دراسة مقارنة لبعض طرائق تقدير مصفوفة التباين والتباين المشترك الحصينة للمعلمات المقدرة بطريقة (OLS) في البيانات المقطعية

أ.د. سلمى ثابت ذاكر الآلوسي / كلية الإدارة والاقتصاد / الجامعة المستنصرية
الباحث/ ديان حميد مجيد الربيعاوي / الجامعة المستنصرية/ كلية الإدارة والاقتصاد

تاريخ التقديم:2016/10/24
تاريخ القبول: 2016/12/14

المستخلص

أن أنموذج الأنحدار الخطي الطبيعي الكلاسيكي(Classical Normal Linear Regression Model) قائم على أساس العديد من الفرضيات من بينها فرضية تجانس التباين، كما هو معروف فأن استخدام طريقة المربعات الصغرى الأعتيادية (OLS)، تحت ظل وجود هذه المشكلة يجعل مقدراتها تفقد بعضاً من خصائصها المرغوب فيها، كما أن الأستدلال الأحصائي غير مقبول، وعليه فقد تم وضع بديلين مهمين الأول طريقة المربعات الصغرى العمومية (Generalized Least Square) والتي يرمز له (GLS)، أما البديل الثاني فهو تقدير مصفوفة التباين والتباين المشترك الحصينة (Robust covariance matrix estimation) للمعلمات المقدرة بطريقة (OLS). تكون حسب نوع البيانات التي يتم التعامل معها، ولقد تناولنا في هذه الدراسة البيانات المقطعية (Cross-Section)، حيث أن مشكلة عدم تجانس في التباين تكون واردة فيها وان مصفوفة التباين والتباين المشترك الحصينة المقدرة لهذا النوع من البيانات هي (HCCME) وتتضمن العديد من الطرائق، ولقد تناولنا في دراستنا بعضاً منها والمتمثلة بـ(HC_0 ، HC_1 ، HC_2 ، HC_3، HC_{3a}) .

ولقد تمت في هذه الدراسة مقارنة هذه الطرائق وتحديد أولوية أدائها بالنسبة لأداء طريقة (GLS) وذلك في حالة البيانات المقطعية وبأستخدام أسلوب المحاكاة في توليد عينات بأحجام مختلفة.

المصطلحات الرئيسة للبحث/عدم تجانس التباين ، الحصين ، الاتساق ، كفوءة ، عدم التحيز ، انحدار، متحيز ، تقدير ، التباين المشترك.

مجلة العلوم
الاقتصادية والإدارية
العدد 98 المجلد 23
الصفحات 384ـ 405

*البحث مستل من رسالة ماجستير

1.المقدمة (Introduction):

أن ابرز المشاكل التي تواجه أنموذج الانحدار الخطي هي مشكلة عدم تجانس التباين (Heteroscedasticity)، مما يجعل مقدرات طريقة المربعات الصغرى الأعتيادية (OLS) لمعلمات الأنموذج، لا تمتلك كافة الخصائص المرغوب فيها، مع ذلك فان هذه المقدرات تبقى محتفظة ببعض الخصائص المهمة مثل خاصية عدم التحيز (Unbiased) وخاصية الأتساق (Consistent)، الا أن الأمر الأكثر خطورة هو أن تباين المقدر لهذه المعلمات يصبح متحيزاً في حالة وجود هذه المشكلة، كما يجعل الأستدلال الأحصائي بأعتماد طريقة (OLS) في هذه الحالة (غير مقبول) ولا يمكن الأعتماد عليه، وعلى هذا الأساس وضعت بدائل لطريقة (OLS) تتلافى هذا الأمر، البديل الأول هو أستخدام طريقة المربعات الصغرى العمومية (Generalized Least Square)، أما البديل الثاني فهو أستخدام مصفوفة التباين والتباين المشترك الحصينة (Robust covariance matrix estimation) للمعلمات المقدرة بطريقة (OLS).

أن البديل الأول (GLS) يعطي مقدرات كفوءة (Efficient) والأستدلال الأحصائي الموضوع على أساسه مقبول وهي طريقة معتمدة وذلك في حالة هنالك درجة من التأكد من طبيعة المشكلة والأنموذج الموضوع لهذه المشكلة، وبخلافه فأن طريقة (GLS) تصبح غير مناسبة [10].

أما البديل الثاني و الذي يتمثل بتقدير مصفوفة التباين والتباين المشترك الحصينة فأنه لا يتطلب معرفة بطبيعة المشكلة والأنموذج الخاص بها هذا أولاً، ثانياً أنها طريقة سهلة والأستدلال الأحصائي الموضوع على أساسها مقبول ولقد لاقت هذه الطريقة اهتماماً متزايداً من قبل الباحثين منذ عام (1980) [15]، حيث تم أيجاد تقدير مصفوفة التباين والتباين المشترك الحصينة للبيانات المقطعية، ولقد حاولنا في هذه الدراسة مقارنة أداء طريقة (GLS) مع هذه الطرائق أولاً، وتحديد أولوية هذه الطرائق من حيث الأداء الأفضل مقارنة مع طريقة (GLS) وذلك في حالة البيانات المقطعية و ذلك بأعتماد أسلوب المحاكاة في توليد العينات بأحجام مختلفة لكل منها .

2.هدف البحث (Purpose of search):

يستهدف هذا البحث تقييم أداء طرائق تقدير مصفوفة التباين والتباين المشترك الحصينة للبيانات المقطعية (Cross-Section) وفي حالة وجود مشكلة عدم تجانس التباين والمعروفة بأسم (HCCME) وهي مختصر لـ (Heteroscedasticity-Consistent Covariance Matrix Estimator) والتي تتضمن الطرائق (HC_0 ،HC_1 ،HC_2 ، HC_3 ، HC_{3a}) وذلك من خلال مقارنة أداء هذه الطرائق مع طريقة المربعات الصغرى العمومية القابلة للتطبيق (FGLS) او ما تسمى بطريقة المربعات الصغرى الموزونة القابلة للتطبيق (FWLS).

الجانب النظري

3. أنموذج الانحدار الخطي الطبيعي الكلاسيكي (CNLRM)[7] :

أن أنموذج الأنحدار الخطي الطبيعي الكلاسيكي Classical Normal Linear Regression (Model)، يعد حجر الزاوية لأغلب نظريات القياسي الاقتصادي والذي يمكن تمثيله بالمعادلة الاتية :

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \cdots + \beta_k X_{ik} + U_i \quad \ldots\ldots (1)$$

$$i = 1،2، \ldots ، n$$

$$k = 0،1،2، \ldots ،K$$

حيث ان:

Y_i : يمثل المتغير المعتمد للقيم المشاهدة.

X_{ik} : يمثل المتغيرات التوضيحية للانموذج.

β_k : يمثل المعلمات للانموذج.

U_i : يمثل حد الخطأ العشوائي والذي يفترض فيه تحقيق كافة الشروط.

ان تقدير معلمات الأنموذج (1) المذكورة آنفاً بطريقة المربعات الصغرى الأعتيادية (OLS) وعند توفر الفرضيات الأساسية تعطي مقدرات تشتمل على كافة الخصائص المرغوب فيها سواء في العينات الصغيرة او في العينات الكبيرة حيث تتصف هذه المقدرات في العينات الصغيرة بكونها (***BLUE***، ***Unbiased***، ***Efficient***) ، وفي العينات الكبيرة تتصف بكونها متسقة (***Consistent***) وكفوء تقريبا (***Asymptotically Efficient***) ومصفوفة التباين والتباين المشترك للمعلمات المقدرة صحيحة، والاستدلال الاحصائي الذي يتضمن اختبارات (F, t) وحدود الثقة مقبولة، كما ان طريقة المربعات الصغرى (OLS) لا تتطلب شرط التوزيع الطبيعي في الأخطاء العشوائية (U_i) ولكن الاستدلال الاحصائي حول التقديرات تقتضي توفر هذه الفرضية [6][7]. هناك العديد من المشاكل التي تواجه أنموذج الانحدار (***CNLRM***) الا ان ابرز هذه المشاكل في اغلب التطبيقات العلمية هي حدوث مشكلة عدم تجانس التباين.

