائج التالية:

$$(Y_i/\sigma_i) = 307.4466\,(1/\sigma_i) + 9.2143(X_i/\sigma_i)$$

$$(25.0225) \qquad (1.800)$$

$$R^2 = 0.66 \qquad n = 10$$

يتضح من هذه النتائج مدى تحسن التقديرات وأخطائها المعيارية رغم أن قيمة (R^2) قد إنخفضت، لكن مقدرة المعادلة الثانية على التنبؤ تبقى أفضل من بديلتها الأولى.

الحالة الثانية: المربعات الصغرى الموزونة في حال عدم معرفة التباين

لنفترض أننا نرغب بتقدير معلمتي معادلة انحدار من الشكل التالي:

$$Y_i = b_0 + b_1 X_i + \varepsilon_i$$

وكان:

$$Var(\varepsilon_i) = E(\varepsilon_i^2) = \sigma^2 X_i^n$$

أي أن تباين الخطأ العشوائي هو دالة في المتغير (X_i)، ويتناسب طرديا معه.

في هذه الحال يتم تحويل معادلة الانحدار أعلاه، بتقسيم طرفيها على المتغير $\left(X_i^{n/2}\right)$، أي

$$Y_i/(X_i)^{n/2} = \frac{b_0}{X_i^{n/2}} + \frac{b_1 X_i}{X_i^{n/2}} + \frac{\varepsilon_i}{X_i^{n/2}} = b_0/(X_i)^{n/2} + b_1 X_i/(X_i)^{n/2} + \nu_i$$

حيث ترمز (ν_i) للخطأ العشوائي الجديد. ومنه يكون

$$Var(\varepsilon_i) = Var(\varepsilon_i)\left(\frac{1}{X_i^{n/2}}\right)^2 = \sigma^2 X_i^n/(X_i)^n = \sigma^2$$

وبالتالي نكون قد حققنا تجانسا في تباين الخطأ العشوائي.

أما في حال أن:

$$Var(\varepsilon_i) = E(\varepsilon_i^2) = \sigma^2(b_0 + b_1 X_i)^n$$

فإن تحويل الدالة يتم بقسمة أطرافها كما يلي:

$$\frac{Y_i}{(b_0 + b_1 X_i)^{n/2}} = \frac{b_0}{(b_0 + b_1 X_i)^{n/2}} + \frac{b_1 X_i}{(b_0 + b_1 X_i)^{n/2}} + \frac{\varepsilon_i}{(b_0 + b_1 X_i)^{n/2}}$$

حيث يتم الحصول على $(Y_i = b_0 + b_1 \quad _i)$ من تقدير معلمتي دالة الانحدار، ثم تستخدم الدالة المقدرة $\left(\hat{Y}_i = \hat{b}_0 + b_1 X_i\right)$ للتحويل كما يلي:

$$\frac{Y_i}{(\hat{Y}_i)^{n/2}} = \frac{b_0}{(\hat{Y}_i)^{n/2}} + \frac{b_1 X_i}{(\hat{Y}_i)^{n/2}} + \frac{\varepsilon_i}{(\hat{Y}_i)^{n/2}}$$

الحالة الثالثة: يمكننا التخلص من مشكلة عدم التجانس بإستخدام اللوغاريتم الطبيعي لتحويل معادلة الانحدار كما يلي:

لنفترض ان لدينا دالة الانحدار التالية :

$$Y_i = b_0 + b_1 X_i + \varepsilon_i$$

وعلمنا، من خلال الاختبارات، بوجود مشكلة " عدم التجانس ". يمكننا استخدام اللوغاريتم الطبيعي في تحويل البيانات للحصول على الصيغة التالية، ثم تقدير معلماتها للتخلص من المشكلة:

$$Ln\, Y_i = \mathrm{A} + B\, Ln\, X_i + \omega_i$$

يتم تقدير معلمتى الدالة الجديدة بإستخدام البيانات المحولة باللوغاريتم الطبيعي، إذ أن اللوغاريتم يقوم بتخفيض الأوزان الكبيرة (المشاهدات الكبيرة) كي تكون قريبة من المشاهدات ذات الأوزان الصغيرة.

تمرين حول مشكل عدم تجانس التباين

لتكن لدينا البيانات التالية:

i	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
y	4	5	6	8	8	6	11	13	15	17	9	13	14	15	21
X	4	4	4	4	4	3.5	3.5	3.5	3.5	3.5	2	2	2	2	2
i	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
y	6	13	16	23	26	11	15	17	22	34	7	21	23	28	38
X	1.5	1.5	1.5	1.5	1.5	1	1	1	1	1	0.5	0.5	0.5	0.5	0.5

المطلوب:

1- قدر معالم النموذج ؟

2- من خلال الرسم البياني لسحابة النقاط ،هل هناك تجانس لتباين الأخطاء ؟

3- اختبر وجود تجانس تباين الخطأ بإجراء الاختبارات التالية:

أ- اختبار white ؟

ب- اختبار ثبات التباين الشرطي ARCH-LM ؟

4-قم بتصحيح النموذج من مشكل عدم تجانس تباين الأخطاء إن وجد ؟

الحل:

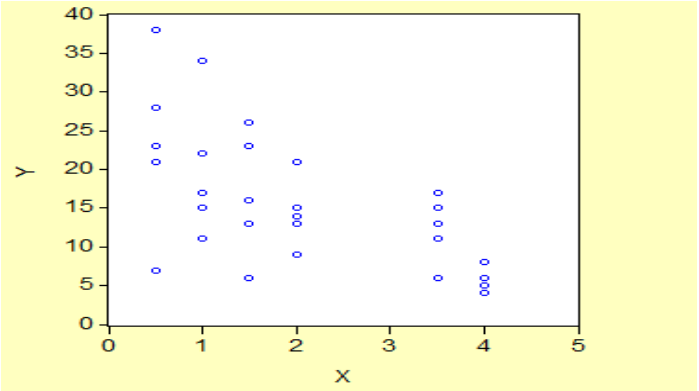
1- تقدير معالم النموذج

$$y = a_0 + a_0 x + \varepsilon_i$$

Dependent Variable: Y
Method: Least Squares
Date: 05/12/14 Time: 20:43
Sample: 1 30
Included observations: 30

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	24.09442	2.397697	10.04898	0.0000
X	-4.125322	0.982272	-4.199774	0.0002

R-squared	0.386478	Mean dependent var	15.50000
Adjusted R-squared	0.364566	S.D. dependent var	8.585272
S.E. of regression	6.843674	Akaike info criterion	6.748867
Sum squared resid	1311.404	Schwarz criterion	6.842280
Log likelihood	-99.23300	F-statistic	17.63810
Durbin-Watson stat	1.808528	Prob(F-statistic)	0.000245



من الرسم البياني نلاحظ تشتت سحابة النقاط x,y وبالتالي يوجد عدم ثبات للتباين (عدم تجانس تباين الأخطأ)

2-اختبارات الكشف عن وجود تجانس تباين الأخطاء:

أ- اختبار White:

White Heteroskedasticity Test:

F-statistic	3.956388	Probability	0.031126
Obs*R-squared	6.799324	Probability	0.033385

Test Equation:
Dependent Variable: RESID^2
Method: Least Squares
Date: 05/12/14 Time: 20:46
Sample: 1 30
Included observations: 30

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	136.0182	43.70370	3.112280	0.0044
X	-78.58151	48.29367	-1.627160	0.1153
X^2	11.98436	10.25804	1.168289	0.2529

R-squared	0.226644	Mean dependent var	43.71348
Adjusted R-squared	0.169359	S.D. dependent var	69.19018
S.E. of regression	63.05961	Akaike info criterion	11.22068
Sum squared resid	107365.9	Schwarz criterion	11.36080
Log likelihood	-165.3102	F-statistic	3.956388
Durbin-Watson stat	1.589880	Prob(F-statistic)	0.031126

نحسب إحصائية مضاعف لغرنج :

$LM=n.R^2=30x0.226644=6.79932$

إيجاد قيمة كي دو المجدولة بدرجة حرية 2K

لدينا K=1 إذا درجة الحرية هي 2

$$\aleph_{\alpha}^{2}(2K) = \aleph_{\alpha}^{2}(2) = 5.991$$

المقارنة :

نلاحظ ان فيمة LM =6.79932 أكبر من قيمة كي دو بدرجة حرية 2 التي هي 5,991 وبالتالي نرفض الفرضية H0 ونقبل الفرضية H1 القائلة بعدم تجانس تباين الأخطاء

ب- اختبار ثبات التباين الشرطي ARCH-LM:

ARCH Test:

F-statistic	1.887374	Probability	0.180798
Obs*R-squared	1.894733	Probability	0.168669

Test Equation:
Dependent Variable: RESID^2
Method: Least Squares
Date: 05/12/14 Time: 20:49
Sample (adjusted): 2 30
Included observations: 29 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	33.41826	15.26088	2.189800	0.0374
RESID^2(-1)	0.311774	0.226940	1.373817	0.1808

R-squared	0.065336	Mean dependent var	44.77564
Adjusted R-squared	0.030718	S.D. dependent var	70.16552
S.E. of regression	69.07943	Akaike info criterion	11.37486
Sum squared resid	128843.1	Schwarz criterion	11.46916
Log likelihood	-162.9355	F-statistic	1.887374
Durbin-Watson stat	1.454612	Prob(F-statistic)	0.180798

نحسب إحصائية مضاعف لغرنج :

$LM=(n-q).R^2=(30-1)x0.065336=1.8947$

إيجاد قيمة كي دو المجدولة بدرجة حرية q

لدينا q=1 إذا درجة الحرية هي 1

$$\aleph_{\alpha}^{2}(2K) = \aleph_{\alpha}^{2}(1) = 3.841$$

المقارنة :

نلاحظ ان فيمة LM =1.8947 أكبر من قيمة كي دو بدرجة حرية 1 التي هي 3.841 وبالتالي نرفض الفرضية H1 ونقبل الفرضية H0 التي تدل على وجود تجانس تباين الأخطاء

ARCH Test:

F-statistic	1.887374	Probability	0.180798
Obs*R-squared	1.894733	Probability	0.168669

2- تصحيح النموذج:

نقوم بحساب تباين الأخطاء

$$\sigma_\varepsilon^2 = \frac{\sum \hat{\varepsilon}_i^2}{n-k-1} = \frac{1311.404}{30-1-1} = 46.835$$

ثم نقوم بتحويل المتغيرات :

$$a_{01} = \frac{a_0}{\sqrt{\sigma_\varepsilon^2}} = \frac{a_0}{46.835}$$

$$Y1 = \frac{Y}{\sqrt{\sigma_\varepsilon^2}} = \frac{Y}{46.835}$$

$$X1 = \frac{X}{\sqrt{\sigma_\varepsilon^2}} = \frac{X}{46.835}$$

ثم نقوم بتقدير النموذج المصحح:

Dependent Variable: Y1
Method: Least Squares
Date: 05/15/14 Time: 06:59
Sample: 1 30
Included observations: 30

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	3.520718	0.350356	10.04898	0.0000
X1	-4.125322	0.982272	-4.199774	0.0002

R-squared	0.386478	Mean dependent var	2.264886
Adjusted R-squared	0.364566	S.D. dependent var	1.254495
S.E. of regression	1.000009	Akaike info criterion	2.902236
Sum squared resid	28.00052	Schwarz criterion	2.995649
Log likelihood	-41.53354	F-statistic	17.63810
Durbin-Watson stat	1.808528	Prob(F-statistic)	0.000245