

تعديل اختبار كولد فيلد- كوانت المعدل لتشخيص مشكلة عدم التجانس بوجود الشواذ

حسن سامي عريبي الشمري

قسم الاحصاء

كلية الادارة و الاقتصاد - جامعة القادسية

الديوانية - العراق

المستخلص:

Abstract:

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 propose 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 2008 تعديل اختبار كولد فيلد-كوانت لتشخيص مشكلة عدم التجانس بوجود قيم شاذة، ذلك بالاستفادة من طريقة المربعات الصغرى المشدبة بتثبيت نسبة التشذيب عند (10%). هذه النسبة قد تسبب خسارة في البيانات اذا كانت نسبة الشواذ اقل من (10%) و فشل في حصانة التقدير اذا كانت النسبة اكثر من ذلك. تهدف هذه الدراسة الى حل هذه المشكلة من خلال مقترح تعديل اخر على اختبار كولد فيلد -كوانت المعدل. لقد اظهرت النتائج كفاءة عالية للتعديل المقترح في تشخيص المشكلة بوجود نسب مختلفة من القيم الشاذة.

الكلمات المفتاحية: كولد فيلد- كوانت، LTS، القيم الشاذة، عدم التجانس، اختبار F .

1. المقدمة

واحدة من فرضيات نموذج الانحدار الخطي ان حد الخطأ العشوائي له تباين ثابت عند كل مشاهدة و التي يطلق عليها ايضا فرضية تجانس تباين الخطأ العشوائي. قد تتحقق هذه الفرضية على ارض الواقع او قد لا تتحقق، من حيث ان تباين حد الخطأ العشوائي في الكثير من الاحيان لا يكون ثابتاً(1).

انتهاك فرضية تجانس تباين الخطأ العشوائي سيؤدي الى ظهور مشكلة عدم ثبات تجانس التباين أي أن

$$\text{Var}(u_i) \neq \sigma_u^2$$

عندئذ ترفض الفرضية الصفرية التي تفترض تجانس التباين كالاتي:

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

و تحقق الفرض البديل للفرضية الصفرية كالاتي:

$$\sigma_i^2 \neq \sigma_j^2 \text{ for at least one case}$$

تحقق الفرضية البديلة بظهور مشكلة عدم التجانس التي يمكن تصنيفها الى نوعين من عدم التجانس الا وهما المتضاعف حسب نموذج هارفي (1976) والذي صيغته العامة

$$\sigma_i^2 = \sigma^2 \exp(z_i/\alpha)$$

حيث z_i متجه المتغير التوضيحي الذي ربما يكون او لا يكون مترامن مع متجه المتغير X ، اما α هو متجه المعالم. ويمكن اعادة الصيغة اعلاه بشكل مبسط كالاتي

$$: (Rana and et.al, 2008) (18):$$

$$\sigma_i^2 = \sigma^2 c_i$$

حيث c_i كمية غير معلومة.

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

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 للمعاملات في ظل وجود مشكلة عدم التجانس.

طرق التشخيص المقترحة صنفتم الى نوعين هما: طرق تخطيطية تعتمد على رسم العلاقة في تحديد المشكلة مثل

كلاهما معا⁽⁶⁾. و اسباب ظهورها متعددة و مختلفة و لكن الحقيقة الثابتة حسب رأي معظم الخبراء في مجال الاحصاء انه لا توجد بيانات حقيقية تخلو من القيم الشاذة لكن تأثيرها مختلف حسب نسبة ظهورها قياسا باحجام العينات.

تختلف مقاومة طرق التقدير للقيم الشاذة من طريقة الى اخرى الا ان طريقة المربعات الصغرى لا تمتلك اية قدرة لمقاومة قيمة شاذة واحدة على الاقل (Hampel,1971)⁽¹¹⁾، لذلك لا بد من استخدام طرق التقدير الحصينة بدلا من طريقة المربعات الصغرى عند ظهور القيم الشاذة.

(Rousseeuw and Leroy,1987)⁽¹⁷⁾.

اقترحوا طريقة المربعات الصغرى المشدبة (LTS) التي يمكن ايجاد مقدراتها من خلال تقليل دالة الهدف ادناه الى اقل مايمكن

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

عندما $r_{(i)}^2$ تمثل مربعات البواقي المرتبة و h يمكن تحديدها كالاتي:

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

حيث α هي النسبة المئوية للقيم الشاذة او هي نسبة تشذيب المشاهدات. اما n

(Imon,2008)⁽¹⁴⁾ و (Chatterjee and Hadi ,2006)⁽⁷⁾ و غيرهم.