4. ماهي مشكلة عدم تجانس تباين (Heteroscedasticity)[7] [10]:

تحدث مشكلة عدم تجانس التباين عند مخالفة الفرضية الاتية لأنموذج الانحدار الخطي الكلاسيكي وعندما يكون

$$Var(U_i\ ,\ X_i) \neq \sigma^2$$

وتكون هناك انعكاسات سلبية على بعض خصائص مقدرات (OLS)،كما انها تؤثر أيضا سلبياً في خصائص التباينات المقدرة للمعلمات المقدرة، وان هذه المشكلة تظهر بصورة خاصة في البيانات المقطعية (Cross-Section Data) و يمكن توضيحها وبالشكل الاتي:-

5. خصائص مقدرات طريقة المربعات الصغرى (OLS)[7] :

تفقد مقدرات طريقة المربعات الصغرى(OLS) بعض من خصائصها سواء في العينات الصغيرة او الكبيرة، الا انها تبقى محتفظة ببعض الخصائص المرغوب فيها مثل خاصية عدم التحيز (***Unbiased***) ، وخاصية الاتساق (***Consistent***) الأمر الذي يقود الى إمكانية اعتمادها كمقدرات ويمكن البرهنة على ذلك كالاتي:-

1- **خاصية عدم التحيز** (*Unbiased*)[10]:

ليكن لدينا أنموذج الانحدار البسيط وبالشكل التالي:

$$Y_i = \beta_0 + \beta_1 X_i + U_i \quad \dots\dots\dots (2)$$

وعليه يمكن برهنة صفة عدم التحيز بالنسبة للمعلمة (β_1) فقط وكالأتي:

$$\hat{\beta}_1 = \frac{\sum_{i=1}^{n} x_i Y_i}{\sum_{i=1}^{n} x_i^2} \quad \dots\dots (3)$$

$$\hat{\beta}_1 = \frac{\sum_{i=1}^{n} x_i Y_i}{\sum_{i=1}^{n} x_i^2} = \beta_1 + \frac{\sum_{i=1}^{n} x_i \epsilon_i}{\sum_{i=1}^{n} x_i^2}$$

$$E(\hat{\beta}_1) = \beta_1 + E\left[\frac{\sum_{i=1}^{n} x_i \epsilon_i}{\sum_{i=1}^{n} x_i^2}\right] = \beta_1$$

تقدير غير متحيز $\therefore E(\hat{\beta}_1) = \beta_1$

وكذلك بالنسبة لـ (β_0) فأن تقديرها ايضاً غير متحيز.

2- **خاصية الأتساق** (*Consistent*) *[10]*:

ويتم اثبات خاصية الأتساق بالاعتماد على أنموذج الانحدار البسيط (2) أعلاه وبالاعتماد على المعلمة (β_1) فقط إذ ان:

$$plim\hat{\beta}_1 = plim\left[\frac{\sum_{i=1}^{n} x_i y_i}{\sum_{i=1}^{n} x_i^2}\right] = \beta_1 + plim\left[\frac{\sum_{i=1}^{n} x_i u_i}{\sum_{i=1}^{n} x_i^2}\right]$$
$$= \beta_1 + \frac{plim(\sum_{i=1}^{n} x_i u_i/n)}{plim(\sum_{i=1}^{n} x_i^2/n)}$$

وبما أن $\frac{\sum_{i=1}^{n} x_i u_i}{n} = 0$

فأن : $\therefore plim\hat{\beta}_1 = \beta_1$

والذي يشير أن $(\hat{\beta}_1)$ مقدر متسق لـ (β_1) .

6. خصائص التباينات المقدرة للمعلمات المقدرة بطريقة (OLS)[6] [10]:

كما ذكرنا ان مقدرات (OLS) تبقى محتفظة ببعض الخصائص المرغوب فيها، ولكن عندما نستخدم هذه المقدرات لأختبار الفرضيات أو وضع فترات الثقة سوف تكون الحاجة ليست فقط ان تكون هذه المقدرات غير متحيزة ولكن التباينات المقدرة لها تكون غير متحيزة ايضاً، وبخلافه فأن الأختبارات وفترات الثقة الموضوعة تكون غير معتمدة وغير مقبولة، أن هذه التباينات المقدرة للمعلمات بطريقة المربعات الصغرى (OLS) تحت ظل وجود مشكلة عدم تجانس التباين ستتصف بكونها متحيزة (biased)، الأمر الذي يقود الى ان تكون الأختبارات الموضوعة وفترات الثقة غير مقبولة وغير معتمدة وهذا بالتأكيد هو السبب الرئيس الى العزوف عن استخدام طريقة المربعات الصغرى (OLS) عند حدوث مشكلة عدم تجانس التباين واللجوء الى طرائق أخرى في التقدير[6]. ويمكن البرهنة على ذلك كالأتي:

بأفتراض ان الأنموذج الخطي البسيط لمعادلة (2) هو:

$$Y_i = \beta_0 + \beta_1 X_i + U_i$$

فأن مقدر (OLS) للميل الحدي هو($\widehat{\beta}_1$) وأن الصيغة المألوفة لأحتساب التباين الخاص به هو[10]:

$$S^2_{\widehat{\beta}_1} = \frac{s^2}{\sum_{i=1}^{n} x_i^2} \quad \dots\dots\dots (4)$$

$$E(s^2) = E[\frac{1}{n-2}] \sum_{i=1}^{n} [\beta_0 + \beta_1 x_i + \epsilon_i - \widehat{\beta}_0 - \widehat{\beta}_1 x_i]^2$$

$$= \frac{1}{n-2} \sum_{i=1}^{n} E[-(\widehat{\beta}_0 - \beta_0) - (\widehat{\beta}_1 - \beta_1) x_i + \epsilon_i]^2$$

$$\therefore \widehat{\beta}_0 - \beta_0 = -(\widehat{\beta}_1 - \beta_1)\overline{X} + \bar{\epsilon}$$

$$\therefore E(s^2) = \frac{1}{n-2} \sum_{i=1}^{n} E[-(\widehat{\beta}_1 - \beta_1) x_i + \acute{\epsilon}_i]^2 \qquad , \quad \epsilon_i - \bar{\epsilon} = \acute{\epsilon}_i$$

$$= \frac{1}{n-2} [E(\widehat{\beta}_1 - \beta_1)^2 \sum_{i=1}^{n} x_i^2 + E \sum_{i=1}^{n} \acute{\epsilon}_i^2 - 2E(\widehat{\beta}_1 - \beta_1) \sum_{i=1}^{n} x_i \acute{\epsilon}_i]$$

حيث ان:

$$E(\widehat{\beta}_1 - \beta_1)^2 = \frac{\sum_{i=1}^{n} x_i^2 \sigma_i^2}{(\sum_{i=1}^{n} x_i^2)^2}$$

$$E(\acute{\epsilon}_i^2) = E(\epsilon_i - \bar{\epsilon})^2 = E\epsilon_i^2 + E\bar{\epsilon}^2 - 2E\epsilon_i \bar{\epsilon}$$

$$= E\epsilon_i^2 + \frac{1}{n^2} E\left[\sum_{i=1}^{n} \epsilon_i^2 + \sum_{i<j}^{n} \epsilon_i \epsilon_j\right] - \frac{2}{n} E\epsilon_i [\epsilon_1 + \epsilon_2 + \cdots + \epsilon_i + \cdots \epsilon_n]$$

$$= \sigma_i^2 + \frac{\sum_{i=1}^{n} \sigma_i^2}{n^2} - \frac{2\sigma_i^2}{n} = \left[\frac{n-2}{n}\right] \sigma_i^2 + \frac{\sum_{i=1}^{n} \sigma_i^2}{n^2}$$

$$E(\widehat{\beta}_1 - \beta_1) \sum_{i=1}^{n} x_i \acute{\epsilon}_i = E\left[\frac{\sum_{i=1}^{n} x_i \acute{\epsilon}_i}{\sum_{i=1}^{n} x_i^2}\right] \sum_{i=1}^{n} x_i \acute{\epsilon}_i$$

$$= \left[\frac{\sum_{i=1}^{n} x_i \acute{\epsilon}_i}{\sum_{i=1}^{n} x_i^2}\right]^2 \left[\sum_{i=1}^{n} x_i^2\right]$$

$$Var(\widehat{\beta}_1)\left[\sum_{i=1}^{n} x_i^2\right] = \frac{\sum_{i=1}^{n} x_i^2 \sigma_i^2}{(\sum_{i=1}^{n} x_i^2)^2} \sum_{i=1}^{n} x_i^2$$

$$E(s^2) = \frac{1}{n-2} \left\{ \frac{\sum_{i=1}^{n} x_i^2 \sigma_i^2}{(\sum_{i=1}^{n} x_i^2)^2} \sum_{i=1}^{n} x_i^2 + \sum_{i=1}^{n} \left[\sigma_i^2 + \frac{\sum_{i=1}^{n} \sigma_i^2}{n^2} - \frac{2\sigma_i^2}{n}\right] - 2 \frac{\sum_{i=1}^{n} x_i^2 \sigma_i^2}{(\sum_{i=1}^{n} x_i^2)^2} \sum_{i=1}^{n} x_i^2 \right\}$$

