

## تعديل اختبار كولد فيلد- كوانط المعدل لتشخيص عدم التجانس بوجود القيم الشاذة

حسن سامي عرببي الشمري

قسم الاحصاء

كلية الادارة و الاقتصاد - جامعة القادسية

الديوانية - العراق

Email:hssn.sami1@gmail.com

### **Abstract:**

**المستخلص:**

Rana et. al. in 2008 proposed modification Goldfield-Quant test for detection of heteroscedasticity in the presence of outliers by using Least Trimmed Squares (LTS). The proposed method assumed trimming ten percent from dataset. In this situation may take place losing in clean data where the real percentage of outliers is less than (10%) or the estimator may be misleading where the outliers percentage is more than (10%). The researcher proposed another modification of Goldfield-Quant test to be robust. The results show our proposed method is better than the last modification of Goldfield-Quant test and the classical one.

اقترح رانا واخرون عام ٢٠٠٨ تعديل اختبار كولد فيلد-كوانط لتشخيص مشكلة عدم التجانس بوجود قيم شاذة، ذلك بالاستفادة من طريقة المربعات الصغرى المشدبة بتثبيت نسبة التشديب عند (١٠%). هذه النسبة قد تسبب خسارة في البيانات اذا كانت نسبة الشوادع اقل من (١٠%) و فشل في حسانة التقدير اذا كانت النسبة اكبر من ذلك. تهدف هذه الدراسة الى حل هذه المشكلة من خلال مقترن تعديل اخر على اختبار كولد فيلد-كوانط المعدل. لقد اظهرت النتائج كفاءة عالية للتعديل المقترن في تشخيص المشكلة بوجود نسب مختلفة من القيم الشاذة.  
الكلمات المفتاحية: كولد فيلد- كوانط ، LTS، القيم الشاذة، عدم التجانس، اختبار F.

## ١. المقدمة

تحقق الفرضية البديلة بظهور مشكلة عدم التجانس التي يمكن صياغتها إلى نوعين من عدم التجانس الا وهما المتضاعف حسب نموذج هارفي (١٩٧٦) والذي صيغته العامة

$$\sigma_t^2 = \sigma^2 \exp(z/\alpha)$$

حيث  $z$  متغير التوضيحي الذي ربما يكون او لا يكون متزامن مع متغير  $X$  ، اما  $\alpha$  هو متغير المعامل. ويمكن اعادة الصيغة اعلاه بشكل مبسط كالتالي :<sup>(١٨)</sup> (Rana and et.al, 2008)

$$\sigma_t^2 = \sigma^2 c_t$$

حيث  $c_t$  كمية غير معلومة.

اما النوع الثاني فيدعى (Groupwise) الذي تتوزع فيه المشاهدات على عدد من المجموعات لكل مجموعة تباين حد العشوائي ثابت ولكن مشكلة عدم التجانس تظهر مابين المجموعات ، لذلك متوسط حد الخطأ العشوائي في هذه الحالة يكون صفرًا الا ان التباين لا يكون ثابتا.

واحدة من فرضيات نموذج الانحدار الخطى ان حد الخطأ العشوائى له تباين ثابت عند كل مشاهدة و التي يطلق عليها ايضا فرضية تجانس تباين الخطأ العشوائى. قد تتحقق هذه الفرضية ارض الواقع او قد لا تتحقق، من حيث ان تباين حد الخطأ العشوائى في الكثير من الاحيان لا يكون ثابتا<sup>(١)</sup>.

انتهاء فرضية تجانس تباين الخطأ العشوائى سيؤدي الى ظهور مشكلة عدم ثبات تجانس التباين أي أن

$$Var(u_i) \neq \sigma_u^2$$

عندننن ترفض الفرضية الصفرية التي تفترض تجانس التباين كالتالي:

$$\sigma_1^2 = \sigma_2^2 = \sigma_3^2 = \sigma_4^2 \dots \dots \sigma_u^2$$

و تتحقق الفرض البديل للفرضية الصفرية كالتالي:

$$\sigma_1^2 \neq \sigma_2^2 \neq \sigma_3^2 \neq \sigma_4^2 \dots \dots \sigma_u^2$$

الغني ، (١٩٨٢) ، Cook and (Weisberg,1983) .<sup>(٥)</sup>

و لغرض معالجة مشكلة عدم تجانس التباين اقترح ترجيح طريقة المربعات الصغرى باوزان محددة و بما يسمى طريقة المربعات الصغرى الموزونة.

خرق اخر ممكن ان يعطى العمل بطريقة المربعات الصغرى هو ظهور القيم الشاذة في بيانات الانحدار والتي تعرف بانها تلك القيم التي تبعد عن خط الانحدار و لا تتبع نفس التوزيع الذي تتبعه معظم البيانات ، Imon (2002,2005,2008,2014). و تصنف الشواد حسب موقع ظهورها في جانب المتغير المعتمد او المستقل او كلاهما معاً. و اسباب ظهورها متعددة و مختلفة و لكن الحقيقة الثابتة حسب رأي معظم الخبراء في مجال الاحصاء انه لا توجد بيانات حقيقة تخلو من القيم الشاذة لكن تأثيرها مختلف حسب نسبة ظهورها قياسا باحجام العينات.

