



## Detecting The Problem of Heteroscedasticity in Cross-Sectional Data

الكشف عن مشكلة عدم ثبات تباين حد الخطأ في البيانات المقطعية

Wafaa A. B. M. Eid and Rania F. Mahmoud

Cent. Lab. for Design & Stat. Analysis Res., A. R. C., Giza, Egypt.

DOI: [10.21608/JALEXU.2024.256059.1181](https://doi.org/10.21608/JALEXU.2024.256059.1181)



### Article Information

Received: December 14<sup>th</sup> 2023

Revised: December 29<sup>th</sup> 2023

Accepted: December 30<sup>th</sup> 2023

Published: December 31<sup>th</sup> 2023

**ABSTRACT:** The current research aims to study the nature of the problem of heteroscedasticity

{instability of the error term variation in the regression analysis}, its causes and effects and to apply the most important methods for detecting this problem in the cross-sectional data. Field data were collected during the 2022/2023 agricultural season from lupine farmers in Al-Husseiniyah District, Sharkia Governorate through a questionnaire prepared for this purpose. Two samples were randomly selected, the first one included 25 farmers and the second included 102 farmers. The analysis was done using the Eviews computer program.

Tests were conducted to detect the presence of a problem of instability of the error term variation in the regression analysis for production and production factors (amount of seeds, quantity of phosphate fertilizer, number of workers, quantity of irrigation water) for the lupine crop. It was found that there is a problem of instability of the error term variation for quantity of fertilizer and amount of seeds in the 25 farmers sample. In the case of the 102 farmers sample, there was a problem of instability for amount of phosphate fertilizer and quantity of irrigation water. It was sufficient to explain only one element (amount of fertilizer) as an example to illustrate how to detect the problem using the collected data.

In case of the 25 farmers sample, by using Park's test, the calculated T value was about 2.084 which is greater than the tabulated value that is of about 2.069, so the alternative hypothesis of the existence of the problem of instability of the error term variation is accepted. Data were processed by performing a transformation of the original model using the weighted least squares method. The calculated value of T was about 0.329, which is less than the tabulated value, so the null hypothesis is accepted where the variance of the error term is constant.

In case of the 102 farmers sample, by using the Park test, the calculated T value was about 2.005 which is greater than the tabulated T value which was about 1.984. The alternative hypothesis of the existence of the instability of error variance is accepted. After treatment, the calculated T value reached about 1.961, which is less than that of tabular T value. Therefore, the null hypothesis that the variance of the error term is stable is accepted. From the previous results, it is cleared that there is no relationship between sample size and presence or absence of the problem of variance instability of the error term in regression analysis using cross-sectional data.

### Recommendations:

The study recommends that researchers should pay attention to conducting the necessary tests to detect the problem of {heteroscedasticity} and the need to detect it and put the remedy to obtain accurate results.

**Keywords:** Heteroscedasticity - Cross section Data -Park Test -Goldfeld Quanat -Spearman

تقدير قيمة المتغير التابع عند قيم معينة للمتغيرات المستقلة وقياس الخطأ في التقدير، بحيث يكون تقدير النموذج الإحصائي في حقيقة الأمر هو عبارة عن محاولة للوصول إلي تقديرات دقيقة لقيم معاملاته  $B'S$ ، ولكي تتم عملية التقدير هذه يلزم تجميع البيانات عن المتغيرات المستقلة  $B'S$  والمتغير التابع  $Y$  وإعدادها للتحليل الإحصائي باستخدام أسلوب تحليل الانحدار، كما يجب التأكد من عدم وجود مشاكل في القياس باستخدام اختبارات الإحصاء القياسي  $ECONOMETRICS$ .

### المشكلة:

يواجه الباحثين عدة مشاكل عند تقدير النماذج الإحصائية والقياسية ومن أبرزها مشكلة عدم ثبات تباين حد الخطأ (HETEROSCEDASTICITY) بمعنى وجود ارتباط

### مقدمة:

يعتبر الأسلوب الإحصائي أحد الأساليب العلمية التي تمكن الباحث من الوصول إلي القرارات المثلي، في ظل الإمكانات المادية المتوفرة لديه وتقييم الآثار المترتبة علي هذه القرارات، لذلك يعتبر الوصف الإحصائي الدقيق والتحليل الإحصائي السليم للعلاقة بين المتغيرات الإحصائية المختلفة، من أهم المؤشرات التي يعتمد عليها الباحث في مجالات العلوم الزراعية، في اتخاذ أي قرار صحيح يستهدف توصية ما مرتبطة بحل المشكلة البحثية، وعليه فإن النماذج الإحصائية هي الصياغة الرياضية للعلاقة بين المتغيرات المستقلة والتابعة موضع البحث المبنية علي البيانات الإحصائية، وتمثل نماذج الانحدار جزء هام في دراسة العلاقة بين المتغيرات، وعند استخدام نماذج الانحدار يمكن

من بين افتراضات نموذج الانحدار الخطي باستخدام طريقة المربعات الصغرى العادية هو ثبات تباين حد الخطأ *HOMOSCEDASTICITY* ويترتب علي إسقاط هذا الافتراض حدوث مشكلة عدم ثبات تباين حد الخطأ *HETEROSCEDASTICITY* أي أن حدود الأخطاء ليس لها نفس التباين، مفهوم مشكلة عدم ثبات التباين في تغير تباين الحد العشوائي مع تغير قيم المتغير التفسيري حيث يلاحظ أن تغير المتغير التفسيري  $\alpha X_i$  يؤدي لتغير المتغير التابع  $Y_i$  ويؤدي أيضاً لتغير تباين الحد العشوائي، ومع وجود هذه المشكلة، فإن المعلمات المقدرة باستخدام طريقة المربعات الصغرى العادية تتصف بعدم الكفاءة وإن كانت تتصف بعدم التحيز والاتساق، ويرتبط عدم ثبات تباين حد الخطأ بالبيانات المقطعية (CROSS SECTION DATA -) أكثر من بيانات السلاسل الزمنية (TIME - SERIES DATA).

#### أسباب عدم ثبات تباين حد الخطأ

توجد عدة أسباب لعدم ثبات تباين حد الخطأ إلا انه يمكن حصرها في الأسباب التي تؤدي إلي زيادة تباين الأخطاء ومنها ما يلي:

- 1- يمكن أن ينشأ اختلاف التباين *HETEROSCEDASTICITY* نتيجة لوجود قيم متطرفة *OUTLIER* وهي قيم إما صغيرة جداً أو قيم كبيرة جداً، إن إدراج أو استبعاد مثل تلك القيم، خاصة إذا كان حجم العينة صغيراً، يمكن أن يغير نتائج تحليل الانحدار بشكل كبير.
- 2- تنشأ أيضاً مشكلة اختلاف التباين من عدم توصيف النموذج المقدر بشكل صحيح على سبيل المثال حذف بعض المتغيرات الهامة من النموذج، وإذا أضيفت تلك المتغيرات المحذوفة للنموذج فإن مشكلة اختلاف التباين تختفي.
- 3- عدم ملائمة أساليب جمع البيانات، وعدم كفاية مفردات العينة وانخفاض كفاءة جامعي البيانات.