$$= \frac{1}{n-2}\left[-\frac{\sum_{i=1}^{n} x_i^2 \sigma_i^2}{\sum_{i=1}^{n} x_i^2} + \sum_{i=1}^{n} \sigma_i^2 + \frac{\sum_{i=1}^{n} \sigma_i^2}{n} - \frac{2\sum_{i=1}^{n} \sigma_i^2}{n}\right]$$
$$= \frac{-n\left(\sum_{i=1}^{n} x_i^2 \sigma_i^2\right) + (n-1)\left(\sum_{i=1}^{n} x_i^2\right)\left(\sum_{i=1}^{n} \sigma_i^2\right)}{n(n-2)\left(\sum_{i=1}^{n} x_i^2\right)}$$

$$E\left(s_{\hat{\beta}_1}^2\right) = \frac{E(s^2)}{\sum_{i=1}^{n} x_i^2} = \frac{-n\left(\sum_{i=1}^{n} x_i^2 \sigma_i^2\right) + (n-1)\left(\sum_{i=1}^{n} x_i^2\right)\left(\sum_{i=1}^{n} \sigma_i^2\right)}{n(n-2)\left(\sum_{i=1}^{n} x_i^2\right)^2} \quad ...(5)$$

$$\therefore Var\left(\hat{\beta}_1\right) = \frac{\sum_{i=1}^{n} x_i^2 \sigma_i^2}{\left(\sum_{i=1}^{n} x_i^2\right)^2} = \frac{\sum_{i=1}^{n}\left(x_i^2 / w_i\right)}{\left(\sum_{i=1}^{n} x_i^2\right)^2} \quad ......(6)$$

بما ان $E\left(S_{\hat{\beta}_1}^2\right)$ في معادلة (5) مختلف عن $Var\left(\hat{\beta}_1\right)$ في معادلة (6) يدل على ان التباين المقدر للمعلمات المقدرة بطريقة (OLS) متحيز ويتم احتساب التحيز[10] بالشكل الاتي:

$$E\left(s_{\hat{\beta}_1}^2\right) - Var\left(\hat{\beta}_1\right) = \frac{-n\left(\sum_{i=1}^{n} x_i^2 \sigma_i^2\right) + (n-1)\left(\sum_{i=1}^{n} x_i^2\right)\left(\sum_{i=1}^{n} \sigma_i^2\right)}{n(n-2)\left(\sum_{i=1}^{n} x_i^2\right)^2}$$
$$- \frac{\sum_{i=1}^{n} x_i^2 \sigma_i^2}{\left(\sum_{i=1}^{n} x_i^2\right)^2}$$

أن العواقب المتأتية من كون التباينات المقدرة للمعلمات بطريقة (OLS) متحيزة، تتمثل كما ذكرنا آنفاً بأن يكون الاستدلال الأحصائي حول معلمات المجتمع غير صحيح، اذ ستكون فترات الثقة الموضوعة ومناطق قبول الفرضية (Acceptance Regions) جميعها خاطئة، وعليه من الضروري معرفة اتجاه الخطأ (Direction of Error) وذلك حتى يكون بالإمكان معرفة فيما اذا كانت فترات الثقة الموضوعة ومناطق القبول للفرضية أوسع (Wider) أو أضيق (Narrower) من المناطق وفترات الثقة الصحيحة، ويمكن تحقيق ذلك من خلال تحديد اتجاه التحيز للتباينات المحتسبة فأذا كان التحيز موجباً فأن فترات الثقة الموضوعة ومناطق القبول سوف تكون أوسع من المناطق والفترات الصحيحة الأمر الذي يعمل الى ان تكون أختبارات (t , F) غير معنوية في حين أنها في واقع الحال غير ذلك. واذا كان اتجاه التحيز سالباً سوف تكون فترات الثقة ومناطق القبول أضيق من نظيرتها الصحيحة الأمر الذي يعمل على جعل الأختبارات (t , F) معنوية وهي في واقع الحال ليست كذلك.

وهذا ما يسمى بالأنحدار المعظم أو الرائق (Surprise regions) وأن التحيز يمكن تمثيله بالمعادلة الأتية:-

$$biased = \frac{-n(n-1)\left(\sum_{i=1}^{n} x_i^2 \sigma_i^2\right) + (n-1)\left(\sum_{i=1}^{n} x_i^2\right)\sum_{i=1}^{n} \sigma_i^2}{n(n-2)\left(\sum_{i=1}^{n} x_i^2\right)^2} \quad ...(7)$$

وعليه عند استخدام طريقة (OLS) تحت ظل وجود مشكلة عدم تجانس التباين فأن الأختبارات (t , F) سواء كانت معنوية او غير معنوية والأخطاء المعيارية للمعلمات المقدرة سواء كانت صغيرة أم كبيرة فأنها غير معتمدة وغير صحيحة[10][6].

7.البدائل المتاحة في حالة وجود مشكلة عدم تجانس التباين:

في ضوء السلبيات التي تم بيانها مفصلاً والتي تتعلق بعدم إمكانية الأعتماد على التباينات المقدرة للمعلمات المقدرة بطريقة (OLS) في حالة وجود مشكلة عدم تجانس التباين ، لكونها متحيزة ومن ثم عدم إمكانية قبول الأختبارات الموضوعة وفترات الثقة حول المقدرات، ويعد هذا هو السبب الرئيس للعزوف أستخدام طريقة (OLS) وعلى هذا الأساس ظهرت طرائق أخرى بديلة تتجاوز هذه المشكلة والتي تقود الى وضع مقدرات بخصائص مرغوب فيها والأستدلال الأحصائي معتمد ومقبول وتتمثل بما يأتي:

أولاً: أستخدام طريقة المربعات الصغرى العمومية (Generalized Least Square).

ثانياً: أستخدام مصفوفة التباين والتباين المشترك الحصينة (Robust Covariance Matrix).

وفيما يلي ادناه التفاصيل الخاصة بكل منها.

7.1 طريقة المربعات الصغرى العمومية (Generalized Least Square)[1][8] :

تمثل هذه الطريقة الأسلوب التقليدي المتبع لوضع التقديرات وأنجاز الأستدلال الأحصائي في حالة وجود مشكلة عدم تجانس التباين ، والذي لا يزال يتبع وعليه لابد من وضع التفصيلات الخاصة بهذا الطريقة وكما يلي:

ويتضمن هذا الأسلوب في الواقع على طريقتين هما:

1- طريقة المربعات الصغرى العمومية والتي يرمز لها (GLS).

2- طريقة المربعات الصغرى العمومية العملية (Feasible Generalized Least Square) والتي يرمز لها (FGLS).

7.1.1 طريقة المربعات الصغرى العمومية (GLS)[1] [6] [8]:

وليكن لدينا أنموذج الخطي العام وبدلالة المصفوفات والأنموذج يكون بالشكل الاتي:-

$$\underline{Y} = X\underline{\beta} + \underline{U} \quad \ldots\ldots\ldots\ldots (8)$$

إذ ان: $\underline{U} \sim N(0\,,\sigma^2\Omega)$

وكما هو معروف فأن:

$\underline{Y}$: يمثل موجه قيم مشاهدات المتغير المعتمد بالرتبة $(n \times 1)$.

$\underline{\beta}$: يمثل موجه المعلمات بالرتبة $[(k+1)\times 1]$.

X : يمثل مصفوفة المتغيرات التوضيحية بالرتبة $[n \times (k+1)]$.

$\underline{U}$: يمثل موجه الأخطاء بالرتبة $(n \times 1)$.

ولتقدير موجه المعلمات($\underline{\beta}$) بطريقة (GLS) نستخدم الصيغة الأتية :

$$\underline{\widehat{\beta}}_{GLS} = (X'\Omega^{-1}X)^{-1}X'\ \Omega^{-1}Y \quad \ldots.\ \ldots\ \ldots (9)$$

وأن الصيغة الخاصة لتقدير مصفوفة التباين والتباين المشترك للمعلمات المقدرة بطريقة (GLS) هي:

$$Var - Cov(\hat{\beta}_{GLS}) = (X' \ \Omega^{-1} \ X)^{-1}\sigma^2 \quad \dots\dots (10)$$

أن عناصر مصفوفة (Ω) تحدد حسب نوع المشكلة التي تواجه أنموذج الأنحدار ففي حالة وجود مشكلة عدم تجانس التباين فأن مصفوفة (Ω) هي عبارة عن مصفوفة قطرية، قطرها يمثل الأوزان وكما مبين فيما يأتي :-

فعندما تكون مصفوفة (Ω) معلومة (*Known*) فأن خصائص طريقة (GLS) هي كالأتي:-

1- مقدرات المعلمات تكون غير متحيزة (*Unbiased*) و (*BLUE*) وكفوءة (*Efficient*) ومتسقة (*Consistent*) وكفوءة في العينات الكبيرة (*Asymptotically Efficient*) .