بناء على ما تقدم سعى الباحثون في مجال العلوم الاحصائية و الاقتصاد القياسي الى ايجاد طرق لتشخيص مشكلة عدم التجانس اولا ومن ثم اقتراح طرق من شأنها الوصول على تقديرات غير متحيزة للمعلمات في ظل وجود مشكلة عدم التجانس.

طرق التشخيص المقترحة صفت الى نوعين هما: طرق تخطيطية تعتمد على رسم العلاقة في تحديد المشكلة مثل Chatterjee (2008)<sup>(٦)</sup> و (and Hadi, 2006)<sup>(٧)</sup> وغيرهم.

اما الطرق الاخرى هي طرق تحليلية مثل كولدفيلد-كواント ( 1965 )<sup>(٤)</sup> ، جليسير ( 1969 )<sup>(٤)</sup> ، بارك ( 1966 )<sup>(٤)</sup> ، اختبار سبيرمان لارتباط White ( 1953 )<sup>(٤)</sup> ، ( 1980 )<sup>(٤)</sup> ، ( 2011 )<sup>(٣)</sup> ، ( الزبيدي، حسين ١٩٧٨ )<sup>(٢)</sup> ، ( العيسوي ، ابراهيم ، طيف ، ١٩٩٥ )<sup>(٤)</sup> ، ( محبوب ، عادل عبد هادي ، ١٩٩٥ )<sup>(٤)</sup> ، ( كاظم ، اموری

و جدا ان النقطة (and Leroy, 1987) التي تنهار فيها مقدرات هذه الطريقة هي:

$$BP = ((n - p)/2 + 1)/n$$

حيث ان  $P$  هي عدد المعلمات للنموذج و  $n$  تمثل حجم العينة.

و لقد اثبتت الدراسات التي بنت هذه الطريقة ان نسبة قدرتها على مقاومة القيم الشاذة تصل الى اقل من ٥٥٪، بقليل تبعا الى نسبة القيم الشاذة في البيانات الاصلية.<sup>(١٨)</sup>

اختبار كولدفيلد-كونانت و احد من اهم الاختبارات المستخدمة في تشخيص مشكلة عدم التجانس الا ان هذا الاختبار يحتاج الى احتساب متوسط مربعات بوافي الانحدار لعينتين جزئيتين لقياس قوة عدم التجانس و خوارزمية هذا الاختبار تعتمد على طريقة المربعات الصغرى للوصول الى متوسط مربعات البوافي، وبما ان هذه الطريقة تتحسن القيم الشاذة و تعد طريقة غير حصينة، لذلك ستعطي مقدرات مطللة و بالتالي

تختلف مقاومة طرق التقدير للقيم الشاذة من طريقة الى اخرى الا ان طريقة المربعات الصغرى لا تمتلك اية قدرة مقاومة قيمة شاذة واحدة على الاقل (Hampel, 1971)، لذلك لابد من استخدام طرق التقدير الحصينة بدلا من طريقة المربعات الصغرى عند ظهور القيم الشاذة. (Rousseeuw and Leroy, 1987).

المربعات الصغرى المشذبة (LTS) التي يمكن ايجاد مقدراتها من خلال تقليل دالة الهدف ادناء الى اقل ما يمكن

$$\min \sum_{i=1}^h r_{(i)}^2,$$

عندما  $r_{(i)}^2$  تمثل مربعات الباقي المرتبة  $i$  يمكن تحديدها كالتالي:

$$h = [(1 - \alpha) \times n] + [\alpha(k + 1)]$$

حيث  $\alpha$  هي النسبة المؤدية للقيم الشاذة او هي نسبة تشذيب المشاهدات. اما  $n$  هي حجم العينة و  $k$  هي عدد معاملات نموذج الانحدار. (Rousseeuw, 1987)

التجانس في بيانات انحدار خطى ملوثة بقيم شاذة. اذ ان الدراسة السابقة افترضت نسبة التشذيب هي (١٠٪) في أي حال من الاحوال. هذه النسبة كانت مثار لعدد من التساؤلات، مثلاً لو كانت نسبة الشواد تقل عن (١٠٪)، الا تتسبب النسبة الافتراضية للتشذيب بخسارة في بعض البيانات النظيفة، من جانب اخر اذا بلغت نسبة الشواد (١٥٪) او اكثراً، الا تتسبب نسبة التشذيب الافتراضية ببقاء (٥٪) او اكثراً من الشواد في البيانات، مما تكون سبباً كافياً الى فشل حصانة طريقة التقدير والتي بالضرورة ستؤدي الى فشل في تشخيص مشكلة عدم التجانس.

هذه التساؤلات شكلت محور المشكلة في هذه الدراسة، و لانها تحتاج الى اجابة من خلال التحقق من المشكلة و معالجة في حالة اثبات وجود المشكلة من خلال اقتراح تعديل طريقة كولد فيلد - كوانط المعدل ليكون اكثراً حصانة. بما ان هناك امكانية لاستخدام الطرق الحصينة استخدمنا طريقة المربعات الصغرى

ربما ستؤدي الى فشل الاختبار في تشخيص مشكلة عدم التجانس.