اما الطرق الاخرى هي طرق تحليلية مثل كولدفيلد-كوانت (1965)⁽⁴⁾، جليسير (1969)⁽⁴⁾، بارك (1966)⁽⁴⁾، اختبار سبيرمان

(1953)⁽⁴⁾، (White ,1980)⁽²⁰⁾ و غيرها من الاختبارات. للمزيد يرجى مراجعة ((العيسوي، ابراهيم، 1978)⁽²⁾، (الزيدي، حسين

لطيف، 2011)⁽³⁾، (كاظم، اموريهادي،

(1995)⁽⁴⁾، (محبوب، عادل عبد الغني، 1982)⁽⁵⁾، Cook (and Weisberg 1983)⁽⁸⁾ .

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 يعطل 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,2005,2008,2002)^(12,13,14)

. و تصنف الشواذ حسب مواقع ظهورها في جانب المتغير المعتمد او المستقل او

من اهم الدراسات السابقة التي تطرقت الى موضوع تحسين اختبار كولدفيلد-كوانت هي الدراسة التي تقدم بها (Rana and et.al, 2008)⁽¹⁸⁾ و التي اختارت توظيف طريقة المربعات الصغرى المشذبة لتحسين وتعديل اختبار كولدفيلد-كوانت الا ان هذه الطريقة اعتمدت على احتساب قيمة الوسيط لمربعات بواقى 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 يمكن مقارنتها بقيمها الحرجة عند درجات حرياتها و مستوى معنوية محدد لرفض او قبول الفرضية القائلة بتجانس التباين. تطرح هذه الدراسة تساؤلا عن قدرة اختبار كولد فيلد- كوانت التقليدي والمعدل في تشخيص مشكلة عدم التجانس في بيانات انحدار خطي ملوثة بقيم شاذة. اذ ان الدراسة السابقة افترضت نسبة التشذيب هي (10%) في أي حال من الاحوال. هذه النسبة كانت مثار لعدد من التساؤلات، مثلا لو كانت نسبة الشواذ تقل عن (10%)، الا تتسبب النسبة

هي حجم العينة و k هي عدد معاملات نموذج الانحدار. (Rousseuw)⁽¹⁷⁾ and Leroy,1987) وجدا ان النقطة 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 تبنت هذه الطريقة ان نسبة قدرتها على مقاومة القيم الشاذة تصل الى اقل من 50% بقليل تبعا الى نسبة القيم الشاذة في البيانات الاصلية⁽¹⁸⁾.

اختبار كولدفيلد-كوانت و احد من اهم الاختبارات المستخدمة في تشخيص مشكلة عدم التجانس الا ان هذا الاختبار يحتاج الى احتساب متوسط مربعات بواقى الانحدار لعينتين جزئيتين لقياس قوة عدم التجانس و خوارزمية هذا الاختبار تعتمد على طريقة المربعات الصغرى للوصول الى متوسط مربعات البواقى، وبما ان هذه الطريقة تتحسس القيم الشاذة و تعد طريقة غير حصينة، لذلك ستعطي مقدرات مظلة و بالتالي ربما ستؤدي الى فشل الاختبار في تشخيص مشكلة عدم التجانس.

ظل وجود نسب مختلفة للقيم الشاذة في بيانات المتغير المعتمد تحديداً.
2. الاستفادة من وسيط مربعات بواقي الانحدار بدلا من متوسط مربعات البواقي او وسيط مربعات الانحدار المحذوفة في ايجاد قيمة الاختبار *F*.

1.2 اهمية الدراسة

اغلب الاحيان تظهر مشكلة عدم التجانس في بيانات الاقتصاد القياسي مثل العلاقة بين الدخل و الانفاق، و نظرا الى ان ظهور هذه المشكلة يجعل من تقديرات المربعات الصغرى تقديرات متحيزة لا يمكن الاعتماد عليها، لذلك يلجأ العديد من الباحثين في مجال الاقتصاد و الاحصاء الى اختبار بيانات العلاقة الخطية للانحدار قبل الشروع بعملية التقدير، لكن الاعتماد على عملية تشخيص التجانس من عدمه غير كافية لان وجود القيم الشاذة ربما يفقد كفاءة طرق التشخيص مثل طريقة كولدفيلد- كوانت التقليدية مما يجعلها لا تتحسس عدم التجانس الحاصل في البيانات. من جانب اخر ابتعاد القيمة الشاذة عن معظم البيانات و اتباعها لتوزيع اخر، يدفع بقية البيانات التي تتبع التوزيع