#### طرق الكشف عن مشكلة اختلاف التباين

##### *HETEROSCEDASTICITY*

تستخدم عدة اختبارات للكشف عن مشكلة اختلاف التباين منها:

##### 1- اختبار بارك *PARK TEST*

يفرض أن:

$$Y_i = \alpha + \beta X_i + \epsilon_i$$

$$I=1,2,\dots,N$$

حيث أن:

بين الحد العشوائي والمتغير التفسيري التي تكون أكثر شيوعاً بالبيانات المقطعية *CROSS SECTION DATA* عن بيانات السلاسل الزمنية *TIME SERIES DATA* ويجب علي الباحث معالجتها قبل القيام بتفسير نتائج النموذج الإحصائي أو القياسي المقدر، لما لها من آثار سلبية على نتيجة التحليل الإحصائي، فعلى الرغم من أن مقدرات النموذج *PARAMETERS* تظل غير متحيزة *UNBIASED* ومتسقة *CONSISTENT* إلا أنها لم تعد لها أقل تباين أي لم تعد كفاء *EFFICIENT* وبمعنى آخر لم تعد تلك المقدرات تمتاز أفضل مقدر خطي غير متحيز "BLUE" مما لاشك يؤثر على دقة النتائج المتحصل عليها وبالتالي تؤدي لقرارات خاطئة لا تتناسب مع المشكلة التي يسعى الباحث لحلها وتؤثر على جودة الأداء العام للنموذج.

#### الهدف:

يهدف البحث إلي دراسة طبيعة مشكلة عدم ثبات تباين حد الخطأ *HETEROSCEDASTICITY*، وأسبابه وآثاره وتطبيق أهم طرق الكشف عن عدم ثبات تباين حد الخطأ في البيانات موضوع البحث حتى يتسنى معالجة هذه المشكلة وإيجاد تقديرات إحصائية ذات كفاءة عالية واختيارها إحصائياً بعد التخلص من هذه المشكلة مع توضيح طرق المعالجة الإحصائية المختلفة.

#### الطريقة البحثية ومصادر البيانات

يعتمد البحث بصفة أساسية علي أساليب التحليل الكمي للكشف عن مشكلة عدم ثبات تباين حد الخطأ من خلال عدة طرق من أهمها اختبار بارك *PARK TEST*، اختبار *GOLDFELD-QUANDT*، اختبار معامل ارتباط الرتب ل *SPEARMAN*، ثم تطبيق الطرق المختلفة لمعالجة عدم ثبات تباين حد الخطأ علي الإنتاج وعناصر الإنتاج (كمية التقاوي)، كمية السماد الفوسفاتي، عدد العمال، كمية مياه الري) ، وقد تم الاعتماد على البيانات الميدانية التي تم تجميعها خلال الموسم الزراعي 2023/2022 من بعض مزارعي الترمس بمحافظة الشرقية عن طريق استمارة استبيان أعدت خصيصاً لهذا الغرض، حيث تم اختيار عينتين مختلفتين في حجم العينة للكشف عن وجود أو عدم وجود مشكلة عدم ثبات تباين حد الخطأ في حالة العينة الصغيرة حجمها 25 مزارع وعينة كبيرة حجمها 102 مزارع بطريقة عشوائية في مركز الحسينية بمحافظة الشرقية، وقد تم التحليل باستخدام برنامج *EViews*.

#### النتائج البحثية والمناقشة:

#### أولاً: الإطار النظري والتحليلي:

#### طبيعة عدم ثبات تباين حد الخطأ

$\hat{b}$  معنوية احصائياً، ويدل هذا علي وجود مشكلة عدم ثبات تباين حد الخطأ.

- اما اذا كانت قيمة T المحسوبة اقل من قيمة T الجدولية، يتم قبول فرض عدم القائل فإن  $\hat{b} = 0$ ، ويقال في هذه الحالة ان  $\hat{b}$  غير معنوية احصائياً، ويدل علي وجود افتراض ثبات تباين حد الخطأ.

## 2- اختبار GOLDFELD-QUANDT:

بافتراض النموذج التالي:

$$I=1,2,\dots,N \quad Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i$$

يمكن بيان كيفية استخدام اختبار GOLDFELD-QUANDT في اكتشاف عدم ثبات تباين الخطأ من خلال الخطوات التالية:

1- ترتيب مشاهدات X ترتيباً تصاعدياً.

2- استبعاد المشاهدات الوسطى لكل من X و Y، ثم تكوين مجموعتين من المشاهدات بحيث يكون لكل مجموعة على حدا معادلة خاصة بها كما يلي :

- المجموعة الأولى: وتتمثل في المشاهدات الخاصة بكل من X و Y الواردة قبل المشاهدات التي تم استبعادها، والمعادلة الخاصة بهذه المجموعة هي :

$$Y_{1i} = a + bX_{1i} + \varepsilon_{1i}$$

- المجموعة الثانية: وتتمثل في المشاهدات الخاصة بكل من X و Y الواردة بعد المشاهدات التي تم استبعادها، والمعادلة الخاصة بهذه المجموعة هي:

$$Y_{2i} = c + dX_{2i} + \varepsilon_{2i}$$

3- تقدير معاملات المعادلتين السابقتين باستعمال طريقة المربعات الصغرى :

$$\hat{Y}_{1i} = \hat{a} + \hat{b}X_{1i} + e_{1i}$$

$$Y_{2i} = \hat{c} + \hat{d}X_{2i} + e_{2i}$$

4- الحصول على القيم المقدرة لحد الخطأ:

$$\hat{\varepsilon}_{1i} = Y_{1i} - \hat{Y}_{1i}$$

$$\hat{\varepsilon}_{2i} = Y_{2i} - \hat{Y}_{2i}$$

5- إيجاد القيمة المحسوبة لإحصائية F كما يلي:

$$\hat{F} = \frac{\sum \hat{\varepsilon}_{2i}^2}{\sum \hat{\varepsilon}_{1i}^2}$$

6- إيجاد درجات الحرية:

$$DF = \frac{n - m - 2(k + 1)}{2}$$

حيث أن:

$N =$  عدد المشاهدات،  $K =$  عدد المتغيرات المستقلة،  $M =$  عدد المشاهدات المستبعدة.

$Y =$  القيمة الفعلية للمتغير التابع،  $X =$  القيمة الفعلية للمتغير المستقل،  $\varepsilon =$  القيمة الفعلية لحد الخطأ،  $\alpha, \beta =$  القيم الفعلية لمعاملات الانحدار،  $N =$  عدد المشاهدات.