من اهم الدراسات السابقة التي تطرقت الى موضوع تحصين اختبار كولدفيلد- كوانط هي الدراسة التي تقدم بها (Rana and et.al, 2008)<sup>(18)</sup> و التي اختارت توظيف طريقة المربعات الصغرى المشذبة لتحسين وتعديل اختبار كولدفيلد-كوانط الا ان هذه الطريقة اعتمدت على احتساب قيمة الوسيط لمربعات بوافي الانحدار المحذوفة لكل مجموعة من البيانات المجزأة ومن ثم ايجاد النسبة بينهما من خلال جعل القيمة العظمى في البسط والصغرى في المقام و هذه النسبة تمثل اختبار فيشر لتجانس التباين، التي ممكن مقارنتها بقيمها الحرجة عند درجات حرياتها و مستوى معنوية محدد لرفض او قبول الفرضية القائلة بتجانس التباين.

طرح هذه الدراسة تساؤلاً عن قدرة اختبار كولد فيلد- كوانط التقليدي والمعدل في تشخيص مشكلة عدم

العلاقة الخطية للانحدار قبل الشروع بعملية التقدير ، لكن الاعتماد على عملية تشخيص التجانس من عدمه غير كافية لأن وجود القيم الشاذة ربما يفقد كفاءة طرق التشخيص مثل طريقة كولدفيلد-كانت التقليدية مما يجعلها لا تتحسن عدم التجانس الحاصل في البيانات. من جانب اخر ابتعاد القيمة الشاذة عن معظم البيانات و اتباعها لتوزيع اخر ، يدفع بقية البيانات التي تتبع التوزيع الطبيعي الى التمركز في زاوية ما ، الامر الذي لايمكن ملاحظته من خلال الرسم التخطيطي لعلاقة الاخطاء العشوائية بالقيم المقدرة باستخدام المربعات الصغرى. ومن هنا تبرز اهمية مقترح هذا البحث في تعديل اختبار كولد فيلد-كانت لتشخيص مشكلة عدم التجانس بوجود القيم الشاذة.

### ٣. خوارزمية تعديل اختبار كولد فيلد – كانت المعدل ( RMGQ )

هذه الخوارزمية هي تعديل لخوارزمية MGQ المقترحة من قبل Rana (18) and et.al, 2008 لغرض تشخيص

المشتبهة وفقاً للصيغة الثانية المتوفرة في برنامج S-PLUS2000 كتعديل على هذا الاختبار. لذلك تتمحور اهداف هذه الدراسة حول النقاط الآتية:

١. تحصين اختبار كولد فيلد – كانت باستخدام طريقة المربعات الصغرى المشتبهة ليكون اكثر حسنة في تشخيص مشكلة عدم التجانس في ظل وجود نسب مختلفة لقيم الشاذة في بيانات المتغير المعتمد تحديداً.
٢. الاستفادة من وسيط مربعات بوافي الانحدار بدلاً من متوسط مربعات البوافي او وسيط مربعات الانحدار المحنوفة في ايجاد قيمة الاختبار .F

## ٤. أهمية الدراسة

أغلب الاحيان تظهر مشكلة عدم التجانس في بيانات الاقتصاد القياسي مثل العلاقة بين الدخل و الانفاق ، ونظراً الى ان ظهور هذه المشكلة يجعل من تقديرات المربعات الصغرى تقديرات متحيزه لايمكن الاعتماد عليها، لذلك يلجأ العديد من الباحثين في مجال الاقتصاد والاحصاء الى اختبار بيانات

لعمل طريقة المربعات الصغرى المشتبهة.

٦. تقدير معلمات كل مجموعة باستخدام طريقة المربعات الصغرى المشتبهة (LTS) بالاعتماد على الحجم الجديد لكل عينة جزئية.

٧. لكلا العينتين الجزئيتين يتم حساب وسيط مربعات البوافي  $MSR$  ، وتحديد أي من القيمتين اعظم  $Max$  و ايهما اقل  $Min$

٨. ايجاد قيمة اختبار  $F$  المحسوبة وفقا للصيغة الآتية:

$$F = \frac{Max}{Min}$$

و بما ان توزيع معظم البيانات هو الطبيعي لذلك بالتأكيد سيكون وسيط مربعات البوافي لكلا المجموعتين سيبتعد توزيع مربع كاي ،

مشكلة عدم التجانس في بيانات الانحدار

لعينات كبيرة الحجم الى حد ما في ظل وجود بعض الشواذ في متغير الاستجابة و كالتالي:

١. رتب مشاهدات المتغير التوضيحي تصاعديا.

٢. حذف (C) من المشاهدات الوسطية من بيانات العينة ويفضل حذف مابنته  $1/5$  من المشاهدات.

٣. تقسم المشاهدات الباقيه  $(n - c)$  الى  $(n - c)/2$  عينتين جزئيتين متساويتين تتطوي الأولى على قيم  $X_i$  الصغيرة ، والثانية على قيم  $X_i$  الكبيرة .

٤. تحديد نسبة القيم الشاذة في كل عينة جزئية بعد تشخيصها باحدى طرق التشخيص الحصينة.