الافتراضية للتشذيب بخسارة في بعض البيانات النظيفة، من جانب اخر اذا بلغت نسبة الشواذ (15%) او اكثر، الا تتسبب نسبة التشذيب الافتراضية ببقاء (5%) او اكثر من الشواذ في البيانات، مما تكون سببا كافية 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 -كوانت المعدل ليكون اكثر حصانة. بما ان هناك امكانية لاستخدام الطرق الحصينة استخدمنا طريقة المربعات الصغرى المشدبة وفقا للصيغة الثانية المتوفرة في برنامج S-PLUS2000 كتعديل على هذا الاختبار. لذلك تتمحور اهداف هذه الدراسة حول النقاط الاتية:

1. تحسين اختبار كولدفيلد - كوانت باستخدام طريقة المربعات الصغرى المشدبة ليكون اكثر حصانة في تشخيص مشكلة عدم التجانس في

المقترحة لا تفترض نسبة ثابتة للتشذيب بل تتعاطى مع النسب الفعلية للشواذ في بيانات العينة الجزئية الأولى أو الثانية. هذا الاجراء يجعل من الخوارزمية المقترحة أكثر واقعية في التعامل مع البيانات و تقدير معاملات انحدار أكثر دقة لأنها لا تسبب بخسارة في بعض البيانات النظيفة و تشذب الشاوذ منها بنسبها الفعلية أو تجاور نسبها الفعلية حسب حجم العينة الجزئية من جانب و نسب الشواذ في هذه العينة من جانب آخر، وكالاتي:

1. ترتيب المشاهدات المتغير التوضيحي تصاعديا.
2. حذف (c) من المشاهدات الوسطية من بيانات العينة ويفضل حذف مانسبته $1/5$ من المشاهدات. حيث (c) عدد صحيح موجب.
3. تقسم المشاهدات الباقية (n - c) الى $(n - c)/2$ عينتين جزئيتين متساويتين تتطوي الأولى على قيم X_i الصغيرة، والثانية على قيم X_i الكبيرة.
4. تحديد نسبة القيم الشاذة في كل عينة جزئية بعد تشخيصها باحدى طرق التشخيص الحصينة. كأن نأخذ انحراف مشاهدات العينة الجزئية عن وسطها و نقسمها على الانحراف المطلق للوسيط.

الطبيعي الى التمركز في زاوية ما، الامر الذي لا يمكن ملاحظته من خلال الرسم التخطيطي لعلاقة الاخطاء العشوائية بالقيم المقدرة باستخدام المربعات الصغرى. لقد عالجت خوارزمية MGQ هذه المشكلة اذا كانت نسبة الشواذ لا تتجاوز (10%) ومن هنا تبرز اهمية مقترح هذا البحث في تعديل اختبار كولد فيلد-كوانت لتشخيص مشكلة عدم التجانس بوجود نسب مختلفة من القيم الشاذة سواء زادت هذه النسبة او قلت عن (10%)، و ذلك بجعل عملية تشخيص الشواذ ديناميكية مما يجعل الاختبار أكثر حصانة وأكثر مقاومة للقيم الشاذة .

3. خوارزمية تعديل اختبار كولد فيلد -

كوانت المعدل (RMGQ)

هذه الخوارزمية هي تعديل لخوارزمية MGQ المقترحة من قبل (Rana⁽¹⁸⁾ and *et.al*, 2008) لغرض تشخيص مشكلة عدم التجانس في بيانات الانحدار لعينات كبيرة الحجم الى حد ما في ظل وجود بعض الشواذ في متغير الاستجابة. من حيث ان خوارزمية MGQ افترضت تشذيب ما نسبته (10%) من البيانات المجزئة بعد ترتيبها تصاعديا. الخوارزمية

الآخرى التي تحتوي على الأقل على مشاهدة شاذة واحدة من شأنها ان تعظم متوسط مربعات الخطأ.

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

7. ايجاد قيمة اختبار F المحسوبة وفقا للصيغة الآتية:

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

و بما ان توزيع معظم البيانات هو الطبيعي لذلك بالتاكيد سيكون وسيط مربعات البواقي لكلا المجموعتين سيتبع توزيع مربع كاي، اذا يمكن ان نستخدم اختبار F تحت مستوي معنوية محدد و بدرجة حرية $(n - c - 2k)/2$ لاختبار صحة فرضية العدم بعدم وجود مشكلة عدم التجانس اذا كانت قيمة F المحسوبة اقل من قيمة F الحرجة في الجداول الاحصائية.

هذا الاجراء في احتساب قيمة F سيتم ايضا باستخدام الطريقة التقليدية للاختبار و باستخدام طريقة MGQ المقترحة من قبل (Rana and et.al, 2008)⁽¹⁸⁾.

و لغرض المقارنة بين الطرق الثلاث، قام الباحث بالاعتماد على ثلاثة مجموعات من البيانات الحقيقية و التي تم دراستها في

بمعنى اخر اخذ الدرجة المعيارية للملاحظات المرتبة بالاعتماد على الوسيط، فاذا كانت القيمة المطلقة للدرجة المعيارية تجاوزت العدد (3)، فتعد هذه القيمة شاذة. عندئذ يمكن تحديد نسبة الشواذ من خلال تحديد عدد الدرجات المعيارية الاكبر من العدد (3) و قسمتها على حجم العينة.

5. تحديد الحجم الجديد لكل عينة جزئية بعد تحديد نسبة الشواذ (α) كالاتي:

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

حيث سيتم الاعتماد على (h) كحجم جديد للعينة الجزئية، الذي قد يختلف بين العينتين و هو نفس حجم المجموعات الفرعية التي ستشكل بتوافق (h) من حجم كل عينة جزئية حسب خوارزمية طريقة (LTS).

ثم يتم تقدير معاملات كل مجموعة باستخدام طريقة المربعات الصغرى المشنبة (LTS) بالاعتماد على الحجم الجديد لكل عينة جزئية. لو دققنا النظر جيدا سنلاحظ ان هناك ثمة مجموعة جزئية بحجم (h) تحتوي على بينانات خالية من الشواذ. هذه المجموعة ستجعل من متوسط مربعات الخطأ اقل مايمكن على العكس من قريناتها من المجموعات

و الدخل وهي نفس البيانات التي تم استخدامها في الدراسة السابقة، و لجعل هذه البيانات مفيدة في هذه الدراسة تعمدنا بتلويث مشاهدات الاتفاق في مواقع مختلفة بنسبة من الشواذ (5%) للمجموعة الاولى (Pindyck and Rubinfeld 1997) و (20%) للمجموعة الثانية (Gujarati 1995) و بنفس النسبة الاخيرة للمجموعة الثالثة Montgomery, (Peck and Vining 2001).

3.1 بيانات Pindyck and Rubinfeld (1997): هذه البيانات قدمت من خلال Pindyck and Rubinfeld (1997)⁽¹⁶⁾ لدراسة العلاقة بين الدخل و الاتفاق المنزلي لعشرين اسرة تتوزع دخولها الى اربعة مجموعات دخل كالاتي:

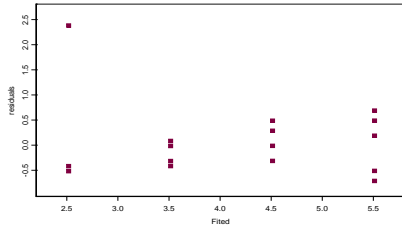
الدراسة السابقة، و سيتم اختبار قدرة التشخيص لكل طريقة على هذه البيانات بدون قيم شاذة، و من ثم تعمد الباحث تلويث بيانات المتغير المعتمد في كل مجموعة بنسبة من القيم الشاذة تتجاوز النسبة المفترضة في الدراسة السابقة و التي كانت (10%) من اجمالي حجم العينة. اما المفاضلة بين الطرق الثلاث سيتحدد على ضوء قدرة الطريقة في رفض الفرضية الصفرية التي تشير الى تجانس التباين في حالة وجود او عدم وجود القيم الشاذة. و للتأكد من صحة ودقة وقوة الاختبار المستخدم سيتم احتساب قيمة p لكل طريقة لبيان أي من الطرق اكثر قوة ودقة في تشخيص مشكلة عدم التجانس للاخطاء العشوائية.

3. تحليل البيانات:

لقد تم دراسة ثلاثة مجاميع من البيانات الحقيقية المختبرة من قبل عدد من الباحثين في مجال الاحصاء و الاقتصاد القياسي تتعلق في دراسة العلاقة الخطية بين الاتفاق

جدول(3-1) بيانات Pindyck and Rubinfeld الاصلية و المعدلة

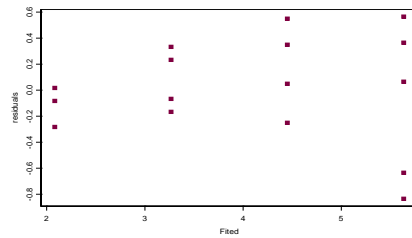
التسلسل	الدخل	الاتفاق المنزلي	التسلسل	الدخل	الاتفاق المنزلي
1	5	1.8 (4.9)	11	15	4.2
2	5	2.0	12	15	4.2
3	5	2.0	13	15	4.5
4	5	2.0	14	15	4.8
5	5	2.1	15	15	5.0
6	10	3.1	16	20	4.8
7	10	3.2	17	20	5.0
8	10	3.5	18	20	5.7
9	10	3.5	19	20	6.0
10	10	3.6	20	20	6.2