يمكن بيان كيفية استخدام اختبار PARK في اكتشاف عدم ثبات تباين حد الخطأ من خلال الخطوات التالية:

1- تطبيق طريقة المربعات الصغرى العادية فينتج ما يلي:

$$\hat{Y}_i = \hat{\alpha} + \hat{\beta}X_i + e_i$$

حيث أن:

$\hat{Y} =$  القيمة المقدرة للمتغير التابع،  $\hat{\alpha} =$  القيمة المقدرة ل  $\alpha$ ،

$\hat{\beta} =$  القيم المقدرة ل  $\beta$ ،  $e =$  القيمة المقدرة لحد الخطأ  $\varepsilon$ .

2- الحصول علي البواقي أو القيم المقدرة لحد الخطأ ( $e_i$ ) من المعادلة التالية:

$$e_i = Y_i - \hat{Y}_i$$

3- استخدام طريقة المربعات الصغرى العادية في إجراء انحدار  $\ln e_i^2$  علي  $\ln X_i$  فينتج ما يلي:

$$\ln \hat{e}_i^2 = \hat{\alpha} + \hat{\beta} \ln X_i$$

4- إيجاد القيمة المحسوبة لاختبار T بالنسبة ل  $\hat{\beta}$  كما يلي:

$$T(\hat{\beta}) = \frac{\hat{\beta}}{SE(\hat{\beta})}$$

حيث:

$$SE(\hat{\beta}) = \sqrt{Var(\hat{\beta})}$$

$$Var(\hat{\beta}) = \frac{\hat{\sigma}^2}{\sum x_i^2}$$

$$\hat{\sigma}^2 = \frac{\sum e_i^2}{DF}$$

$$X_i = X_i - \bar{X}$$

$$DF = N - K + 1$$

حيث أن:

$SE =$  الخطأ المعياري،  $Var =$  التباين،  $\hat{\sigma}^2 =$  التباين المقدر

لحد الخطأ،  $X =$  الوسط الحسابي ل X،  $DF =$  درجات الحرية،  $K =$  عدد المتغيرات المستقلة،  $K+1 =$  عدد معاملات الانحدار المقدرة.

5- إيجاد القيمة الجدولية لاختبار T، ويتم بالبحث في الجدول عند درجات حرية معينة  $(N-K+1)$  ومستوي معنوية معين.

6- مقارنة القيمة المحسوبة لاختبار T بالقيمة الجدولية:

- فإذا كانت قيمة T المحسوبة اكبر من قيمة T الجدولية، يتم قبول الفرض البديل القائل بان  $\hat{b} \neq 0$  ويقال في هذه الحالة أن

- فإذا كانت قيمة T المحسوبة أكبر من قيمة T الجدولية، يتم قبول الفرض البديل القائل بأن  $\hat{b} \neq 0$  ويقال في هذه الحالة أن  $\hat{b}$  معنوية إحصائياً، ويدل هذا علي وجود مشكلة عدم ثبات تباين حد الخطأ.

- أما إذا كانت قيمة T المحسوبة أقل من قيمة T الجدولية، يتم قبول فرض العدم القائل فإن  $\hat{b} = 0$ ، ويقال في هذه الحالة أن  $\hat{b}$  غير معنوية إحصائياً، ويدل علي وجود افتراض ثبات تباين حد الخطأ.

لاحظ انه إذا كان نموذج الانحدار يتضمن أكثر من متغير مستقل واحد، يتم تقدير  $R_s$  بين  $|e_i|$  وكل متغير مستقل علي حدة ثم القيام باختبار المعنوية الإحصائية بواسطة اختبار T للحكم علي وجود أو عدم وجود افتراض ثبات تباين حد الخطأ.

**معالجة عدم ثبات تباين حد الخطأ REMEDY OF HETEROSCEDASTICITY** من أبرز الطرق المستخدمة لتصحيح هذه المشكلة هي طريقة المربعات الصغرى العامة أو المرجحة (GLS) **GENERALIZED (OR) WEIGHTED LEAST SQUARES** وتقوم فكرة هذه الطريقة على إعطاء القيم ذات الانحراف الأقل عن خط الانحدار وزناً أكبر من القيم ذات الانحراف الأكبر في تقدير العلاقة محل الاعتبار ولذا فإن الوزن الذي تتخذه هو مقلوب الانحراف المعياري للبواقي  $E_t$  أي أن الوزن:

$$W_t = \frac{1}{\delta_t}$$

من الملاحظ أنه كلما قل تباين البواقي زاد الوزن  $W_t$  والعكس صحيح ومن ثم فإذا كان النموذج الأصلي هو:

$$\hat{Y}_t = \beta_1 + \beta_2 X_{2t} + \beta_3 X_{3t} + u_t$$

فإن النموذج المعدل الذي يتم تقديره لتلاشى مشكلة عدم ثبات التباين إن وجدت هو:

$$\frac{Y_t}{\delta_t} = \beta_1 \frac{1}{\delta_t} + \beta_2 \frac{X_{2t}}{\delta_t} + \beta_3 \frac{X_{3t}}{\delta_t} + \frac{u_t}{\delta_t}$$

وهي نفس الصيغة السابقة:

$$Y_t^* W_t = \beta_1 W_t + \beta_2 (W_t X_{2t}) + \beta_3 (W_t X_{3t}) + W_t u_t$$

$$Y_t^* = \beta_1 W_t + \beta_2 X_{2t}^* + \beta_3 X_{3t}^* + u_t^*$$

ثانياً: نتائج الدراسة التطبيقية:

أولاً: تقدير النموذج باستخدام العينة الصغيرة

تم الاعتماد على البيانات الميدانية التي تم تجميعها خلال الموسم الزراعي 2023/2022 من بعض مزارعي الترمس من مركز الحسينية بمحافظة الشرقية، حيث تم اختيار 25 مزارع بطريقة عشوائية حيث تم تجميع الإنتاج وعناصر الإنتاج (كمية التقاوي،

7- إيجاد القيمة الجدولية لإحصائية F عند درجات الحرية لكل من البسط والمقام، ومستوى معنوية معين.

8- مقارنة بين القيم المحسوبة لإحصائية F والقيمة الجدولية لها:

- فإذا كانت F المحسوبة أكبر من F الجدولية، نقبل الفرضية البديلة أي فرضية عدم ثبات تباين الأخطاء.

- أما إذا كانت F المحسوبة أقل من F الجدولية، يتم قبول فرضية العدم المتمثل في وجود افتراض ثبات تباين حد الخطأ.

لاحظ أن اختبار GOLDFELD-QUANDT لا يمكن تطبيقه إلا في حالة ما إذا كانت إحدى المتغيرات المستقلة هي المتسببة في وجود مشكلة عدم ثبات تباين حد الخطأ.