٥. ضرب نسبة القيم الشاذة في كل عينة جزئية بحجم تلك العينة لتحديد حجم جديد

النسبة المفترضة في الدراسة السابقة و التي كانت (١٠٪) من اجمالي حجم العينة.اما المفضلة بين الطرق الثلاث سيتعدد على ضوء قدرة الطريقة في رفض الفرضية الصفرية التي تشير الى تجانس التباين في حالة وجود او عدم وجود القيم الشاذة و للتأكد من صحة ودقة وقوه الاختبار المستخدم سيعتمد احتساب قيمة  $p$  لكل طريقة لبيان أي من الطرق اكثر قوته ودقة في تشخيص مشكلة عدم التجانس للاخطاء العشوائية.

٣. تحليل البيانات :

لقد تم دراسة ثلاثة مجاميع من البيانات الحقيقة المختبرة من قبل عدد من الباحثين في مجال الاحصاء و الاقتصاد القياسي تتعلق في دراسة العلاقة الخطية بين الانفاق و الدخل وهي نفس البيانات التي تم استخدامها في الدراسة السابقة ، و لجعل

اذا يمكن ان نستخدم اختبار  $F$  تحت مستوى معنوية محدد و بدرجة حرية  $(n - c - 2k)/2$  فرضية عدم بوجود مشكلة عدم التجانس اذا كانت قيمة  $F$  المحتسبة اقل من قيمة  $F$  الحرجة في الجداول الاحصائية. هذا الاجراء في احتساب قيمة  $F$  سيعتمد ايضا باستخدام الطريقة التقليدية للاختبار و باستخدام طريقة MGQ المقترنة من قبل (Rana and et.al, 2008)<sup>(18)</sup>. و لغرض المقارنة بين الطرق الثلاث، قام الباحث بالاعتماد على ثلاثة مجموعات من البيانات الحقيقة و التي تم دراستها في الدراسة السابقة، و سيتم احتساب قدرة التشخيص لكل طريقة على هذه البيانات بدون قيم شاذة، و من ثم تعمد الباحث تلويث بيانات المتغير المعتمد في كل مجموعة بنسبة من القيم الشاذة تتجاوز

هذه البيانات مفيدة في هذه الدراسة تعمدنا بتلويث مشاهدات الإنفاق في موقع مختلف  
بنسبة من الشواد (%) للمجموعة الأولى  
1995، ) و بنفس النسبة الأخيرة  
(Pindyck and Rubinfeld ,1997)  
Montgomery, للمجموعة الثالثة  
Gujarati (%) للمجموعة الثانية (Peck and Vining ,2001)

### ٣.١ بيانات Pindyck and Rubinfeld (١٩٩٧) : هذه البيانات قدمت من خلال

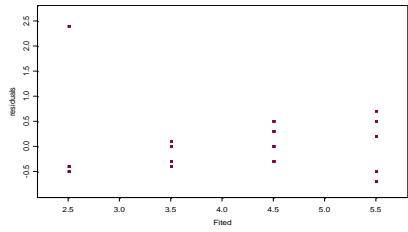
Pindyck and Rubinfeld (1997)<sup>(١)</sup> لدراسة العلاقة بين الدخل و الإنفاق المنزلي

لعشرين أسرة تتوزع دخولها إلى أربعة مجموعات دخل كالتالي:

جدول(٣-١) بيانات Pindyck and Rubinfeld (١٩٩٧) الأصلية و المعدلة

النوع	النوع	النوع	النوع	النوع	النوع
١	٥	١.٨ (٤.٩)	١١	١٥	٤.٢
٢	٥	٢.٠	١٢	١٥	٤.٢
٣	٥	٢.٠	١٣	١٥	٤.٥
٤	٥	٢.٠	١٤	١٥	٤.٨
٥	٥	٢.١	١٥	١٥	٥.٠
٦	١٠	٣.١	١٦	٢٠	٤.٨
٧	١٠	٣.٢	١٧	٢٠	٥.٠
٨	١٠	٣.٥	١٨	٢٠	٥.٧
٩	١٠	٣.٥	١٩	٢٠	٦.٠
١٠	١٠	٣.٦	٢٠	٢٠	٦.٢

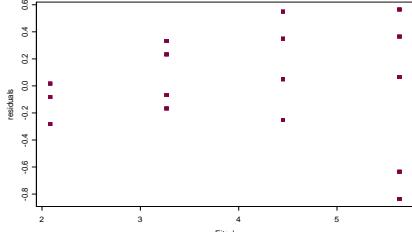
بوضع قيمة شاذة واحدة بتعتمد و هي القيمة المصنفة (١) ، والشكل(٢) يوضح ظهور مشكلة عدم التجانس و القيم الشاذة من خلال رسم العلاقة بين بوافي الانحدار و القيم المقدرة.



الشكل(٣-٢) بوافي الانحدار و القيم المقدرة لبيانات Pindyck and Rubinfeld المعدلة.

بيانات **Gujarati (١٩٩٥)** : هذه البيانات قدمت بواسطة Gujarati (١٠) ، تتكون من ٣٠ مشاهدة (1995) تبين الانفاق المتتنوع للأشخاص حسب مستويات

من المفترض الاشخاص ذوي الدخل المرتفع يكون انفاقهم المنزلي اكثر تنوعا نسبيا قياسا بذوي الدخول المنخفضة. الشكل(٣-١) ادناه يبين عرضا لنمط عدم التجانس في هذه البيانات.