2- مصفوفة التباين والتباين المشترك للمعلمات المقدرة صحيحة (*Correct*) وعليه فأن الأخطاء المعيارية المقدرة هي غير متحيزة (*Unbiased*)، ومتسقة (*Consistent*).

3- الأختبارات الأحصائية الموضوعة تكون مقبولة (*Valid*).

ففي حالة وجود مشكلة عدم تجانس التباين فأن مصفوفة (Ω) تكون :

$$\Omega_{n\times n} = \begin{bmatrix} \frac{1}{w_1} & 0 & \cdots & 0 \\ 0 & \frac{1}{w_2} & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \frac{1}{w_n} \end{bmatrix} \quad ، \quad \Omega^{-1}_{n\times n} = \begin{bmatrix} w_1 & 0 & \cdots & 0 \\ 0 & w_2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & w_n \end{bmatrix}$$

ولكن من النادر ان لم يكن من المستحيل أن تكون مصفوفة (Ω) معلومة وعليه لابد من تقديرها، وعندئذ فأن طريقة (GLS) تسمى (FGLS).

7.1.2 طريقة المربعات الصغرى العمومية العملية (FGLS) [6][7] [8] [9]:

أن هذه الطريقة تقتضي تقدير مصفوفة (Ω) وهي الحالة الأكثر شيوعاً وعليه فأن الصيغة الخاصة بتقدير موجه المعلمات ($\underline{\beta}$) بهذه الطريقة هي:-

$$\underline{\hat{\beta}}_{FGLS} = (X' \ \hat{\Omega}^{-1}X)^{-1}X'\hat{\Omega}^{-1}Y \quad \dots\dots\dots (11)$$

والصيغة الخاصة بتقدير مصفوفة التباين والتباين المشترك للمعلمات هي:

$$\mathrm{Var} - \mathrm{Cov}(\hat{\beta}_{FGLS}) = (X' \ \hat{\Omega}^{-1}X)^{-1}\hat{\sigma}^2_{FGLS} \quad \dots\dots (12)$$

وأن:

$$\hat{\sigma}^2_{FGLS} = \frac{Y'\hat{\Omega}^{-1}Y - \hat{\beta}'_{FGLS}X'Y}{n-k-1} \quad \dots\dots\dots (13)$$

أن خصائص مقدرات (FGLS) والأستدلال الأحصائي الموضوع على أساسها يعتمد وبشكل كبير على تقدير مصفوفة(Ω) ، فأذا كانت $(\widehat{\Omega})$ متسقة (*Consistent*) فأن خصائص (FGLS) هي كالأتي[8][9]:-

1-المقدرات تكون غير متحيزة (*Unbiased*) وكفوءة (*Efficient*) و (*Asymptotically Efficient*) ومتسقة (*Consistent*).

2-الأختبارات الموضوعة (F , t) مقبولة في العينات الكبيرة.

وحتى تكون $(\widehat{\Omega})$ متسقة (*Consistent*) يجب أن يكون هناك توصيف متقن لنموذج عدم تجانس التباين ، وأن تقدير مصفوفة (Ω) يعتمد على طبيعة المشكلة اذا كانت عدم تجانس التباين والأتي التفصيل الخاص بكيفية تقديرها ومعالجتها وكما يأتي:-

أن تقدير طريقة (FGLS) في حالة وجود مشكلة عدم تجانس التباين بالنسبة لموجه المعلمات ($\underline{\beta}$) أو تقدير مصفوفة التباين والتباين المشترك للمعلمات المقدرة بهذه الطريقة وبحسب المعادلات (11) و(12) يقتضي إيجاد تقدير $(\widehat{\Omega})$ و$(\widehat{\Omega}^{-1})$ حيث أن:-

$$\widehat{\Omega}_{n\times n}=\begin{bmatrix}\frac{1}{\widehat{W}_1} & 0 & \cdots & 0\\ 0 & \frac{1}{\widehat{W}_2} & \cdots & 0\\ \vdots & \vdots & \vdots & \vdots\\ 0 & 0 & 0 & \frac{1}{\widehat{W}_n}\end{bmatrix} \quad ، \quad \widehat{\Omega}^{-1}_{n\times n}=\begin{bmatrix}\widehat{W}_1 & 0 & \cdots & 0\\ 0 & \widehat{W}_2 & \cdots & 0\\ \vdots & \vdots & \vdots & \vdots\\ 0 & 0 & 0 & \widehat{W}_n\end{bmatrix}$$

وعليه نجد أن الأوزان تلعب دوراً مهماً وأن عملية تقديرها يجب أن يكون وفق دراسة طبيعة أنموذج عدم تجانس التباين الموجود في أنموذج الأنحدار قيد الدراسة، وهناك العديد من الأفتراضات الموضوعة حول انموذج عدم تجانس التباين والتي تتم في ضوء طبيعة العلاقة مابين (σ_i^2) وأي من المتغيرات التوضيحية (X_i) أو مابين (σ_i^2) و $(\hat{Y}_i)$ فقد تكون هذه العلاقة خطية، أو اسية أو غير ذلك[7].

أن أنموذج عدم تجانس التباين المفترض الأكثر شيوعاً ولاسيما للبيانات في مجال القياسي الأقتصادي هو:

$$\sigma_i^2=\sigma^2\cdot Z_i^{\gamma} \quad \dots\dots\dots (14)$$

حيث أن :

γ : تمثل معيار قوة عدم تجانس $\gamma>0$

Z_i : وقد تمثل المتغير التوضيحي (X_i) أو $(\hat{Y}_i)$ أو أي متغير خارجي .

في ضوء الأنموذج (14) و عندما تكون $(Z_i=X_i)$ فأن الوزن (W_i) يتحدد بالشكل الأتي:

$$W_i=\frac{1}{X_i^{\gamma}} \quad \dots\dots\dots (15)$$

أن عملية تقدير الوزن تعتمد على تقدير (γ) أي أن :

$$\widehat{W}_i = \frac{1}{X_i^{\hat{\gamma}}} \quad .... (16)$$

والطريقة التي يمكن اتباعها لأيجاد $(\hat{\gamma})$ هي :-

طريقة بارك (Parke)[7][6]:

أن هذه الطريقة تفترض العلاقة الدالية بين (σ_i^2) و (X_i) وكالأتي:-

$$\sigma_i^2 = \sigma^2 X_i^{\gamma} e^{u_i} \quad ... ... (17)$$

وعند أخذ اللوغارتم الطبيعي للمعادلة المذكورة آنفاً تكون:

$$\ln \sigma_i^2 = \ln \sigma^2 + \gamma \ln X_i + u_i \quad .... (18)$$

$$u_i \sim N(0 , \sigma_{u_i}^2)$$

وعليه يمكن أستخدام طريقة (OLS) لتقدير المعلمة (γ) حيث أن الصيغة العملية للمعادلة (15)هي:

$$\ln \hat{u}_i^2 = \ln \sigma^2 + \gamma \ln X_i + u_i$$

وأن $(\hat{u}_i)$ تحسب بطريقة (OLS) ويتم ايضاً تقدير المعلمة (γ) كالأتي:

$$\hat{\gamma} = \frac{\sum_{i=1}^{n}(\ln X_i)(\ln \hat{u}_i^2) - \sum_{i=1}^{n}(\ln X_i)\sum_{i=1}^{n}(\ln \hat{u}_i^2)/n}{\sum_{i=1}^{n}(\ln X_i)^2 - \left(\sum_{i=1}^{n}\ln X_i\right)^2/n} \quad ... ... (19)$$

كما يمكن بيان معنوية $(\hat{\gamma})$ المقدرة بطريقة (OLS) هذه من خلال أختبار فرضية العدم :

$$H_0 = \hat{\gamma} = 0$$

وتتصف طريقة بارك بكونها طريقة سهلة وهي الأكثر شيوعاً عندما تكون هناك درجة من التأكد بأن الأفتراض المذكورة آنفاً المتمثل بالمعادلة (17) صحيح.

7.2 **تقدير مصفوفة التباين والتباين المشترك الحصينة** (Robust Covariance Matrix Estimation)[16] [15] [14]:

أن السبب الرئيس وكما ذكرنا آنفاً في عدم أستخدام طريقة المربعات الصغرى (OLS) في حالة وجود مشكلة عدم تجانس التباين على الرغم من وجود بعض الخصائص الجيدة في مقدراتها ، هو عدم أمكانية الأعتماد على الأختبارات الموضوعة اذ يكون الأستدلال الأحصائي غير مقبول وذلك بسبب كون مصفوفة التباين والتباين المشترك للمعلمات المقدرة بطريقة (OLS) متحيزة، وقد امكن تجاوز هذه الثغرة أو هذا النقص من خلال اجراء تعديل على مصفوفة التباين والتباين المشترك هذه وجعلها غير متحيزة وبالشكل الذي يقود الى جعل الأختبارات المستخرجة وفترات الثقة الموضوعة مقبولة.