## 2- اختبار معامل ارتباط الرتب لـ SPEARMAN

يقيس معامل ارتباط الرتب لـ SPEARMAN درجة

الارتباط بين مجموعتين من الرتب، ويفرض أن:

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

يمكن بيان كيفية استخدام معامل ارتباط الرتب لـ

SPEARMAN في اكتشاف عدم ثبات تباين حد الخطأ من خلال الخطوات التالية:

1- تطبيق طريقة المربعات الصغرى العادية فينتج ما يلي:

$$\hat{Y}_i = \hat{\alpha} + \hat{\beta} X_i + e_i$$

2- الحصول علي البواقي أو القيم المقدرة لحد الخطأ ( $e_i$ ) من المعادلة التالية:

$$e_i = Y_i - \hat{Y}_i$$

3- إهمال إشارة  $e_i$  أي اخذ القيم المطلقة لـ  $|e_i|$ ، ثم ترتيب

كل من  $X_i$ ،  $|e_i|$  طبقاً لتزايد أو تناقص الرتب.

4- تقدير معامل ارتباط الرتب لـ SPEARMAN ( $R_s$ ) كما يلي:

$$r_s = 1 - \left[ \frac{6 \sum d_i^2}{N(N^2 - 1)} \right]$$

حيث أن:

$d_i$  = الفرق بين كل رتبتين متناظرتين ( $X_i - |e_i|$ ).  $N =$  عدد المشاهدات.

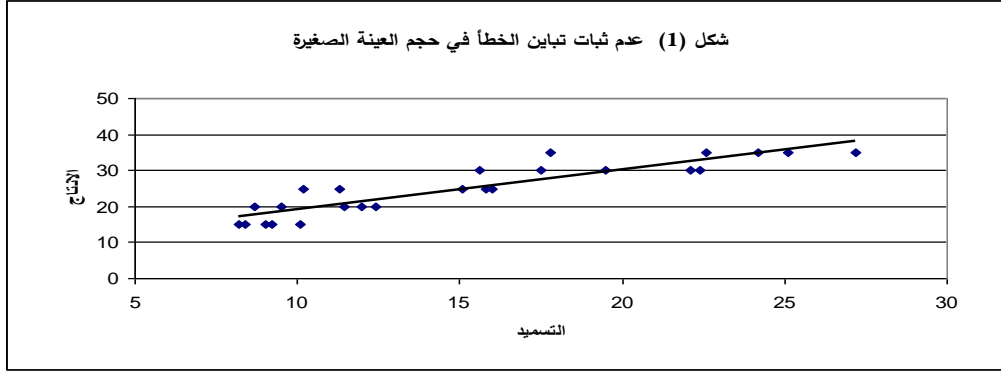
5- إيجاد القيمة المحسوبة لاختبار T كما يلي:

$$\hat{r} = \frac{r_s \sqrt{N-2}}{\sqrt{1-r_s^2}}$$

6- إيجاد القيمة الجدولية لاختبار T، ويتم ذلك بالبحث في جدول عند درجات حرية  $(N-K+1)$  ومستوى معنوية 5%.

7- مقارنة القيمة المحسوبة لاختبار T بالقيمة الجدولية:

كمية السماد الفوسفات، عدد العمال، كمية مياه الري) للمحصول وتبين من الشكل (1) حالة عدم ثبات التباين لحد الخطأ لبيانات وذلك للكشف عن وجود مشكلة عدم ثبات تباين حد الخطأ حيث العينة الصغيرة



المصدر: نتائج تحليل بيانات استمارة استبيان عينة البحث بمحافظة الشرقية موسم 2023/2022.

#### أولاً:- اختبار بارك PARK TEST

بتطبيق طريقة المربعات الصغرى العادية لبيانات الإنتاج لمحصول الترمس Y والتسميد X فينتج ما يلي:

$$\hat{Y}_i = -3.452 + 0.748 X_i$$

(-1.828)<sup>-</sup> (10.294)<sup>\*\*</sup>

ثم إيجاد القيم المقدرة لحد الخطأ ( $e_i$ )، وباستخدام طريقة المربعات الصغرى العادية في إجراء انحدار  $\ln X_i$  على  $\ln e_i^2$  فينتج ما يلي:

#### جدول (1) قيم $\ln X_i$ وقيم $\ln e_i^2$

$\ln e_i^2$	$\ln X_i$
-1.68	2.71
4.81	2.71
4.66	2.71
4.77	2.71
4.97	2.71
4.85	3.00
4.78	3.00
4.68	3.00
4.06	3.00
4.26	3.00
4.97	3.22
4.93	3.22
3.64	3.22
3.97	3.22
4.81	3.22
5.05	3.40
4.71	3.40
5.47	3.40
4.31	3.40
5.43	3.40
5.44	3.56
5.08	3.56
4.13	3.56
5.70	3.56
5.32	3.56

المصدر: جمعت وحسبت من نتائج تحليل بيانات استمارة استبيان عينة البحث بمحافظة الشرقية موسم 2023/2022.

$$\ln \hat{e}_i^2 = -1.232 + 1.811 \ln X_i$$

(-0.444)<sup>-</sup> (2.083)<sup>+</sup>

إيجاد القيمة المحسوبة لاختبار T بالنسبة ل  $\hat{\beta}$  والتي بلغت نحو 2.084، وبالكشف في جداول T تبين أن القيمة الجدولية بلغت نحو 2.069، يتم قبول الفرض البديل القائل بان  $\hat{\beta} \neq 0$  ويقال في هذه الحالة أن  $\hat{\beta}$  معنوية إحصائياً، ويدل هذا علي وجود مشكلة عدم ثبات تباين حد الخطأ.

#### ثانياً: اختبار GOLDFELD-QUANDT:

بافتراض النموذج التالي:

$$I=1,2,\dots,N, Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i$$

وللكشف عن مشكلة عدم ثبات تباين الخطأ باستخدام اختبار GOLDFELD-QUANDT لابد من إتباع الخطوات التالية:

- 1- ترتيب مشاهدات X ترتيباً تصاعدياً.
- 2- استبعاد المشاهدات الوسطى لكل من التسميد X والإنتاج Y ثم تكوين مجموعتين من المشاهدات بحيث يكون لكل مجموعة على حدا معادلة خاصة بها كما يلي :
- المجموعة الأولى: وتتمثل في المشاهدات الخاصة بكل من التسميد X و الإنتاج Y الواردة قبل المشاهدات التي تم استبعادها، والمعادلة الخاصة بهذه المجموعة هي :

$$Y_{1i} = 3.496 + 0.366 X_{1i}$$

(1.238)<sup>-</sup> (20289)<sup>+</sup>

- المجموعة الثانية: وتتمثل في المشاهدات الخاصة بكل من X التسميد والإنتاج Y الواردة بعد المشاهدات التي تم استبعادها، والمعادلة الخاصة بهذه المجموعة هي:

$$Y_{2i} = -4.340 + 0.792 X_{2i}$$

(-0.324)<sup>-</sup> (1.929)<sup>+</sup>

- 4- ثم الحصول على القيم المقدرة لحد الخطأ لإيجاد القيمة المحسوبة لإحصائية F لمقارنة بين القيم المحسوبة لإحصائية F والقيمة الجدولة لها :

جدول (4) رتبة كل من  $E$  ،  $X$ 

x	e	رتبة x	رتبة e
15	0.432	3	1
15	11.072	3	13.5
15	10.272	3	8
15	10.872	3	11
15	11.972	3	17.5
20	11.322	8	15
20	10.922	8	12
20	10.362	8	9
20	7.622	8	4
20	8.422	8	6
25	11.972	13	17.5
25	11.772	13	16
25	6.172	13	2
25	7.272	13	3
25	11.072	13	13.5
30	12.522	18	19
30	10.522	18	10
30	15.422	18	24
30	8.622	18	7
30	15.122	18	22
35	15.172	23	23
35	12.672	23	20
35	7.872	23	5
35	17.272	23	25
35	14.272	23	21

المصدر: جمعت وحسبت من نتائج تحليل بيانات استمارة استبيان عينة البحث بمحافظة الشرقية موسم 2023/2022.