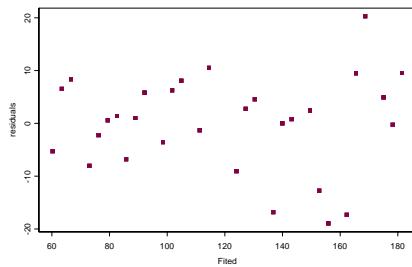
الشكل(٣-١) البوافي و القيم المقدرة لبيانات Pindyck and Rubinfeld الاصلية

الشكل(٣-٢) لا يبدو واضحا مثل الشكل(٣-١) في بيان مشكلة عدم التجانس ، لذلك نحن بالتأكيد بحاجة الى اختبارات للحصول على استنتاج من وجود المشكلة او عدمها.لقد قام الباحث

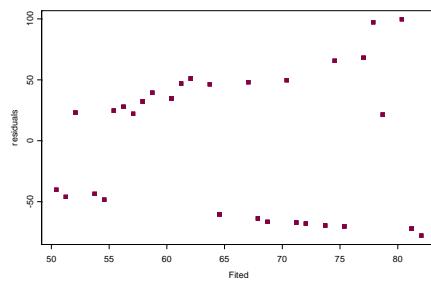
دخلهم، ولذلك يمكن ان نتوقع هذا النوع او التغير ليس ثابتا. وقد تعمدنا وضع ستة قيم شاذة في مجموعة البيانات حل محل مشاهدات الانفاق على لاستهلاك (1,2,3,4,29,30) الظاهرة بين قوسين في الجدول ادناء.

جدول (٣-٢) بيانات Gujarati الاصلية و المعدلة

		النوع	النوع	النوع	النوع			النوع	النوع	النوع
		النوع	النوع	النوع	النوع	النوع	النوع	النوع	النوع	النوع
1	55 (10)	80	11	74	105	21	152	220		
2	65(10)	100	12	110	160	22	144	210		
3	70(10)	85	13	113	150	23	175	245		
4	80(10)	110	14	125	165	24	180	260		
5	79	120	15	108	145	25	135	190		
6	84	115	16	115	180	26	140	205		
7	98	130	17	140	225	27	178	265		
8	95	140	18	120	200	28	191	270		
9	90	125	19	145	240	29	137(100)	230		
10	75	90	20	130	185	30	189 (100)	250		



الشكل (٣-٣) الباقي مقابل القيم المقدرة لبيانات Gujarati الأصلية



الشكل(٤-٤) بواقي الانحدار مقابل القيم المقدرة لبيانات Gujarati المعدلة

الشكل(٣-٣) و(٤-٤) يظهر رسمًا لبواقي الانحدار و القيم المقدرة لبيانات الأصلية و المعدلة للإنفاق على الاستهلاك ، و يبدو مشكلة عدم تجانس التباين واضحة في الشكل- ٣- لبيانات الأصلية على عكس الشكل (٤-٤) الذي لا يظهر هذه المشكلة بسبب وجود القيم الشاذة.

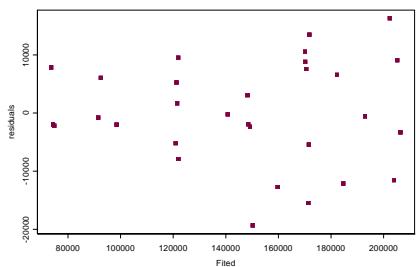
المطاعم وقيمة الانفاق على الاعلان،  
مرة اخرى وضعنا ستة قيم شاذة في  
مجموعة البيانات باستبدال الحالات ( 1,2,3,27,28,30 )  
المبنية داخل الاقواس في جدول ( ٣-٣ ) للبيانات الاصلية  
العلاقة بين دخل او ايرادات احد والمعندة.

**Montgomer, Peck and Vining** : هذه البيانات قدمت  
بواسطة Montgomery, Peck and Vining ( ٢٠٠١ )<sup>(١٥)</sup> و التي تبين  
البيانات الاصلية ( ٣-٣ ) جدول ( ٣-٣ ) بيانات Montgomer, Peck and Vining

جدول ( ٣-٣ ) بيانات Montgomer, Peck and Vining الاصلية و المعدلة

الدخل	الانفاق على الاعلان	الدخل	الانفاق على الاعلان	الدخل	الانفاق على الاعلان
81464 (300000)	3000	131434	9000	178187	15050
72661(300000)	3150	140564	11345	185304	15200
72344(300000)	3085	151352	12275	155931	15150
90743	5225	146926	12400	172579	16800
98588	5350	130963	12525	188851	16500
96507	6090	146630	12310	192424	17830
126574	8925	147041	13700	203112(300000)	19500
114133	9015	179021	15000	192482(21431)	19200
115814	8885	166200	15175	218715	19000
123181	8950	180732	14995	214317 (21431)	19350

و المعدلة لبيانات مبيعات اغذية المطعم  
و يبدو عدم تجانس التباين واضحا في  
البيانات الاصلية الشكل(٣-٥) و لا يبدو  
ذلك بظهور القيم الشاذة.