وأن هذه المصفوفة المعدلة تسمى مصفوفة التباين والتباين المشترك الحصينة **(Robust Covariance Matrix)** لأمكانية استخدامها على الرغم من وجود مشكلة عدم تجانس التباين.

ولقد لاقت هذه الطريقة اهتماماً واسعاً و رواجاً كبيراً منذ عقدين أو اكثر وذلك لسهولتها مقارنة مع طريقة (GLS) التي تتطلب دراسة متقنة لطبيعة أنموذج عدم تجانس و وضع الأفتراضات، حيث أن استخدام مصفوفة التباين والتباين المشترك الحصينة لا تتطلب وضع أي افتراضات، وأن تطبيقها والاعتماد على نتائجها يقتضي أن تكون العينات كبيرة.

وعليه فأن أستخدام مصفوفة التباين والتباين المشترك الحصينة يقوم على ثلاثة مفاهيم أساسية هي:

1. البقاء على مقدرات طريقة المربعات الصغرى(OLS) لأنها تتصف بكونها غير متحيزة(***Unbiased***) و متسقة (***Consistent***).
2. تقدير مصفوفة التباين والتباين المشترك الحصينة للمعلمات المقدرة بطريقة (OLS).
3. أستخدام أو التعامل مع العينات الكبيرة.

أن تقدير مصفوفة التباين والتباين المشترك الحصينة للمعلمات المقدرة بطريقة (OLS)، يكون بحسب نوع البيانات التي يتم التعامل معها وذلك فيما اذا كانت بيانات مقطعية (***Cross-Sectional***) أو بيانات السلاسل الزمنية (***Time Series***) أو بيانات المزدوجة (***Panel Data***) والأتي التفاصيل الخاصة بأيجاد هذه المصفوفة لكل نوع من البيانات.

ولبيان المفهوم الأساسي الذي نقوم عليه عملية تقدير مصفوفة التباين والتباين المشترك الحصينة سواء في البيانات المقطعية أو بيانات الاتي:

$$\underline{Y} = X\underline{\beta} + \underline{u} \quad \dots\dots\dots (20)$$

حيث أن:

Y : موجه المشاهدات $(n \times 1)$

X : مصفوفة $(n \times K)$

β : موجه المعلمات $(K \times 1)$

u : موجه الأخطاء $(n \times 1)$

وبأستخدام طريقة المربعات الصغرى الأعتيادية (OLS) يتم تقدير $(\underline{\beta})$ للأنموذج (20) أعلاه:

$$\underline{\hat{\beta}} = (X'X)^{-1}X'Y \quad \dots\dots\dots\dots (21)$$

وعندما يتحقق الشرط $E(u/X) = 0$

فأنه تقدير غير متحيز والذي يقود مباشرة الى التباين الخاص بهذا التقدير يكون بالصيغة الأتية:

$$Var(\hat{\beta}) = E\left[(\hat{\beta} - \beta)(\hat{\beta} - \beta)'\right] = (X'X)^{-1}X'\Omega X\,(X'X)^{-1} \quad \dots(22)$$

حيث أن:

$E(u'u) = \Omega$ والتي تمثل مصفوفة التباين والتباين المشترك للأخطاء.

وفي حالة أستقلالية الأخطاء وتجانس التباين (*Independently and Identically Distributed*) فأن $\Omega = \sigma^2 I_n$ وعندئذ المعادلة (22) أعلاه تتمثل بالصيغة التقليدية الأتي

$$Var(\hat{\beta}) = \sigma^2 (X'X)^{-1} \quad \text{.... ... ... (23)}$$

وعليه عندما توجد مشكلة عدم تجانس التباين فأن الصيغة (22) سوف تعتمد في تقدير مصفوفة التباين والتباين المشترك الحصينة وذلك من خلال تقدير عناصر القطر الرئيس لمصفوفة (Ω) في حالة وجود مشكلة عدم تجانس التباين.

والأتي التفاصيل الخاصة لتقدير مصفوفة التباين والتباين المشترك الحصينة .

<u>تقدير مصفوفة التباين والتباين المشترك الحصينة في حالة البيانات المقطعية (Cross-Sectional Data and HCCME):</u>

في البيانات المقطعية وكما هو معروف فأن الأنحراف عن الفرضية كون توزيع الأخطاء(*iid*) يتمثل بعدم تجانس التباين، على هذا الأساس اقترح الباحث (*White*) عام (1980) [15] تقدير مصفوفة التباين والتباين المشترك الحصينة في حالة وجود مشكلة عدم تجانس التباين والتي اسماها (مقدر مصفوفة التباين والتباين المشترك المتسقة) لحالة عدم تجانس التباين (Heteroscedasticity-Consistent Covariance Matrix Estimation) والتي يرمز لها (HCCME) وذلك وفق الصيغة الاتية:

$$\widetilde{Var}_h(\hat{\beta}) = (X'X)^{-1} X'\hat{\Omega}X \, (X'X)^{-1} \quad \text{.... ... (24)}$$

حيث أن $(\hat{\Omega})$ في هذه الحالة عبارة عن مصفوفة قطرية، حيث أن عناصر القطر الرئيسي تقدر بواسطة مربع البواقي لطريقة (OLS) وهو $(\hat{u}_i^2)$.

وتعد هذه الصيغة هي الأصلية ويرمز لها (HC_0) وذلك لوجود صيغ أخرى تم أقتراحها،وتتمثل جميعها بأجراء تعديلات على صيغة (HC_0) لتحسين أدائها في العينات المحدودة أو الصغيرة، حيث أقترح الباحث (*Mackinnon And White*) عام (1985) [11] صيغتين هما $(HC_1 \text{ ، } HC_2)$ وكما موضح فيما يأتي:

HC_1 : تم وضعها من قبل الباحثان (*Mackinnon And White*) في ضوء المقترح الموضوع للتباينات في حالة وجود عدم تجانس التباين للباحث (*Hinkley*) عام (1977) [5] وعليه فأن (HC_1) تتمثل بضرب المصفوفة (HC_0) بالمقدار الاتي: $n/(n-K)$ الذي يمثل درجة الحرية المصحح.

HC_2 : تم وضعها في ضوء المقترح الموضوع لتقدير التباينات في حالة عدم تجانس التباين من قبل الباحث (*Horn*) عام (1975) [13]، وعليه فأن صيغة (HC_2) تتضمن تعديل على تقدير عناصر القطر الرئيسي للمصفوفة $(\hat{\Omega})$ فبدلاً من أستخدام $(\hat{u}_i^2)$ يتم أستخدام الاتي:

$$\frac{\hat{u}_i^2}{(1-h_{ii})}$$

حيث أن (h_{ii}) يمثل العنصر (i^{th}) في قطر المصفوفة (*Hat-Matrix*) والتي يرمز لها بـ (H) أو تسمى في مجال القياسي الأقتصادي (*Projection Matrix*) والتي يرمز لها بـ (P) ، وتعرف (H) على انها:

$$H = X(X'X)^{-1}X'$$

وهي مصفوفة صماء (Idempotent) وتتصف بالأتي:

$$HY = \widehat{Y}$$

HC_3 : تم وضعها من قبل الباحثان (*Davidson and Mackinnon*) عام (1993) وذلك على أساس التقدير الموضوع من قبل (Efron) عام (1982) [2] والتي تقود الى أجراء تعديل على تقدير عناصر القطر للمصفوفة $(\widehat{\Omega})$ وكالأتي:

$$\frac{\widehat{u}_i^2}{(1-h_{ii})^2}$$

HC_4 : هذه الصيغة مقترحة من قبل الباحث (Cribari-Neto) عام (2004) [3] والتعديل يتمثل بالأتي:

$$\frac{\widehat{u}_i^2}{(1-h_{ii})^{\delta_i}}$$

حيث أن:

$$\delta_i = min\left(4\,,\frac{h_{ii}}{\overline{H}}\right) \quad ، \quad \overline{H} = \frac{\sum_{i=1}^{n} h_{ii}}{n}$$

HC_5 : وهذه الصيغة أيضاً مقترحة من قبل الباحث (Cribari-Neto) عام (2007) [4] حيث أجرى التعديل الأتي:

$$\frac{\widehat{u}_i^2}{(1-h_{ii})^{\alpha_i}}$$

حيث أن:

$$\alpha_i = min\left[\frac{h_{ii}}{\overline{H}}\,, Max\left(4\,,\frac{kh_{Max}}{\overline{H}}\right)\right]$$