جدول (5) رتبة كل من  $E$  ،  $X$  والفروق بين الرتب

$d^2$	d	رتبة x	رتبة e
4	2	3	1
110.25	-10.5	3	13.5
25	-5	3	8
64	-8	3	11
210.25	-14.5	3	17.5
49	-7	8	15
16	-4	8	12
1	-1	8	9
16	4	8	4
4	2	8	6
20.25	-4.5	13	17.5
9	-3	13	16
121	11	13	2
100	10	13	3
0.25	-0.5	13	13.5
1	-1	18	19
64	8	18	10
36	-6	18	24
121	11	18	7
16	-4	18	22
0	0	23	23
9	3	23	20
324	18	23	5
4	-2	23	25
4	2	23	21

المصدر: جمعت وحسبت من نتائج تحليل بيانات استمارة استبيان عينة البحث بمحافظة الشرقية موسم 2023/2022.

جدول (2) قيم  $e_{1i}^2$  وقيم  $e_{2i}^2$ 

$e_{1i}^2$	$e_{2i}^2$
0.618	0.006
0.046	3.686
0.343	8.880
0.000	14.592
1.241	7.182
2.509	2.958
1.402	0.608
0.389	31.136
4.477	14.592
1.732	0.672

المصدر: جمعت وحسبت من نتائج تحليل بيانات استمارة استبيان عينة البحث بمحافظة الشرقية موسم 2023/2022.

- فكانت  $F$  المحسوبة (6.61) أكبر من  $F$  (3.44) الجدولية،  
نقبل الفرضية البديلة أي فرضية عدم ثبات تباين الأخطاء  
(وجود مشكلة (HETEROSCEDASTICITY)).

ثالثاً: اختبار معامل ارتباط الرتب لـ SPEARMAN

تقدير معامل ارتباط الرتب لـ SPEARMAN(RS) كما يلي:

1- يتم ترتيب كل من  $E, X$  ترتيب تصاعدي

جدول (3) ترتيب قيم  $X$  و  $E$  ترتيب تصاعدي

مرتبة ترتيب تصاعدي $E$	مرتبة ترتيب تصاعدي $X$
0.432	15
6.172	15
7.272	15
7.622	15
7.872	15
8.422	20
8.622	20
10.272	20
10.362	20
10.522	20
10.872	25
10.922	25
11.072	25
11.072	25
11.322	25
11.772	30
11.972	30
11.972	30
12.522	30
12.672	30
14.272	35
15.122	35
15.172	35
15.422	35
17.272	35

المصدر: جمعت وحسبت من نتائج تحليل بيانات استمارة استبيان عينة البحث بمحافظة الشرقية موسم 2023/2022.

2- إيجاد رتبة كل من  $E$  ،  $X$  تبعا لتزايد أو تناقص الرتب

- تقدير معامل ارتباط الرتب ل Spearman ( $r_s$ ) كما يلي:

$$r_s = 1 - \left[ \frac{6 \times 1329}{25(624)} \right] = 0.5$$

- إيجاد القيمة المحسوبة لاختبار T كما يلي:

$$\hat{T} = \frac{0.5\sqrt{23}}{\sqrt{1 - (0.5)^2}} = 2.76$$

بمقارنة القيمة المحسوبة لاختبار T بالقيمة الجدولية والتي تقدر بنحو 2.069 تبين أن قيمة T المحسوبة أكبر من قيمة T الجدولية التي بلغت نحو 2.76، وبذلك يتم قبول الفرض البديل القائل بأن  $\hat{b} \neq 0$  ويقال في هذه الحالة أن  $\hat{b}$  معنوية إحصائياً، ويدل هذا على وجود مشكلة عدم ثبات تباين حد الخطأ.

معالجة مشكلة عدم ثبات تباين حد الخطأ

يتم معالجة مشكلة عدم ثبات تباين حد الخطأ من خلال إجراء تحويل للنموذج الأصلي بطريقة المربعات الصغرى المرجحة (Weighted Least Squares (WLS)، وتقوم هذه الفكرة على إعطاء القيم ذات الانحراف الأقل على خط الانحدار وزناً أكبر من القيم ذات الانحراف الأكبر كما هو موضح بالشكل (2)، وبحساب معادلة الانحدار بعد معالجة البيانات كما هو موضح بالجدول لبيانات الإنتاج لمحصول الترمس Y والتسميد X ما يلي:

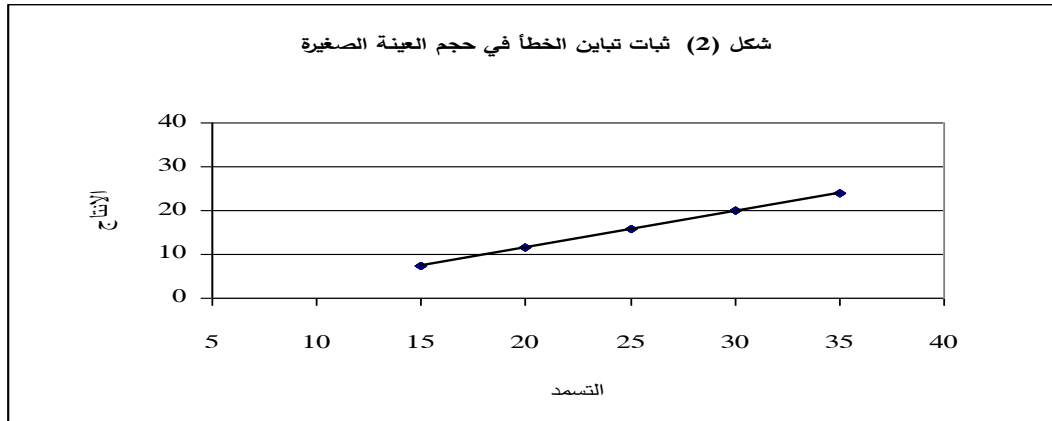
$$\hat{Y}_i = -5.060 + 0.832 X_i$$

(-1.944)<sup>-</sup> (9.643)<sup>\*\*</sup>

جدول (6) نتائج معالجة مشكلة عدم ثبات تباين حد الخطأ باستخدام برنامج EVIEWS

Dependent Variable: Y				
Method: Least Squares				
Included observations: 25				
Weighting series: Y				
Weight type: Inverse standard deviation (EViews default scaling)				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-5.06005	2.602311	-1.944446	0.0642
X	0.831773	0.086256	9.643078	0
Weighted Statistics				
R-squared	0.801705	Mean dependent var	17.48708	
Adjusted R-squared	0.793084	S.D. dependent var	13.20448	
S.E. of regression	2.788108	Akaike info criterion	4.965222	
Sum squared resid	178.7915	Schwarz criterion	5.062732	
Log likelihood	-60.0653	Hannan-Quinn criter.	4.992267	
F-statistic	92.98894			
Prob(F-statistic)	0			

المصدر: برنامج EVIEWS.