#### ٤. مناقشة و تحليل النتائج

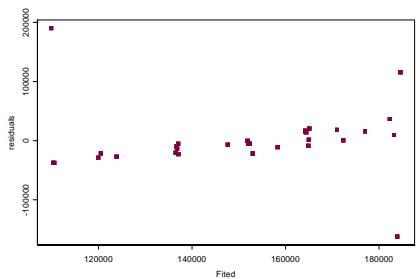
##### ٤.١ نتائج بيانات Pindyck and Rubinfeld

(١٩٩٧): الجدول (١-١) يبين النتائج التي تم الحصول عليها  
لأقيام اختبار فيشر للتجانس وقوة  
الاختبار  $P$  من خلال تطبيق الاختبار  
التقليدي GQ و المعدل  
ال MGQ و المقترن اختبارنا

على ٢٠ مشاهدة غير متجانسة لبيانات  
(١٩٩٧) Pindyck and Rubinfeld  
الخالية من الشوادرة و الملوثة بنسبة  
٥٥٪ مرة اخرى عند مستوى معنوية  
٠٠٥

الشكل(٣-٥) الباقي مقابل القيم

المقدرة لبيانات Montgomery, Peck  
and Vining الاصلية



الشكل (٣-٦) الباقي مقابل القيم

المقدرة لبيانات Montgomery, Peck  
and Vining المعدلة

الشكل (٣-٥) و (٣-٦) يظهران رسميا  
للباقي و القيم المقدرة لبيانات الاصلية

جدول (٤-١) نتائج الاختبارات لبيانات Pindyck and Rubinfeld الاصلية و المعدلة.

Test <b>FC=4.2838</b>	بدون قيم شاذة		بوجود القيم الشاذة	
	F	p - value	F	p - value
<b>Classical GQ</b>	١٢.١٨٦	٠.٠٠٣٨	٤.٠٩٢٨	٠.٠٥٥
<b>MGQ</b>	٣٦.٠٠٠	٠.٠٠٠٢	١٨	٠.٠٠٣٥
<b>RMGQ</b>	<u>٥٦.٥٤٢</u>	<u>٠.٠٠٠١</u>	<u>٥٦.٥٤٢</u>	<u>٠.٠٠٠١</u>

الجدول (٤-١) يظهر ان طرق الاختبار  $F$  (٥٦.٥٤٢) و هي الاعلى بين قيم  $F$  التي تناظر القيمة الصغر (٠.٠٠٠١) في البيانات الاصلية وال معدلة. وذلك ترفض الفرضية  $H_0$  بين نظيراتها في الاختبارات الثلاث قد شخصت مشكلة عدم التجانس في البيانات الاصلية الخالية من القيم الشاذة على الرغم من اختلاف درجة الشدة التجانس من اختبار الى اخر، اذ ان جميع قيم  $F$  المحسوبة بالطرق الثلاث تجاوزت القيمة الحرجة لها (٤.٢٨٣٨) و التي بدت في الطريقة المقترحة نحو الصفرية القائلة بتجانس التباين على الاخرى. لذلك ترفض الفرضية  $H_0$  الصافية القائلة بتجانس التباين على صو النتائج الظاهرة للاختبارات الثلاث في هذه الدراسة.

بعد تلویث البيانات بقيمة شاذة واحدة شكلت مابنسبة ٥٥% تقريباً منها، الملاحظ ان الطريقة التقليدية للاختبار فشلت في تشخيص المشكلة ، أي قبلت الفرضية الصفرية ، بينما رفضت هذه الفرضية من الطريقتين MGQ و RMGQ اللتان تشخصتا المشكلة ولكن بدرجة مختلفة من الشدة. الواضح ان RMGQ شخصت المشكلة بنفس درجة

التجانس في هذه العينة المكونة من  
٢٠ مشاهدة فقط.

**Gujarati** ٤.٢ نتائج بيانات  
(١٩٩٥): بين الجدول(٤-٢) نتائج  
اختبارات عدم التجانس في هذه الدراسة  
على بيانات Gujarati (١٩٩٥) قبل  
وبعد تلویتها بنسبة (%)٢٠ من القيم  
الشاذة عند مستوى معنوية (٠٠٠٥). من  
الواضح ان طرق الاختبار الثلاث  
شخصت مشكلة عدم التجانس في هذه  
البيانات قبل تلویتها و ذلك برفضها  
للفرضية الصفرية بسبب ان قيم اختبار  
**F** للطرق الثلاث قد تجاوزت القيمة

شدة عدم التجانس فيما لو كانت البيانات  
خالية من القيم الشاذة و بنفس الدقة  
والقوّة ، بينما MGQ شخصتها بدرجة  
شدة اقل و السبب يعود في ذلك الى  
نسبة التشذيب الافتراضية التي اعتمدتها  
هذه الطريقة هي (%)١٠ و هذا يؤدي  
بالتأكيد الى خسارة مابنسبة ٥% من  
البيانات النظيفة بمعنى بتر مشاهدتين  
من هذه البيانات واحدة شاذة و الاخرى  
اعتيادية. أما RMGQ لقد شخصت  
النسبة الفعلية للشاذ و كانت قيمة شاذة  
واحدة فقط و قامت بتشذيب هذه القيمة  
ما انعكس على درجة شدة عدم

الفرضية الصفرية على الرغم من تلوث البيانات. من ناحية أخرى يبين الجدول ان قوة الاختبار للطريقة المقترحة كانت افضل من الطرق الأخرى سواء كانت البيانات ملوثة او غير ملوثة.