الشكل (3-2) بواقي الانحدار و القيم المقدره لبيانات Pindyck and Rubinfeld المعدلة

الشكل(3-2) لا يبدو واضحا مثل الشكل(3-1) في بيان مشكلة عدم التجانس، لذلك نحن بالتاكيد بحاجة الى اختبارات للحصول على استنتاج من وجود المشكلة او عدمها. لقد قام الباحث بوضع قيمة شاذة واحدة بتعمد وهي

من المفترض الاشخاص ذوي الدخل المرتفع يكون انفاقهم المنزلي اكثر تنوع نسبيا قياسا بذوي الدخول المنخفضة. الشكل(3-1) ادناه يبين عرضا لنمط عدم التجانس في هذه البيانات.



الشكل(3-1) البواقي و القيم المقدره لبيانات Pindyck and Rubinfeld الاصلية

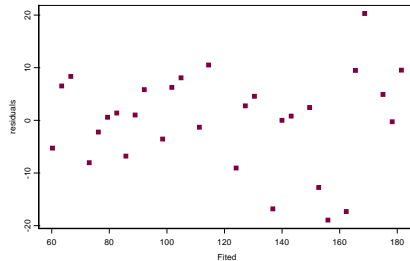
للاشخاص حسب مستويات دخولهم ولذلك يمكن ان نتوقع هذا التنوع او التغير ليس ثابتاً وقد تعمدنا وضع ستة قيم شاذة في مجموعة البيانات حلت محل مشاهدات الانفاق علماً لاستهلاك (1,2,3,4,29,30) الظاهرة بين قوسين في الجدول ادناه.

القيمة المصنفة (1)، و الشكل (2-3) يوضح ظهور مشكلة عدم التجانس و القيم الشاذة من خلال رسم العلاقة بين بواقي الانحدار والقيم المقدره.

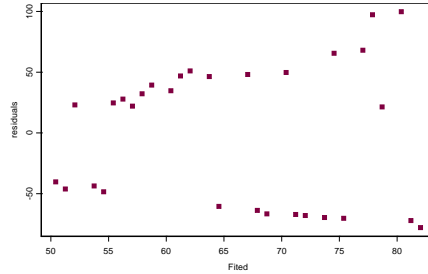
Gujarati بيانات (1995): هذها البيانات تقدمت بواسطة (1995) Gujarati⁽¹⁰⁾، تتكون من 30 مشاهدة تبين الانفاق المتنوع

جدول (3-2) بيانات Gujarati الاصلية و المعدلة

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

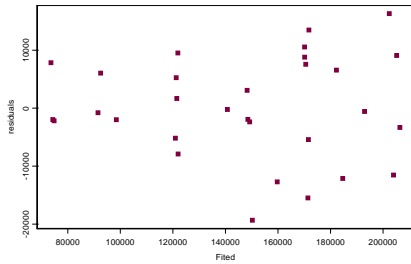


الشكل (3-3) البواقي مقابل القيم المقدره لبيانات Gujarati الاصلية



الشكل (3-4) بواقي الانحدار مقابل القيم المقدرة لبيانات Gujarati المعدلة

مجموعة البيانات باستبدال الحالات (1,2,3,27,28,30) المبينة داخل الاقواس في جدول (3-3) للبيانات الاصلية والمعدلة.



الشكل (3-5) البواقي مقابل القيم المقدرة لبيانات Montgomery, Peck and Vining الاصلية

الشكل (3-3) و (3-4) يظهر رسماً لبواقي الانحدار و القيم المقدرة للبيانات الاصلية و المعدلة للاتفاق على الاستهلاك، ويبدو مشكلة عدم تجانس التباين واضحة في الشكل - 3 - للبيانات الاصلية على عكس الشكل (4-3) الذي لا يظهر هذه المشكلة بسبب وجود القيم الشاذة.

3.3 بيانات Montgomery, Peck and Vining

هذه البيانات قدمت بواسطة Montgomery, Peck and Vining (2001)⁽¹⁵⁾ و التي تبين العلاقة بين دخل او ايرادات احد المطاعم وقيمة الانفاق على الاعلان، مرة اخرى وضعنا ستة قيم شاذة في

جدول (3-3) بيانات Montgomery, 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

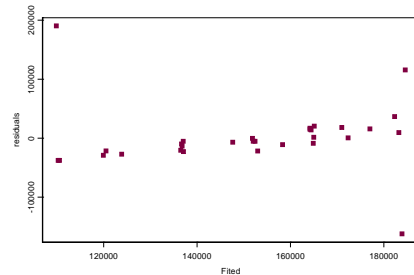
و يبدو عدم تجانس التباين واضحا في البيانات الاصلية الشكل (5-3) و لا يبدو كذلك بظهور القيم الشاذة.

4. مناقشة و تحليل النتائج

4.1 نتائج بيانات Pindyck and

Rubinfeld (1997): الجدول (1-4

) يبين النتائج التي تم الحصول عليها لاقيام اختبار فيشر للتجانس وقوة الاختبار P من خلال تطبيق الاختبار التقليدي Classical GQ و المعدل MGQ ومقترح اختبارنا RMGQ على 20 مشاهدة غير متجانسة لبيانات الخالية من الشواذ مرة و الملوثة بنسبة



الشكل (3-6) البواقي مقابل القيم المقدرة لبيانات Montgomery, Peck and المعدلة Vining

الشكل (5-3) و (3-6) يظهران رسما للبواقي و القيم المقدرة للبيانات الاصلية و المعدلة لبيانات مبيعات اغذية المطعم

5% مرة اخرى عند مستوى معنوية
0.05 .

جدول(4-1)نتائج الاختبارات لبيانات Pindyck and Rubinfeld الاصلية و المعدلة.

Test	بدون قيم شاذة		بوجود القيم الشاذة	
	<i>F</i>	<i>p</i> – value	<i>F</i>	<i>p</i> – value
FC=4.2838				
Classical GQ	12.186	0.0038	4.0928	0.055
MGQ	36.000	0.0002	18	0.0035
RMGQ	<u>56.542</u>	<u>0.0001</u>	<u>56.542</u>	<u>0.0001</u>

بعد تلويث البيانات بقيمة شاذة واحدة شكلت مانسبته 5% تقريبا 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 شخضت المشكلة بنفس درجة شدة عدم التجانس فيما لو كانت البيانات خالية من القيم الشاذة و بنفس الدقة والقوة، بينما MGQ شخضتها بدرجة شدة اقل و السبب يعود في ذلك الى نسبة التشذيب الافتراضية التي اعتمدها هذه الطريقة هي (10%) و هذا يؤدي بالتأكيد الى خسارة مانسبته 5% من

الجدول(4-1) يظهر ان طرق الاختبار الثلاث قد شخضت مشكلة عدم التجانس في البيانات الاصلية الخالية من القيم الشاذة على الرغم من اختلاف درجة شدة التجانس من اختبار الى اخر، اذ ان جميع قيم *F* المحسوبة بالطرق الثلاث تجاوزت القيمة الحرجة لها (4.2838) و التي بدت في الطريقة المقترحة نحو (56.542) و هي الاعلى بين قيم *F* التي تناظر القيمة الاصغر (0.0001) لـ *P* بين نظيراتها في الاختبارات الاخرى. لذلك ترفض الفرضية الصفرية القائلة بتجانس التباين على ضوء النتائج الظاهرة للاختبارات الثلاث في هذه الدراسة.

F للطرق الثلاث قد تجاوزت القيمة الحرجة لها (2.978)، إلا ان الطريقة التقليدية و طريقة MGQ قد اخفقتا في تشخيص المشكلة بقبولها للفرضية الصفرية بعد تلويث البيانات بـ (20%) من القيم الشاذة، في حين الطريقة المقترحة بقت على رفضها للفرضية الصفرية على الرغم من تلويث البيانات. من ناحية اخرى يبين الجدول ان قوة الاختبار للطريقة المقترحة كانت افضل من الطرق الاخرى سواء كانت البيانات ملوثة او غير ملوثة.

البيانات النظيفة بمعنى بتر مشاهدتين من هذه البيانات واحدة شاذة و الاخرى اعتيادية. أما RMGQ لقد شخصت النسبة الفعلية للشواذ و كانت قيمة شاذة واحدة فقط و قامت بتشذيب هذه القيمة مما انعكس على درجة شدة عدم التجانس في هذه العينة المكونة من (20) مشاهدة فقط.

4.2 نتائج بيانات Gujarati

(1995): بين الجدول (2-4) نتائج اختبارات عدم التجانس في هذه الدراسة على بيانات Gujarati (1995) قبل وبعد تلويثها بنسبة (20%) من القيم الشاذة عند مستوى معنوية (0.05). من الواضح ان طرق الاختبار الثلاث شخصت مشكلة عدم التجانس في هذه البيانات قبل تلويثها و ذلك برفضها للفرضية الصفرية بسبب ان قيم اختبار

جدول (2-4) نتائج الاختبارات لبيانات Gujarati الاصلية و المعدلة

Test	بدون قيم شاذة		بوجود القيم الشاذة	
	F	p – value	F	p – value
FC=2.978				
Classical GQ	3.889	0.022	1.370	0.314
MGQ	3.933	0.021	1.042	0.474
RMGQ	<u>15.245</u>	<u>0.0001</u>	<u>3.903</u>	<u>0.021</u>

مانسبته (20%) من القيم الشاذة فشل الاختبار الاول والثاني في تشخيص المشكلة من خلال قبوله للفرضية الصفرية لان قيم فيشر التي اظهرتها لم تتجاوز قيمتها الحرجة عند مستوى معنوية (0.05) و بدرجات حرياتها. لكن الطريقة المقترحة تمكنت من رفض الفرضية الصفرية و تشخيص مشكلة عدم التجانس على الرغم من وجود القيم الشاذة بالاضافة الى ان قيمة P التي كانت اصغر من الاخرى بوجود القيم الشاذة او بدونها، الامر الذ يفسر دقة وقوة تشخيص الطريقة المقترحة.

4.3 بيانات Montgomery, Peck and Vining and Vining: الجدول (3-4) يتضمن نتائج اختبارات الدراسة لتشخيص مشكلة عدم تجانس التباين لبيانات Montgomery, Peck (2001) and Vining بوجود و عدم وجود القيم الشاذة، من النتائج يتضح جميع قيم F للاختبارات الثلاث هي اكبر من قيمتها الحرجة (2.978) عند مستوى معنوية (0.05) بدون قيم شاذة و هذا يعني الطرق قيد الدراسة رفضت الفرضية الصفرية و شخصت وجود مشكلة عدم التجانس في هذه البيانات و لكن بدرجات مختلفة لشدة عدم التجانس. بوجود

جدول (3-4) 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>

5. الاستنتاجات و التوصيات

بينما نسبتها في العينة الاصل (30) مشاهدة) هي (20%). اما بيانات Pindyck and Rubinfeld (1997) شكلت الشواذ في العينة الجزئية الاولى نسبة قدرها (0.125) و الثانية (0.00). ارتفاع نسبة القيم الشاذة الى اكثر من (10%) بعض البيانات هو السبب الاساسي في فشل طريقة MGQ في تشخيص مشكلة عدم التجانس لان طريقة المربعات الصغرى المشدبة تعمل عمل طريقة المربعات الصغرى الاعتيادية بعد تشذيب القيم الشاذة و حيث ان تشذيب مانسبته (10%) في بيانات تتجاوز هذه النسبة يجعلها تتحسس المتبقى من الشواذ في البيانات، وعلى العكس من ذلك عندما كانت نسبة القيم الشاذة في العينة الجزئية الاولى لبيانات Pindyck and Rubinfeld (1997) لاتتجاوز (13%) تمكنت MGQ من تشخيص مشكلة عدم التجانس و لكن ليس بالدقة والقوة الكافية لانها (3%) من اصل (8) مشاهدات لاترقي الى الواحد الصحيح بمعنى (13%) تكافيء (1.04) مشاهدة، لذلك لا يوجد تأثير قوي في بقاء (0.04) من المشاهدة، عندئذ يكون التشذيب لقيمة شاذة واحدة

5.1 الاستنتاجات: من النتائج المبينة في الجداول (4-1)، (4-2)، (4-3)، نستنتج ان طريقة RMGQ المقترحة هي الافضل و الاقوى في تشخيص مشكلة عدم التجانس في البيانات قيد الدراسة بوجود او بعدم وجود القيم الشاذة. الملاحظ ايضا ان نسبة القيم الشاذة الى حجم العينة قد تتعاضد عند تجزئة العينة الى عينتين جزئيتين بعد ترتيب قيم المتغير التوضيحي تصاعديا. اذ لاحظنا عند تلويث بيانات Montgomery, (2001) Peck and Vining (20%) من الشواذ في المتغير التابع و بعد ترتيب قيم المتغير التوضيحي تصاعديا تمركزت ثلاث من القيم الشاذة على راس القيم العظمى و الثلاث الاخرى تذيلت القيم الصغرى، وبعد حذف (6) مشاهدات من وسط الترتيب وتجزئة باقي المشاهدات (24) الى عينتين جزئيتين كل عينة (12) مشاهدة شكلت القيم الشاذة مانسبته (25%) في كل عينة. اما بالنسبة لبيانات Gujarati (1995) شكلت الشواذ في العينة الجزئية الاولى (34%) تقريبا و العينة الجزئية الثانية (17%) تقريبا،

تحسين هذه الاختبارات لمقاومة القيم الشاذة من خلال دراسات جديدة.

المراجع و المصادر

1. الفتلاوي، كاظم كامل، القياس الاقتصادي النظرية و التحليل، دار الصفاء للنشر و التوزيع، 2011
2. العيسوي، ابراهيم، القياس و التنبؤ الاقتصادي، دار النهضة العربية، القاهرة، 1978
3. الزبيدي، حسين لطيف، القياس الاقتصادي و النظرية الاقتصادية، دار الصفا للنشر ،عمان، 2011.
4. كاظم، اموري هادي، الاقتصاد القياسي التطبيقي، دار الحكمة، الموصل، 1995.
5. محبوب، عادل عبد الغني، الاقتصاد القياسي، دار الكتب للطباعة والنشر، الموصل، 1982.
- 6.Chatterjee, S. and Hadi, A.S. (1988). *Sensitivity Analysis in Linear Regression*, New York.: Wiley.
- 7.Cook, R. D. and Weisberg S. (1983). Diagnostics for heteroscedasticity in regression, *Biometrika*, 70, 1 – 10.
- 8.Goldfeld, S.M. and Quandt, R.E. (1965). Some tests for

فقط. أي العودة الى تشذيب مانسبته (10%) الامر الذي دفع طريقة MGQ لتشخيص المشكلة بوجود نسبة من الشواذ تتوافق مع النسبة المقترضة لديها.بينما طريقة RMGQ المقترحة تشخيص مشكلة عدم التجانس بدقة وقوة مهما اختلفت نسبة القيم الشاذة.

5.2التوصيات : بناء على نتائج البحث

- و استنتاجاته يوصي الباحث بما يأتي:
1. تبني مقترح تعديل هذا الاختبار لتشخيص مشكلة عدم التجانس في حالة وجود قيم شاذة في البيانات.
 2. لا يعالج هذا المقترح مشكلة القيم الشاذة في جانب المتغيرات التوضيحية مما يعطي فرصة اخرى للباحثين من دراسة هذه الحالة.
 3. مشكلة عدم التجانس و بوجود القيم الشاذة تظهر ايضا في موضوع السلاسل الزمنية، لذلك يوصي الباحث بدراسة تحسين اختبار كولد فيلد- كوانت لتشخيص مشكلة عدم التجانس في بيانات السلسلة الزمنية.
 4. لقد تناول الباحث اختبار كولد فيلد - كوانت من بين عدة اختبارات اخرى تستخدم لتشخيص مشكلة عدم التجانس، لذلك يوصي الباحث الى

- Linear Regression Analysis, 3rd ed., New York: Wiley.
15. Pindyck, S. R and Rubinfeld, L. D. (1997). *Econometric Models and Econometric Forecasts*, 4th ed., New York: Irwin/McGraw-Hill.
 16. Rousseeuw, P.J. and Leroy, A. (1987). *Robust Regression and Outlier Detection*, New York.: Wiley.
 17. Rana;Md. Sohel, Midi;Habshah, Imon;A.H.M. A Robust Modification of the Goldfield-Quandt Test for Detection of Heteroscedasticity in the Presence of Outliers. *Journal of Mathematics and Statistics* 4 (4): 277-283, 2008.
 18. Ryan, T. P., (1997). *Modern Regression Methods*, New York: Wiley.
 19. White, H. (1980). Heteroscedasticity-consistent covariance matrix estimator and a direct test for heteroscedasticity, *Econometrica*, 48, 817-838.
 - homoskedasticity, *Journal of the American Statistical Association*, 60, 539 – 547.
 9. Gujarati, D. (1995). *Basic Econometrics*, 4rd ed. New York: McGraw-Hill.
 10. Hampel, F.R. 1971. A general Qualitative Definition of robustness, *The Annals Mathematical of Statistics.*, 42,1887-1896.
 11. Imon, A. H. M. R. (2002). On deletion residuals, *Calcutta Statistical Association Bulletin*, 52, 65–79.
 12. Imon, A.H.M.R. (2005). Identifying multiple influential observations in linear regression, *Journal of Applied Statistics*, 32, 929 – 946.
 13. Imon, A.H.M.R. (2008). Deletion residuals in the detection of heterogeneity of variances in linear regression, (Accepted for publication), *Journal of Applied Statistics*.
 14. Montgomery, D.C., Peck, E.A. and Vining, G.G. (2001). *Introduction to*