المصدر: نتائج تحليل بيانات استمارة استبيان عينة البحث بمحافظة الشرقية موسم 2023/2022.

للتأكد من معالجة مشكلة عدم ثبات تباين حد الخطأ يتم الكشف 1- اختبار بارك Parak test من خلال القيم الجديدة لحد الخطأ باستخدام:  $(e_i)$  التي تم الحصول عليها من المعالجة، وباستخدام طريقة

هذه الحالة أن  $\hat{b}$  غير معنوية إحصائياً، وبدل علي وجود افتراض ثبات تباين حد الخطأ.

تم إجراء اختبارات الكشف علي وجود مشكلة عدم ثبات تباين حد الخطأ ما بين الإنتاج وعناصر الإنتاج (كمية التقاوي، كمية السماد الفوسفات، عدد العمال، كمية مياه الري) الخاصة بمحصول الترمس كلا علي حدا فتبين وجود مشكلة عدم ثبات تباين حد الخطأ في كمية السماد و كمية التقاوي وتم الاكتفاء بشرح علي عنصر واحد فقط كمثال لتوضيح كيفية الكشف عن المشكلة بالبيانات.

#### ثانياً: تقدير النموذج باستخدام العينة الكبيرة

تم الاعتماد على البيانات الميدانية التي تم تجميعها خلال الموسم الزراعي 2023/2022 من بعض مزارعي الترمس من مركز الحسينية بمحافظة الشرقية، حيث تم اختيار 102 مزارع بطريقة عشوائية حيث تم تجميع بيانات الإنتاج والتسميد للمحصول وذلك للكشف عن وجود مشكلة عدم ثبات تباين حد الخطأ حيث تبين من الشكل (3) حالة عدم ثبات التباين لحد الخطأ لبيانات العينة الكبيرة

المربعات الصغرى العادية تم إجراء انحدار  $\ln e_i^2$  علي  $\ln X_i$  كانت كالتالي:

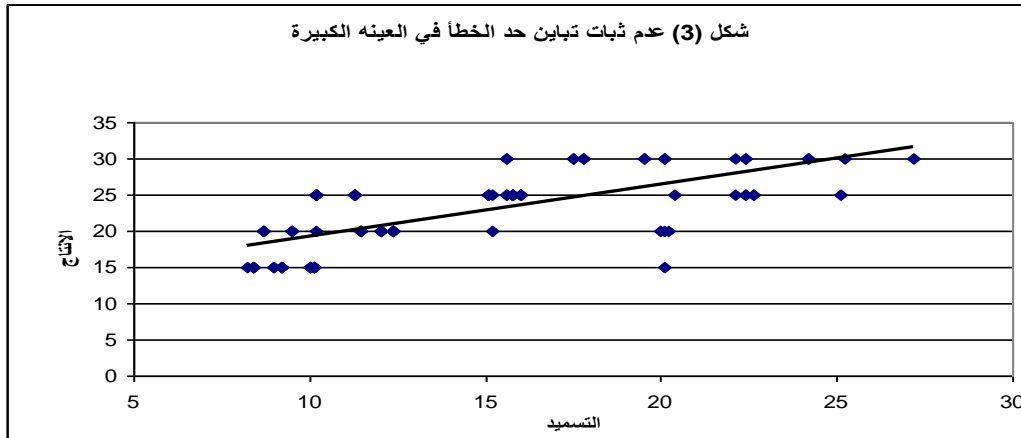
$$\ln \hat{e}_i^2 = -1.429 + 0.546 \ln X_i$$

(-0.270)<sup>-</sup> (0.329)<sup>-</sup>

وإيجاد القيمة المحسوبة لاختبار T بالنسبة ل  $\hat{\beta}$  والتي بلغت نحو 0.329 ، وبالكشف في جداول T تبين أن القيمة الجدولية بلغت نحو 2.069، يتم قبول فرض عدم القائل فإن  $\hat{b} = 0$  ويقال في هذه الحالة أن  $\hat{b}$  غير معنوية إحصائياً، وبدل علي وجود افتراض ثبات تباين حد الخطأ.

2- اختبار Goldfeld-Quandt من خلال القيم الجديدة لحد الخطأ ( $e_i$ ) التي تم الحصول عليها من المعالجة فكانت F المحسوبة (3.12) أقل من F الجدولية (3.44)، يتم قبول فرضية عدم المتمثل في وجود افتراض ثبات تباين حد الخطأ.

1- اختبار معامل ارتباط الرتب ل Spearman ( $r_s$ ) تبين بمقارنة القيمة المحسوبة لاختبار T بالقيمة الجدولية والتي تقدر بنحو 2,069 تبين أن قيمة T المحسوبة أقل من قيمة T الجدولية، يتم قبول فرض عدم القائل فإن  $\hat{b} = 0$  ويقال في



المصدر: نتائج تحليل بيانات استمارة استبيان عينة البحث بمحافظة الشرقية موسم 2023/2022.



$$\hat{Y}_i = -0.861 + 0.675 X_i$$

(-0.537)<sup>\*</sup>      (9.673)<sup>\*\*</sup>

أولاً: -الكشف باختبار بارك Park Test

بتطبيق طريقة المربعات الصغرى العادية لبيانات الإنتاج  
لمحصول الترمس Y والتسميد X فينتج ما يلي:

جدول (7) قيم  $\ln X_i$  وقيم  $\ln e_i^2$

العينة	lne2	lnx	العينة	lne2	lnx	العينة	lne2	lnx
1	-0.61305	2.70805	35	-8.5374	3.218876	69	-0.8957	2.995732
2	-5.49774	2.70805	36	-3.08356	3.218876	70	0.362976	2.995732
3	-0.29237	2.70805	37	3.520538	3.218876	71	2.741854	2.995732
4	-0.61305	2.70805	38	3.101074	3.218876	72	2.287809	2.995732
5	-0.35825	2.70805	39	3.708216	3.218876	73	-8.5374	3.218876
6	-2.86258	2.995732	40	-1.76378	3.218876	74	-3.08356	3.218876
7	-0.8957	2.995732	41	3.611982	3.218876	75	3.520538	3.218876
8	0.362976	2.995732	42	4.41347	3.218876	76	3.101074	3.218876
9	2.741854	2.995732	43	3.769892	3.218876	77	-0.17985	3.218876
10	2.287809	2.995732	44	0.92621	3.401197	78	-4.39645	3.401197
11	-8.5374	3.218876	45	3.519505	3.401197	79	1.272095	3.401197
12	-3.08356	3.218876	46	3.14181	3.401197	80	2.204544	3.401197
13	3.520538	3.218876	47	4.765748	2.70805	81	2.664204	3.401197
14	-0.41159	3.218876	48	-0.35825	2.70805	82	-0.68217	3.401197
15	-0.17985	3.218876	49	-2.86258	2.995732	83	-5.49774	2.70805
16	-4.39645	3.401197	50	-0.8957	2.995732	84	-0.29237	2.70805
17	1.272095	3.401197	51	0.362976	2.995732	85	-2.66361	2.70805
18	2.956835	3.218876	52	4.019379	2.995732	86	-0.35825	2.70805
19	2.664204	3.401197	53	3.992392	2.995732	87	-2.86258	2.995732
20	1.994635	3.401197	54	-8.5374	3.218876	88	-3.08356	3.218876
21	-0.68217	3.401197	55	-3.08356	3.218876	89	3.520538	3.218876
22	3.769892	3.218876	56	3.520538	3.218876	90	3.101074	3.218876
23	0.92621	3.401197	57	3.101074	3.218876	91	3.708216	3.218876
24	4.111066	3.401197	58	-0.17985	3.218876	92	-1.76378	3.218876
25	3.14181	3.401197	59	-3.08356	3.218876	93	-0.61305	2.70805
26	-5.49774	2.70805	60	3.520538	3.218876	94	-0.35825	2.70805
27	-0.29237	2.70805	61	3.101074	3.218876	95	-2.86258	2.995732
28	-0.35825	2.70805	62	2.204544	3.401197	96	-0.8957	2.995732
29	-0.35825	2.70805	63	0.124071	2.70805	97	0.362976	2.995732
30	-2.86258	2.995732	64	-5.49774	2.70805	98	2.741854	2.995732
31	-0.8957	2.995732	65	-0.29237	2.70805	99	2.287809	2.995732
32	4.046007	2.995732	66	-2.66361	2.70805	100	-8.5374	3.218876
33	1.880796	2.995732	67	-0.35825	2.70805	101	-3.08356	3.218876
34	1.783176	2.995732	68	-2.86258	2.995732	102	3.520538	3.218876

المصدر: جمعت وحسبت من نتائج تحليل بيانات استمارة استبيان عينة البحث بمحافظة الشرقية موسم 2022/2023.

ثم إيجاد القيم المقدرة لحد الخطأ ( $e_i$ )، وباستخدام طريقة المربعات الصغرى العادية في إجراء انحدار  $\ln e_i^2$  علي  $\ln X_i$  فينتج ما يلي:

$$\ln \hat{e}_i^2 = -8.459 + 2.743 \ln X_i$$

(-1.999)<sup>\*</sup>      (2.005)<sup>\*</sup>

ويقال في هذه الحالة أن  $\hat{b}$  معنوية إحصائياً، ويدل هذا علي وجود مشكلة عدم ثبات تباين حد الخطأ.

هو موضح بالشكل (4)، وبحساب معادلة الانحدار بعد معالجة البيانات كما هو موضح بالجدول لبيانات الإنتاج لمحصول الترمس Y والتسميد X ما يلي:

$$\hat{Y}_i = -0.861 + 0.675 X_i$$

(-0.537)<sup>-</sup> (9.673)<sup>\*\*</sup>

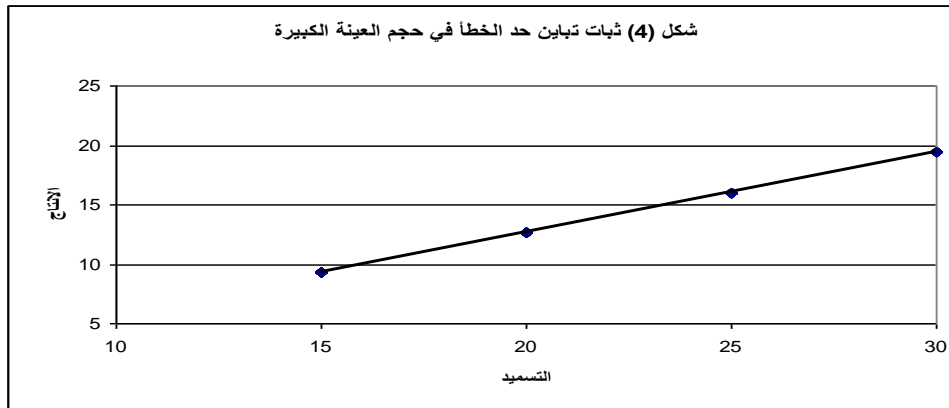
معالجة مشكلة عدم ثبات تباين حد الخطأ

يتم معالجة مشكلة عدم ثبات تباين حد الخطأ من خلال إجراء تحويل للنموذج الاصلي بطريقة المربعات الصغرى المرجحة، وتقوم هذه الفكرة على إعطاء القيم ذات الانحراف الأقل على خط الانحدار وزناً أكبر من القيم ذات الانحراف الأكبر كما

جدول (8) نتائج معالجة مشكلة عدم ثبات تباين حد الخطأ باستخدام برنامج Eviews

Dependent Variable: Y				
Method: Least Squares				
Sample: 1 102				
Included observations: 102				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.86102	1.602045	-0.53745	0.5922
X	0.674916	0.069773	9.673094	0
<b>R-squared</b>	0.483388	<b>Mean dependent var</b>		14.25843
<b>Adjusted R-squared</b>	0.478222	<b>S.D. dependent var</b>		4.912464
<b>S.E. of regression</b>	3.54848	<b>Akaike info criterion</b>		5.390329
<b>Sum squared resid</b>	1259.171	<b>Schwarz criterion</b>		5.441799
<b>Log likelihood</b>	-272.907	<b>Hannan-Quinn criter.</b>		5.411171
<b>F-statistic</b>	93.56875			
<b>Prob(F-statistic)</b>	0			

المصدر: برنامج Eviews.



المصدر: نتائج تحليل بيانات استمارة استبيان عينة البحث بمحافظة الشرقية موسم 2023/2022.

ويقال في هذه الحالة أن  $\hat{b}$  غير معنوية إحصائياً، ويدل علي وجود افتراض ثبات تباين حد الخطأ.