الدرجة لها (٢.٩٧٨) ، الا ان الطريقة التقليدية و طريقة MGQ قد اخفقتا في تشخيص المشكلة بقولها لفرضية الصفرية بعد تلوث البيانات بـ (%)٢٠ من القيم الشاذة ، في حين الطريقة المقترحة بقى على رفضها.

جدول(٤-٢) نتائج الاختبارات لبيانات Gujarati الاصلية و المعدلة

Test FC=2.978	بدون قيم شاذة		بوجود القيم الشاذة	
	F	p - value	F	p - value
Classical GQ	3.889	0.022	١.٣٧٠	٠.٣١٤
MGQ	3.933	0.021	١.٠٤٢	٠.٤٧٤
RMGQ	<u>15.245</u>	<u>0.0001</u>	<u>٣.٩٠٣</u>	<u>٠.٠٢١</u>

(2001) Montgomery, Peck and Vining يوضح وجود و عدم وجود القيم الشاذة ، من النتائج يتضح جميع قيم F لاختبارات الثلاث هي اكبر من قيمتها

٤.٣ بيانات Montgomery, Peck and Vining يتضمن الجدول(٣-٤) نتائج اختبارات الدراسة لتشخيص مشكلة عدم تجانس التباين لبيانات

اظهرتها لم تتجاوز قيمتها الحرجة عند الحرجة (٢.٩٧٨) عند مستوى معنوية (٠٠٥) بدون قيم شاذة و هذا يعني مستوى معنوية (٠٠٥) و بدرجات حرياتها. لكن الطريقة المقترحة تمكنت الطرق قيد الدراسة رفضت الفرضية الصفرية و شخصت وجود مشكلة عدم من رفض الفرضية الصفرية و التجانس في هذه البيانات و لكن تشخيص مشكلة عدم التجانس على بدرجات مختلفة لشدة عدم التجانس. الرغم من وجود القيم الشاذة بالإضافة إلى ان قيمة  $P$  التي كانت اصغر من بوجود مانسيته (%) من القيم الشاذة الاخرى بوجود القيم الشاذة او بدونها ، الامر الذب يفسر دقة وقوه تشخيص فشل الاختبار الاول والثاني في تشخيص المشكلة من خلال قبوله للفرضية الصفرية لأن قيم فيشر التي

الطريقة المقترحة.

جدول(٤) نتائج الاختبارات لبيانات Montgomery, Peck and Vining الاصلية

و المعدلة

Test	بدون قيم شاذة		بوجود القيم الشاذة	
	F	p - value	F	p - value
FC=2.978				
Classical GQ	4.037	0.019	1.468	0.278
MGQ	3.681	0.026	2.825	0.058
RMGQ	<u>4.037</u>	<u>0.019</u>	<u>4.715</u>	<u>0.011</u>

المتغير التابع و بعد ترتيب قيم المتغير التوضيحي تصاعديا تمركزت ثلاثة من القيم الشاذة على رأس القيم العظمى و الثلاث الأخرى تذيلت القيم الصغرى ، وبعد حذف (٦) مشاهدات من وسط الترتيب وتجزئة باقى المشاهدات (٢٤) إلى عينتين جزئيتين كل عينة (١٢) مشاهدة شكلت القيم الشاذة مابنسبة (٢٥٪) في كل عينة. أما بالنسبة لبيانات Gujarati (١٩٩٥) شكلت الشواد في العينة الجزئية الأولى (٣٤٪) تقريبا و العينة الجزئية الثانية (١٧٪) تقريبا ، بينما نسبتها في العينة الأصل (٣٠ مشاهدة) هي (٢٠٪). أما بيانات Pindyck and Rubinfeld (١٩٩٧) شكلت الشواد في العينة الجزئية الأولى نسبة قدرها (١٢٥٪) و الثانية (٠٠٠٪). ارتفاع نسبة القيم الشاذة

## ٥. الاستنتاجات و التوصيات

٥.١ الاستنتاجات: من النتائج المبوبة في الجداول (١-٤)، (٤-٢)، (٤-٣) تستنتج ان طريقة RMGQ المقترحة هي الافضل و الاقوى في تشخيص مشكلة عدم التجانس في البيانات قيد الدراسة بوجود او بعده وجود القيم الشاذة. الملاحظ ايضا ان نسبة القيم الشاذة الى حجم العينة قد تتعاظم عند تجزئة العينة الى عينتين جزئيتين بعد ترتيب قيم المتغير التوضيحي تصاعديا. اذ لاحظنا عند تلويث بيانات (2001) Montgomery، Peck and Vining → (٢٠٪) من الشواد في

لا يوجد تأثير قوي في بقاء (٤٠٠٤) من المشاهدة ، عندئذ يكون التشذيب لقيمة شاذة واحدة فقط. أي العودة الى تشذيب مانسيته (١٠%) الامر الذي دفع طريقة MGQ لتشخيص المشكلة بوجود نسبة من الشواد تتوافق مع النسبة المفترضة لديها. بينما طريقة RMGQ المقترحة تشخيص مشكلة عدم التجانس بدقة وقوة مهما اختلفت نسبة القيم الشاذة.

٥.٢ التوصيات : بناء على نتائج البحث و استنتاجاته يوصي الباحث بما يأتي :

١. تبني مقترح تعديل هذا الاختبار لتشخيص مشكلة عدم التجانس في حالة وجود قيم شاذة في البيانات.

٢. لا يعالج هذا المقترن مشكلة القيم الشاذة في جانب المتغيرات التوضيحية مما يعطي فرصة

إلى اكثر من (١٠%) بعض البيانات هو السبب الاساسي في فشل طريقة MGQ في تشخيص مشكلة عدم التجانس لأن طريقة المربعات الصغرى المشذبة تعمل عمل طريقة المربعات الصغرى الاعتيادية بعد تشذيب القيم الشاذة و حيث ان تشذيب مانسيته (١٠%) في بيانات تتجاوز هذه النسبة يجعلها تتحسس المتبقى من الشواد في البيانات، وعلى العكس من ذلك عندما كانت نسبة القيم الشاذة في العينة الجزئية الاولى Pindyck and Rubinfeld لبيانات (١٩٩٧) لاتتجاوز (١٣%) تمكنت MGQ من تشخيص مشكلة عدم التجانس و لكن ليس بالدقة والقوة الكافية لأنها (٣%) من اصل (٨) مشاهدات لا ترقى إلى الواحد الصحيح بمعنى (١٣%) تكافيء (٤٠٠٤) مشاهدة ، لذلك

اخرى للباحثين من دراسة هذه

الحالة.

### المراجع و المصادر

١. الفلاوى، كاظم كامل، القياس  
الاقتصادي النظرية و التحليل، دار  
الصفاء للنشر و التوزيع، ٢٠١١

٢. العيسوى، ابراهيم، القياس و التبؤ  
الاقتصادي، دار النهضة العربية،  
القاهرة، ١٩٧٨

٣. الزبيدي، حسين لطيف، القياس  
الاقتصادي و النظرية الاقتصادية،  
دار الصفا للنشر و التوزيع، عمان ،  
٢٠١١

٤. كاظم، امورى هادى، الاقتصاد  
القياسى التطبيقى، دار الحكمة،  
الموصل، ١٩٩٥ .

٣. مشكلة عدم التجانس و بوجود القيم

الشاذة تظهر ايضا في موضوع  
السلالس الزمنية ، لذلك يوصى  
الباحث بدراسة تحصين اختبار  
كولد فيلد- كوانت لتخفيض مشكلة  
عدم التجانس في بيانات السلسلة  
الزمنية.

٤. لقد تناول الباحث اختبار كولد فيلد  
- كوانت من بين عدة اختبارات  
اخرى تستخدم لتشخيص مشكلة  
عدم التجانس، لذلك يوصى الباحث  
إلى تحصين هذه الاختبارات  
لمقاومة القيم الشاذة من خلال  
دراسات جديدة.

٥. محبوب، عادل عبد الغني ،  
الاقتصاد القياسي، دار الكتب للطباعة  
والنشر، الموصل، ١٩٨٢.
10. Gujarati, D. (1995). Basic Econometrics, 4rd ed. New York: McGraw-Hill.
11. Hampel, F.R. 1971., A general Qualitative Definition of robustness, The Annals Mathematical of Statistics., 42,1887-1896.
12. Imon, A. H. M. R. (2002). On deletion residuals, Calcutta Statistical Association Bulletin, 52, 65–79.
13. Imon, A.H.M.R. (2005). Identifying multiple influential observations in linear regression, Journal of Applied Statistics, 32, 929 – 946.
14. Imon, A.H.M.R. (2008). Deletion residuals in the detection of heterogeneity of variances in linear regression, (Accepted for
- 6.Chatterjee, S. and Hadi, A.S. (1988). *Sensitivity Analysis in Linear Regression*, New York.: Wiley.
- 7.Chatterjee, S. and Hadi, A.S. (2006). Regression Analysis By Examples, 4th ed., New York.: Wiley.
- 8.Cook, R. D. and Weisberg S. (1983). Diagnostics for heteroscedasticity in regression, Biometrika, 70, 1 – 10.
- 9.Goldfeld, S.M. and Quandt, R.E. (1965). Some tests for homoskedasticity, Journal of the American Statistical Association, 60, 539 – 547.

- Presence of Outliers. *Journal of Mathematics and Statistics* 4 (4): 277-283, 2008.
19. Ryan, T. P., (1997). Modern Regression Methods, New York: Wiley.
20. White, H. (1980). Heteroscedasticity-consistent covariance matrix estimator and a direct test for heteroscedasticity, *Econometrica*, 48, 817-838.
- publication), *Journal of Applied Statistics*.
15. Montgomery, D.C., Peck, E.A. and Vining, G.G. (2001). Introduction to Linear Regression Analysis, 3rd ed., New York: Wiley.
16. Pindyck, S. R and Rubinfeld, L. D. (1997). Econometric Models and Econometric Forecasts, 4th ed., New York: Irwin/McGraw-Hill.
17. Rousseeuw, P.J. and Leroy, A. (1987). Robust Regression and Outlier Detection, New York.: Wiley.
18. Rana;Md. Sohel, Midi;Habshah, Imon;A.H.M. A Robust Modification of the Goldfield-Quandt Test for Detection of Heteroscedasticity in the