8. الكشف عن مشكلة عدم تجانس التباين[6][7] :

نستعرض في أدناه بعض الطرق الموضوعة للكشف عن هذه المشكلة وكالأتي:

1. أختبار ويت (Whites Test)
2. طريقة العروض البيانية (Graphical Method)
3. أختبار بارك (Park Test)
4. أختبار أرتباط الرتب لسبيرمان (Spearman's Rank Correlation)
5. أختبار بارك – كليجسر(Part-Glejser Test)
6. أختبار كولد فليد – كوانت (Goldfeld-Quandt Test)

7.أختبار باغان (Breusch-Pagan-Godfrey)
8.أختبار كونيكر (Koenker-Bassett Test)

8.1 **أختبار ويت** (Whites Test):

اقترح هذا الاختبار من قبل الباحث (*halbert .white*) في عام (1980)م [6][7] ويعد هذا الأختبار من الأختبارات المستخدمة للكشف عن مشكلة عدم تجانس التباين الخطأ في نموذج الأنحدار، وأن هذا الأختبار يفترض أن (σ_i^2) دالة خطية من المتغيرات التوضيحية (X_s) من الدرجة الثانية أو أكثر ومن حاصل ضربهما (*Cross Products*)، كما أن هذه الطريقة لاتفترض أن (u_i) يتوزع توزيعاً طبيعياً ولبيان خطوات هذه الطريقة نفترض أن لدينا نموذج الأنحدار في معادلة (25) الأتي:

$$Y_i = \beta_0 + \beta_1 X_{1i} + \cdots + \beta_k X_{ki} + u_i \quad \dots\dots\dots \quad (25)$$

ثم يتم أحتساب ما يلي:
1- نجد $(\hat{u}_i)$ بطريقة (OLS).

2-يتم وضع أنموذج انحدار مساعد (*Auxiliary*) وعلى افتراض وجود متغيرين $(X_1 ، X_2)$ فقط في الأنموذج الأصلي وكما مبين أدناه:

$$\hat{u}_i^2 = \alpha_0 + \alpha_1 X_{1i} + \alpha_2 X_{2i} + \alpha_3 X_{1i}^2 + \alpha_4 X_{2i}^2 + \alpha_5 X_{1i} X_{2i} + v_i$$

وتجدد الأشارة الى أن الأنموذج المذكورة آنفاً (*Auxiliary*) يتضمن المعلمة الثابتة (α_0) بغض النظر فيما أذا أن النموذج الأصلي في معادلة (25) يتضمن معلمة ثابتة أو لا.
3- يتم أيجاد (R^2) لأنموذج الأنحدار المساعد (*Auxiliary*) حيث أن:

$$nR^2 \sim \chi^2_{(d\cdot f)}$$

وأن $(d \cdot f)$ هي درجة الحرية التي تمثل عدد المعلمات بأستثناء المعلمة (α_0) وفي النموذج المساعد المفترض فأن عدد المعلمات هي (5) وتقارن قيمة (χ^2) المحتسبة في الخطوة (3) مع قيمة (χ^2) الجدولية فأذا كانت أكبر هناك مشكلة عدم تجانس التباين، أما أذا كانت أصغر فانه لا توجد مشكلة عدم تجانس التباين وأن فرضية العدم هي

$$H_0 : \alpha_1 = \alpha_2 = \cdots = \alpha_5 = 0$$

والجدير بالذكر أن بعض المصادر تسمي هذه الطريقة بأختبار عدم تجانس العام (White's General Heteroscedasticity).

الجانب التجريبي

في هذا الجانب تم توليد بيانات المتغيرات الخاصة بأنموذج أنحدار التي تتصف بوجود مشكلة عدم تجانس التباين بحجوم عينات مختلفة وبتكرار واحد فقط، ولم تكرر التجربة بل تم الاكتفاء بتكرار واحد وذلك لأن التقديرات من عينات هذا التكرار كلها تصبح غير متحيزة ونود ان نبين طبيعة المقارنة في هذا البحث لا تحتاج الى تكرار ايضاً، وبعد ذلك سيتم تطبيق طريقة (FGLS) وطرائق تقدير مصفوفة التباين والتباين المشترك الحصينة (HCCME) على مجموعة البيانات التي تتصف بوجود مشكلة عدم تجانس التباين، مع أجراء مقارنات بين هذه الطرائق.

وذلك بأستخدام برنامج (Statistica-Version 10) أخذين بنظر العناية أن العينة المطلوبة (n) تتكون من عينتين جزئيتين، العينة الجزئية الأولى (n_1) والعينة الجزئية الثانية (n_2) والأتي الخطوات الخاصة بتوليد بيانات كلاً من العينة n_1 و n_2 :-

أولاً : العينة الجزئية الأولى (n_1) QUOTE (n_1)

1.توليد بيانات الخطأ العشوائي (u_{i1}) والذي أفترض أنه يتوزع توزيعاً طبيعياً بـ

$$\sigma^2_{u_1} = 0 \cdot 1825741644130 \quad ، \quad \mu_1 = 0$$

2.توليد بيانات المتغيرات التوضيحي (X_{i1}) والذي أفترض أنه يتوزع توزيعاً بـ

$$\sigma^2_{x_1} = 2 \cdot 63523138347 \quad ، \quad \mu_1 = 7 \cdot 5$$

3.أعتماد أنموذج الأنحدار بالمعلمات الحقيقية $\beta_{11} = 0 \cdot 276000$ ، $\beta_{01} = 0 \cdot 600000$ والأنموذج هو:

$$Y_{i1} = \beta_{01} + \beta_{11} X_{i1} + u_{i1} \qquad \dots\dots\dots (26)$$

4.أعتماد القيم المولدة والخاصة بـ (u_{i1}) ، (X_{i1}) في الأنموذج (26) بالمعلمات الحقيقية، لأيجاد المتغير المعتمد (Y_{i1}) .

ثانياً: العينة الجزئية الثانية (n_2) QUOTE (n_2)

1.توليد بيانات الخطأ العشوائي (u_{i2}) والذي أفترض أنه يتوزع توزيعاً طبيعياً بـ

$$\sigma^2_{u_2} = 0 \cdot 4742244837393 \quad ، \quad \mu_2 = 0$$

2.توليد بيانات المتغيرات التوضيحي (X_{i2}) والذي أفترض أنه يتوزع توزيعاً بـ

$$\sigma^2_{x_2} = 2 \cdot 63523138347 \quad ، \quad \mu_2 = 17 \cdot 5$$

3.أعتماد أنموذج الأنحدار بالمعلمات الحقيقية $\beta_{11} = 0 \cdot 2000000$ ، $\beta_{01} = 1 \cdot 540000$ والأنموذج هو:

$$Y_{i2} = \beta_{01} + \beta_{11} X_{i2} + u_{i2} \qquad \dots\dots\dots (27)$$

4. أعتماد القيم المولدة والخاصة بـ (u_{i2}) ، (X_{i2}) في الأنموذج (27) بالمعلمات الحقيقية، لأيجاد المتغير المعتمد (Y_{i2}) .

ثم يتم جمع بيانات العينة الجزئية الأولى و الثانية لأيجاد العينة المطلوبة، $n = n_1 + n_2$.

وبتطبيق الخطوات المذكورة آنفاً وبأستخدام البرنامج الجاهز(Statistica) تم توليد (6) عينات، هي المتوسطة $(n_1 = 50)$ ، $(n_2 = 80)$ والكبيرة $(n_3 = 100)$ ، $(n_4 = 150)$ والكبيرة جداً ، $(n_5 = 300)$ ، $(n_6 = 400)$.

والجدير بالذكر أنه وكما ذكرنا بالجانب النظري أن الأستدلال الأحصائي لطريقة (FGLS) وكذلك طرائق (HCCME) ، تقتضي التعامل مع البيانات الكبيرة، وحيث أن العديد من المصادر اشارت الى أن حجم العينة يعد صغيراً أذا كان أقل من (40) بصورة عامة ونظراً لخصوصية هذه الطرائق تم البدء بعينة محددة $(n = 50)$.

وللتأكد من أن العينات التي تم توليدها تتصف بوجود مشكلة عدم تجانس التباين تم تطبيق أختبارين هما (Koenker) وأختبار (White's) وكانت النتائج كما مبينة في الجدول رقم (1).

جدول رقم (1)

نتائج اختبارات عدم تجانس التباين لمستوى معنوي $(\alpha = 0 \cdot 05)$

حجم العينة	أختبار(White' s)	
n	χ^2	*p*
50	6.93276	0.0312298
80	15.728	0.000384325
100	13.9581	0.000931203
150	17.5642	0.000153455
200	26.6277	1.65143E-006
300	46.4059	8.37689E-011
400	50.918	8.77618E-012

وبمقارنة قيم الأحتمال (P) المستحصلة في الجدول المذكورة وللأختبارين مع مستوى معنوية $(\alpha = 0 \cdot 05)$ نجد أن $(P < 0 \cdot 05)$ وبذلك ترفض فرضية العدم (H_0) وهناك عدم تجانس في التباين لكافة العينات المولدة.

وبهدف تطبيق طريقة (FGLS) على العينات المولدة لا بد من تقدير الأوزان $(\widehat{w}_i)$ ، والتي أقتضت تقدير معيار قوة عدم التجانس $(\widehat{\gamma})$ وذلك في ضوء الأفتراض الموضوع لأنموذج عدم تجانس التباين لأسلوب بارك (*Park*) والموضوع في الجانب النظري معادلة (10) حيث تم وفق هذا الأفتراض وبأستخدام البرنامج الجاهز (Statistica) تقدير وأختبار (γ) وكما مبين في الجدول رقم (2).

جدول رقم (2) نتائج تقدير وأختبار

معيار قوة عدم تجانس التباين لمستوى معنوي $(\alpha = 0 \cdot 05)$

n	γ	*t*	*P -value*
50	1.94452	3.01464	0.004102
80	1.65236	3.32	0.001332
100	1.26951	3.60977	0.000485
150	1.26324	3.77124	0.000234
300	1.07063	4.1970	0.000001
400	1.03245	5.0783	0.000001

ومن ملاحظة الجدول رقم (2) نجد أن تقديرات الـ $(\hat{\gamma})$ كلها معنوية وذلك عند مقارنة قيم الأحتمال المحتسبة (P) مع $(\alpha = 0\cdot05)$ نجد أن $(P < 0\cdot05)$ مما يشير الى أن هذا المعيار معنوي، لوجود معنوي لمشكلة عدم تجانس التباين.

ولقد تم أستخدام $(\hat{\gamma})$ المقدرة لأيجاد الأوزان $(\hat{w}_i)$، ومن ثم أستخدام هذه الأوزان لتطبيق طريقة (FGLS) لوضع التقديرات المطلوبة وأختبارها.

كما تم تطبيق طرائق $(HC_0 ، HC_1 ، HC_2 ، HC_3 ، HC_{3a})$ على بيانات العينات المولدة وكانت النتائج التي تم الحصول عليها كما موضحة في الجداول الاتية:-

جدول رقم (3)

نتائج طرائق التقدير للعينات المتوسطة

Size	*Method*	γ	□	*S.E*	*t*	*P-value*
n_1=50	*FGLS*	1.94452	0.694558 0.260194	0.087043 0.009739	7.979 26.71	0.000000 0.000000
	HC_0	–	0.884145 0.243456	0.100957 0.00934716	8.758 26.05	1.62E-011 5.69E-030
	HC_1	–	0.884145 0.243456	0.103039 0.00953990	8.581 25.52	2.96E-011 1.42E-029
	HC_2	–	0.884145 0.243456	0.104021 0.00966323	8.500 25.19	3.91E-011 2.52E-029
	HC_3	–	0.884145 0.243456	0.107222 0.00999344	8.246 24.36	9.38E-011 1.13E-028
	HC_{3a}	–	0.884145 0.243456	0.106144 0.00989298	8.330 24.61	7.03E-011 7.20E-029
n_2=80	*FGLS*	1.65236	0.746119 0.245745	0.068793 0.006933	10.845 35.446	0.000000 0.000000
	HC_0	–	0.867849 0.240056	0.0820144 0.00886189	10.58 27.09	9.74E-017 1.99E-041
	HC_1	–	0.867849 0.240056	0.0830592 0.00897479	10.45 26.75	1.75E-016 4.86E-041
	HC_2	–	0.867849 0.240056	0.0834385 0.00901424	10.40 26.63	2.15E-016 6.62E-041
	HC_3	–	0.867849 0.240056	0.0848916 0.00916953	10.22 26.18	4.71E-016 2.20E-040
	HC_{3a}	–	0.867849 0.240056	0.0843593 0.00911204	10.29 26.34	3.55E-016 1.41E-040

جدول رقم (4)
نتائج طرائق التقدير للعينات الكبيرة

Size	Method	γ	□	S.E	t	P-value
	FGLS	1.26951	0.654314 0.256095	0.050810 0.005483	12.87 46.70	0.000000 0.000000
	HC_0	–	0.838803 0.240500	0.0581141 0.00583361	14.43 41.23	5.41E-026 1.02E-063
n_3=100	HC_1	–	0.838803 0.240500	0.0587042 0.00589284	14.29 40.81	1.06E-025 2.59E-063
	HC_2	–	0.838803 0.240500	0.0589763 0.00591113	14.22 40.69	1.44E-025 3.45E-063
	HC_3	–	0.838803 0.240500	0.0598545 0.00598997	14.01 40.15	3.83E-025 1.18E-062
	HC_{3a}	–	0.838803 0.240500	0.0595544 0.00595995	14.08 40.35	2.75E-025 7.39E-063
	FGLS	1.26324	0.819156 0.242593	0.037485 0.004050	21.85 59.85	0.000000 0.000000
n_4=150	HC_0	–	1.01675 0.226332	0.0660063 0.00634627	15.40 35.66	1.48E-032 1.49E-074
	HC_1	–	1.01675 0.226332	0.0664507 0.00638901	15.30 35.43	2.74E-032 3.62E-074
	HC_2	–	1.01675 0.226332	0.0667147 0.00641255	15.24 35.30	3.93E-032 5.89E-074
	HC_3	–	1.01675 0.226332	0.0674323 0.00647965	15.08 34.93	1.03E-031 2.33E-073
	HC_{3a}	–	1.01675 0.226332	0.0672071 0.00645802	15.13 35.05	7.64E-032 1.50E-073

جدول رقم (5)
نتائج طرائق التقدير للعينات الكبيرة جداً

Size	Method	γ	□	S.E	t	P-value
n_5=300	FGLS	1.07063	0.681258 0.253159	0.028404 0.002902	23.98 87.23	0.000000 0.000000
	HC_0	–	0.922264 0.233514	0.0437733 0.00433906	21.07 53.82	5.64E-061 1.56E-155
	HC_1	–	0.922264 0.233514	0.0439199 0.00435359	21.00 53.64	1.02E-060 3.85E-155
	HC_2	–	0.922264 0.233514	0.0440264 0.00436202	20.95 53.53	1.58E-060 6.49E-155
	HC_3	–	0.922264 0.233514	0.0442814 0.00438514	20.83 53.25	4.39E-060 2.70E-154
	HC_{3a}	–	0.922264 0.233514	0.0442075 0.00437783	20.86 53.34	3.27E-060 1.72E-154
n_6=400	FGLS	1.03245	0.73055 0.249648	0.025341 0.002520	28.82 99.06	0.000000 0.000000
	HC_0	–	0.895791 0.236386	0.0348131 0.00335030	25.73 70.56	1.09E-086 4.26E-227
	HC_1	–	0.895791 0.236386	0.0349005 0.00335870	25.67 70.38	2.02E-086 1.07E-226
	HC_2	–	0.895791 0.236386	0.0349444 0.00336187	25.63 70.31	2.77E-086 1.52E-226
	HC_3	–	0.895791 0.236386	0.0350764 0.00337349	25.54 70.07	7.04E-086 5.41E-226
	HC_{3a}	–	0.895791 0.236386	0.035025 0.00336927	25.57 70.16	5.16E-086 3.41E-226

ومن ملاحظة الجداول رقم (3) و(4) و (5) نجد ما يلي:-

1. أن أداء طريقة (FGLS) كان هو الأفضل حيث أنها حققت أقل خطأ معياري ولكافة أحجام العينات عدا عينة حجم (50) وبالنسبة للمعلمة (β_1) ، وهذا متوقع جداً لكون طبيعة عدم التجانس في التباين قد تم تعيينها بدرجة عالية من التأكد من خلال التوليد، وبأحجام عينات ملائمة مما جعلها تعطي أفضل النتائج ومن ثم أمكانية أستخدامها كمعيار أساسي يحدد في ضوءها الأداء الأفضل لطرائق (HCCME).
2. عند مقارنة الطرائق $(HC_0 ، HC_1 ، HC_2 ، HC_3 ، HC_{3a})$ مع أداء طريقة (FGLS) نجد أن طريقة (HC_0) هي الأفضل لكونها الأكثر قرباً من (FGLS) حيث أن الخطأ المعياري المقدر لها هو الأقل من بين طرائق (HCCME) الأخرى والأقرب الى (FGLS)، وهذا الأمر ينسحب على كافة أحجام العينات عدا عينة حجم (50) وبالنسبة للمعلمة (β_1) فقط ، ويمكن تفسير ذلك أن أداء (FGLS) ليس جيداً عند حجم (50) لأن هذه الطريقة تتطلب أحجام عينات كبيرة حتى تتحقق الصفات المرغوبة في المقدرات وكما ذكرنا في الجانب النظري.
3. كلما يزداد حجم العينة وخاصة (300) ،(400) فأن أداء طرائق (HCCME) تقترب من بعضها البعض والفروقات ضئيلة.
4. أجمالاً نجد أن كافة طرائق (HCCME) تتطابق مع طريقة (FGLS) في الأستدلال الأحصائي حيث أن كلاً منها يعطي النتائج المتمثلة برفض فرضية العدم H_0 .

الاستنتاجات والتوصيات

الأستنتاجات (Conclusions):

في ضوء مجريات هذا البحث ان استخدام أنموذج انحدار بسيط او متعدد له ذات الانعكاسات لمشكلة عدم تجانس التباين على مقدرات النموذج والاستدلال الاحصائي الخاص به، كما ان طريقة (FGLS) تعتمد على الاوزان وادائها واحد سواء في الأنموذج البسيط او المتعدد فيمكن تعميم النتائج وهذا الاجراء متبع من قبل العديد من الباحثين منهم الباحث (Kmenta)[10]، وقد تم التوصل الى الأستنتاجات الأتية:

1. أن أداء طريقة (FGLS) كان هو الأفضل حيث أنها حققت أقل خطأ معياري ولكافة أحجام العينات عدا عينة حجم (50) وبالنسبة للمعلمة (β_1) ، وهذا متوقع جداً لكون طبيعة عدم التجانس في التباين قد تم تعيينها بدرجة عالية من التأكد من خلال التوليد، وبأحجام عينات ملائمة مما جعلها تعطي أفضل النتائج ومن ثم أمكانية أستخدامها كمعيار أساسي يحدد في ضوئها الأداء الأفضل لطرائق (HCCME).
2. عند مقارنة الطرائق $(HC_0$ ، HC_1 ، HC_2 ، HC_3 ، $HC_{3a})$ مع أداء طريقة (FGLS) نجد أن طريقة (HC_0) هي الأفضل لكونها الأكثر قرباً من (FGLS) حيث أن الخطأ المعياري المقدر لها هو الأقل من بين طرائق (HCCME) الأخرى والأقرب الى (FGLS)، وهذا الأمر ينحسب على كافة أحجام العينات عدا عينة حجم (50) وبالنسبة للمعلمة (β_1) فقط ، ويمكن تفسير ذلك أن أداء (FGLS) ليس جيداً عند حجم (50) لأن هذه الطريقة تتطلب أحجام عينات كبيرة حتى تتحقق الصفات المرغوبة في المقدرات وكما ذكرنا في الجانب النظري.
3. كلما يزداد حجم العينة وخاصة (300)،(400) فأن أداء طرائق (HCCME) تقترب من بعضها البعض والفروقات ضئيلة.
4. أجمالاً نجد أن كافة طرائق (HCCME) تتطابق مع طريقة (FGLS) في الأستدلال الأحصائي حيث أن كلاً منها يعطي النتائج المتمثلة برفض فرضية العدم H_0 .

التوصيات (Recommendations):

1. في حالة البيانات المقطعية و وجود مشكلة عدم تجانس التباين وبالشكل الذي يمكن فيه تحديد نموذج عدم تجانس التباين وهناك درجة معينة من التأكد من طبيعة عدم التجانس موجود، فأنه يمكن الأعتماد على طريقة (FGLS) لان أداءها في التقدير متفوقاً بالنسبة لطرائق التقدير (HCCME) وان كان الأستدلال الأحصائي متطابق معها.
2. في حالة عدم التأكد من طبيعة عدم تجانس التباين الموجود وعدم أمكانية تحديد النموذج الخاص به يمكن الأعتماد على طرائق التقدير (HCCME) كبديل مناسب جداً في تقدير مصفوفة التبايا والتباين المشترك الحصينة ووضع الأستدلال الأحصائي مقبول.
3. يمكن الأعتماد على طريقة (HC_0) من طرائق تقدير (HCCME) تحديداً وذلك لأن أداءها متفوقاً على الطرائق الأخرى وعند أزدياد حجم العينة بشكل كبيراً فأن طرائق تقدير (HCCME) لا يوجد أختلاف كبير بينها.

المصادر

[1]- Amemiya, Takeshi (1985). "Generalized Least Squares Theory". Advanced Econometrics. Harvard University Press. ISBN 0-674-00560-0.

[2]- B. Efron, The Jackknife, the Bootstrap, and Other Resampling Plans, Society for Industrial and Applied Mathematics, Philadelphia, 1982.

[3]- Cribari- Neto F. "Asymptotic inference under heteroskedasticity of unknown form", Computational Statistics and Data Analysis, Vol. 45, pp. 215{233, 2004.

[4]- Cribari- Neto F., T.C. Souza, K.L.P. Vasconcellos, ''Inference under heteroskedasticity and leveraged data", Communications in StatisticsTheory and Methods/, Vol. 36, pp. 1877{1888, 2007.

[5]- D.V. Hinkley, " Jackknifing in unbaanced situations", Technometrics, Vol. 19, pp. 285{292, 1977.

[6]- Gujarati, D. N. and Porter, D. C. (2014)," Business & Economics. " See, Chapter 11, op cit., pp. 365–411.

[7]- Greene, W. H. (2003). Econometric Analysis Chapter 10 &11, pp.198-240 (5th ed.) Upper Saddle River, NJ: Prentice Hall.

[8]- Hansen, Christian B. (2007). "Generalized Least Squares Inference in Panel and Multilevel Models with Serial Correlation and Fixed Effects". Journal of Econometrics 140 (2): 670–694

[9]- Johnston, John (1972). "Generalized Least-squares". Econometric Methods (Second ed.). New York: McGraw-Hill. pp. 208–242

[10]-Kmenta, J. ,1971,''Elements of Econometrics''. New York: Macmillan.

[11]-Mackinnon, J.G., White, H., (1985)," Some heteroskedasticity–consistent covariance matrix estimator with improved finite sample properties". J. Econometrics 29, 305-325.

[12]-Newey, W.K. and West, K.D., (1987),"A simple positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix.", Economatrica 55, 703-708.

[13]- S.D. Horn, R.A. Horn, D.B. Duncan, " Estimating heteroscedastic variances in linear models", Econometrica, Vol. 60, pp. 539{547, 1975.

[14]- Startz, R. (2015)," Bayesian Heteroskedasticity-Robust Regression." Department of Economics, 2127 North Hall, University of California, Santa Barbara, CA 93106, email: startz@econ.ucsb.edu. Advice from Sid Chib, Ed Greenberg, and Doug Steigerwald is gratefully acknowledged

[15]-White, H., (1980), "A heteroscedasticity- consistent covariance estimator and a direct test for heteroscedasticity" econometrica 48, 817-838

[16]-Zeileis, A. (2004), " Econometric computing with HC and HAC covariance matrix estimators" Journal of Statistical Software, November 2004, Volume 11, Issue 10

A Comparative Study of Some Methods of Estimating Robust Variance Covariance Matrix of the Parameters Estimated by (OLS) in Cross-Sectional Data

Abstract

The Classical Normal Linear Regression Model Based on Several hypotheses, one of them is Heteroscedasticity as it is known that the wing of least squares method (OLS), under the existence of these two problems make the estimators, lose their desirable properties, in addition the statistical inference becomes unaccepted table. According that we put tow alternative, the first one is (Generalized Least Square) Which is denoted by (GLS), and the second alternative is to (Robust covariance matrix estimation) the estimated parameters method(OLS), and that the way (GLS) method neat and certified, if the capabilities (Efficient) and the statistical inference Thread on the basis of an acceptable but this method requires knowledge and knowledge of the nature of the problem and the private model of the problem, whether the Heteroscedasticity otherwise, the method (GLS) become inappropriate. The second alternative is a matrix contrast common variation fortified it does not require prior knowledge of the nature of your problem model, it's also an easy way and by this method has met with popular and interest in more than two decades by researchers, that the estimated of Robust covariance matrix the estimated parameters method(OLS), shall be according to data that is handled type, and we have dealt with in this study, cross-sectional data where the problem of Heteroscedasticity in contrast may be contained therein(Heteroscedasticity- Consistent Covariance matrix Estimation) and symbolized by the (HCCME) and includes many ways, and we have dealt with in our study of these methods $(HC_0 ، HC_1 ، HC_2 ، HC_3 ، HC_{3a})$.

Keywords: **Heteroscedasticity, Robust, Consistent, Efficient, Unbiased, Regression, Biased, Estimator, Covariance .**