تم إجراء اختبارات الكشف علي وجود مشكلة عدم ثبات تباين حد الخطأ الإنتاج وعناصر الإنتاج (كمية التقاوي، كمية السماد الفوسفات، عدد العمال، كمية مياه الري) الخاصة بمحصول الترمس فتبين وجود مشكلة عدم ثبات تباين حد الخطأ في كمية السماد الفوسفاتي وكمية مياه الري وتم الاكتفاء بشرح علي عنصر واحد فقط كمثال لتوضيح كيفية الكشف عن المشكلة بالبيانات.

#### الملخص:

يهدف البحث إلي دراسة طبيعة مشكلة عدم ثبات تباين حد الخطأ، أسبابه وآثاره وتطبيق أهم طرق الكشف عن عدم الثبات في البيانات المقطعية، وتم الاعتماد على البيانات الميدانية التي تم تجميعها خلال الموسم الزراعي 2023/2022 من مزارعي

للتأكد من معالجة مشكلة عدم ثبات تباين حد الخطأ يتم الكشف باستخدام:

1- اختبار بارك من خلال القيم الجديدة لحد الخطأ ( $e_i$ ) التي تم الحصول عليها من المعالجة، وباستخدام طريقة المربعات الصغرى العادية تم إجراء انحدار  $\ln e_i^2$  علي  $\ln X_i$  كانت كالتالي:

$$\ln \hat{e}_i^2 = -8.417 + 2.722 \ln X_i$$

(-1.961)<sup>-</sup> (1.961)<sup>-</sup>

وإيجاد القيمة المحسوبة لاختبار T بالنسبة ل  $\hat{\beta}$  والتي بلغت نحو 1.961، وبالكشف في جداول T تبين أن القيمة الجدولية بلغت نحو 1.984، يتم قبول فرض العدم القائل فإن  $\hat{b} = 0$

2013/24053 التقييم الدولي: 2-082-735-977-978)، الإسكندرية.

3- عبد القادر محمد عبد القادر عطية، الحديث في الاقتصاد القياسي بين النظرية والتطبيق، (2005)، الدار الجامعية، الإسكندرية.

4- وليد اسماعيل السينو، احمد محمد مشعل، (2003)، الاقتصاد القياسي التحليلي، دار مجدلاوي للنشر والتوزيع.

5- وليد إسماعيل السيفو، (2006)، أساسيات الاقتصاد القياسي التحليلي، الأهلية للنشر والتوزيع والطباعة- لبنان.

6- Andreas G. Klein , Carla Gerhard , Rebecca D. Büchner , Stefan Diestel & Karin Schermelleh-Engel.(2015). The Detection of Heteroscedasticity in Regression Models for Psychological Data, Psychological Test and Assessment Modeling, Volume 58, (4): 542-568.

7-Gujarati,D.N.(2004) Basic Econometrics, 4<sup>th</sup> Edition, the mc Grow-Hill Companies,new York, 387-426.

8- Kiefer, N. and Vogelsang, T. (2002a). Heteroskedasticity-autocorrelation robust testing using bandwidth equal to sample size. Econometric Theory, 18:1350-1366.

9- Kiefer, N. and Vogelsang, T. (2002b). heteroskedasticity-autocorrelation robust standard errors using the bartlett kernel without truncation. Econometrica, 70:2093-2095.

10- Kim, M. and Sun, Y. (2013). Heteroskedasticity and spatiotemporal dependence robust inference for linear panel models with fixed effects. Journal of Econometrics, 177:85- 108.

11- Rosopa, P. J., Schaffer, M. M., & Schroeder, A. M. (2013). Managing heteroscedasticity in general linear models. Psychological Methods, 18(3): 335-351.

12- Stanislaus S. Uyanto,(2022) Monte Carlo power comparison of seven most commonly used heteroscedasticity tests, Communications in Statistics - Simulation and Computation, 2065-2082

13- Sun, Y. (2013). A heteroskedasticity and autocorrelation robust f test using orthonormal series variance estimator. Econometrics Journal, 16:1-26.

الترمس بمحافظة الشرقية عن طريق استمارة استبيان أعدت لهذا الغرض، حيث تم اختيار عينتين الأولى 25 مزارع والثانية 102 مزارع بطريقة عشوائية في مركز الحسينية بمحافظة الشرقية، وقد تم التحليل باستخدام برنامج Eviews.

تم إجراء اختبارات الكشف علي وجود مشكلة عدم ثبات تباين حد الخطأ الإنتاج وعناصر الإنتاج (كمية التقاوي، كمية السماد الفوسفات، عدد العمال، كمية مياه الري) لمحصول الترمس فتبين وجود مشكلة عدم ثبات تباين حد الخطأ في كمية السماد وكمية التقاوي بالعينة 25 مزارع، بينما في حالة العينة 102 مزارع تبين وجود مشكلة عدم الثبات في كمية السماد الفوسفاتي وكمية مياه الري وتم الاكتفاء بشرح علي عنصر واحد (كمية السماد) فقط كمثال لتوضيح كيفية الكشف عن المشكلة بالبيانات.

فتبين في حالة العينة (25 مزارع) باستخدام اختبار بارك أن قيمة T المحسوبة بلغت نحو 2.084 وهي اكبر من القيمة الجدولية التي بلغت نحو 2.069، فيتم قبول الفرض البديل بوجود مشكلة عدم الثبات، وتم معالجة البيانات من خلال إجراء تحويل للنموذج الأصلي بطريقة المربعات الصغرى المرجحة فبلغت قيمة T المحسوبة نحو 0.329 وهي اقل من القيمة الجدولية فيتم قبول فرض عدم وجود ثبات تباين حد الخطأ.

وفي حالة العينة (102 مزارع) باستخدام اختبار بارك ان قيمة T المحسوبة بلغت نحو 2.005، وهي اكبر من قيمة T الجدولية التي بلغت نحو 1.984 فيتم قبول الفرض البديل بوجود مشكلة عدم الثبات، وبعد المعالجة بلغت قيمة T المحسوبة نحو 1.961 وهي اقل من قيمة T الجدولية فيتم قبول فرض عدم الثبات بوجود ثبات تباين حد الخطأ، ويتضح من ذلك أن حجم العينة ليس له أي علاقة بوجود أو عدم وجود مشكلة عدم ثبات تباين حد الخطأ بالبيانات المقطعية.

#### التوصيات:

توصي الدراسة بضرورة اهتمام الباحثين بإجراء الاختبارات اللازمة للكشف عن مشكلة عدم ثبات تباين حد الخطأ وضرورة معالجتها قبل البدء في التحليل الإحصائي للحصول علي نتائج دقيقة ثم استخدام البيانات بالصورة المناسبة لطبيعة البيانات المقطعية.

#### المراجع:

- 1- جابر أحمد بسيوني & سالم الحاسية(2007)، الإحصاء التطبيقي، منشورات جامعة عمر المختار. البيضاء.
- 2- جابر أحمد بسيوني (2014)، الإحصاء العام، دار الوفاء لدنيا الطباعة والنشر، الطبعة الأولى (رقم الإيداع: