استعمال طريقتي التقدير الحصينة المربعات الصغرى المشذبة و M لتقدير معلمات أنموذج الانحدار الخطي باخطاء عشوائية غير متجانسة التباين

أ.م.د. احمد شاكر محمد طاهر
الجامعة المستنصرية/ كلية الادارة والاقتصاد

الباحث رضا قاسم محمد تميم
الجامعة المستنصرية/ كلية الادارة والاقتصاد

تاريخ استلام البحث: 13 / 04 / 2022

تاريخ قبول البحوث: 10 / 05 / 2022

نشر البحث في العدد السابع عشر: ديسمبر / كانون اول 2022

رمز التصنيف ديوي / النسخة الالكترونية (Online): 2522-64X/519.5

رمز التصنيف ديوي / النسخة الورقية (Print): 2519-948X/519.5

استعمال طريقتي التقدير الحصينة المربعات الصغرى المشذبة و M لتقدير معلمات أنموذج الانحدار الخطي باخطاء عشوائية غير متجانسة التباين

أ.م.د. احمد شاكر محمد طاهر
الجامعة المستنصرية/ كلية الادارة والاقتصاد

الباحث رضا قاسم محمد تميم
الجامعة المستنصرية/ كلية الادارة والاقتصاد

المستخلص

تعتبر مشكلة عدم تجانس تباين حدود الخطأ العشوائي احدى اهم واكثر المشاكل القياسية التي تعاني منها اغلب البيانات الاقتصادية والمالية. ان الاعتماد على مقدرات المربعات الصغرى الاعتيادية (OLS) في هذه الحالة تؤدي الى نتائج مظللة، فبالرغم من بقاء تلك المقدرات غير متحيزة لكنها لا تملك اقل تباين. لذا يجب البحث عن طرائق بديلة تكون حصينة ضد مشكلة عدم تجانس تباين حدود الخطأ العشوائي. في هذا البحث تم استعمال طريقتين حصينتين, الاولى هي طريقة المربعات الصغرى المشذبة (TLS), والثانية طريقة (M), ومن ثم اجراء مقارنة بين تلك الطريقتين من خلال التطبيق على بيانات واقعية تمثل الكميات المنتجة والمصدرة من النفط الخام في العراق. افرزت نتائج التقدير لكلا الطريقتين الى افضلية طريقة المربعات الصغرى المشذبة بالاعتماد على معيار المفاضلة متوسط مربعات الخطأ (MSE).

Abstract:

The problem of heterogeneity of the random error terms variance for the regression model is one of the most important standard problems that most economic and financial data suffer from. The reliance on Ordinary Least Squares (OLS) estimators in this case lead to shaded results, although these estimators remain unbiased, but they do not have the least variance. Therefore, it is necessary to search for alternative methods that are robust against to the problem of heterogeneity of the variance of the random error terms. In this research, two robust methods were used, the first is the Trimmed Least Squares (TLS) method, and the second is the (M) method.

Then a comparison was made between these two methods, through the application on realistic data representing the quantities produced and exported of crude oil in Iraq. The estimation results for both methods resulted in the preference of the trimmed least squares method, based on the mean of squared error (MSE) criterion.

الكلمات الافتتاحية: الانحدار الخطي البسيط، عدم تجانس التباين، طريقة المربعات الصغرى المشذبة، طريقة M.

1. المبحث الاول: منهجية البحث

1-1 المقدمة

تعد طريقة المربعات الصغرى الاعتيادية (OLS) من طرائق التقدير المفضلة لتقدير معاملات أنموذج الانحدار لبساطة فكرتها ولان تقديراتها تتصف بانها افضل تقدير خطي غير متحيز (Best Linear Unbiased Estimator (BLUE)) وذلك في حالة تحقق الفرضيات الاساسية لأنموذج الانحدار ومنها تجانس تباين حدود الخطأ العشوائي. ولكن في الواقع العملي قد يحدث انتهاك لهذه الفرضية إذ يكون تباين حدود الخطأ العشوائي غير متجانس وهذا ما يعرف بمشكلة عدم تجانس التباين (Heteroscedasticity Problem)، والتي تعتبر واحدة من اهم واكثر المشاكل القياسية. وفي هذه الحالة لا يكون من الجيد الاعتماد على طريقة المربعات الصغرى الاعتيادية، فبالرغم من بقاء مقدراتها غير متحيزة لكنها لا تملك اقل تباين. لذا يجب البحث عن طرائق تقدير بديلة ذات كفاءة عالية تكون حصينة لمشكلة عدم تجانس التباين, من تلك الطرائق الحصينة طريقة المربعات الصغرى المشذبة وطريقة (M) اللتان تم استعمالهما لتقدير معاملات أنموذج الانحدار باخطاء عشوائية غير متجانسة التباين. لغرض اجراء مقارنة بين هاتين الطريقتين تم الاعتماد على بيانات واقعية التي تمثل الكميات المنتجة والمصدرة للنفظ الخام في العراق. افرزت نتائج تقدير انموذج الانحدار الخطي البسيط الذي يصف العلاقة بين الكميات المصدرة من النفط الخام والكميات المنتجة منه الى افضلية طريقة المربعات الصغرى المشذبة وذلك بالاعتماد على معيار المفاضلة متوسط مربعات الاخطاء.

تضمن البحث اربع مباحث, الاول تضمن على المقدمة ومشكلة البحث وهدف البحث، والثاني فتضمن الجانب النظري للبحث حيث تم عرض مشكلة عدم تجانس التباين وطريقة المربعات الصغرى المشذبة وطريقة (M)، و الثالث تضمن التطبيق العملي لطريقتي التقدير موضوع البحث، وختم البحث بالمبحث الرابع الذي شمل على ما تم التوصل اليه من الاستنتاجات والتوصيات.

2-1 مشكلة البحث

لغرض التوصل الى تقديرات لمعاملات أنموذج الانحدار تتصف بالخصائص المرغوبة وذات كفاءة عالية يفترض تحقق جميع افتراضات التحليل الخاصة بذلك الأنموذج. غير إن ذلك قد يكون غير متحقق في التطبيق العملي إذ قد تحصل خروق لبعض تلك الافتراضات او جميعها كعدم تحقق فرضية تجانس تباين حدود الخطأ العشوائي، مما يؤدي الى ان تقديرات المربعات الصغرى الاعتيادية تكون غير كفوءة وتؤدي الى نتائج مظللة، الامر الذي يدعو الى البحث عن طرائق تقدير بديلة تمكننا من الحصول على تقديرات تتصف بالخصائص المرغوبة وذات كفاءة عالية.

3-1 هدف البحث

يهدف البحث الى استعمال طريقتي التقدير الحصينة المربعات الصغرى المشذبة و (M) كبديل لطريقة المربعات الصغرى الاعتيادية لتقدير معلمات أنموذج الانحدار الخطي البسيط بوجود مشكلة عدم تجانس التباين لحدود الخطأ العشوائي، واجراء المقارنة بينهما بالاعتماد على بيانات حقيقية بحجوم عينات مختلفة.

2. المبحث الثاني الجانب النظري

يتضمن هذا المبحث بعض المفاهيم النظرية المتعلقة بأنموذج الانحدار الخطي البسيط ومشكلة عدم تجانس التباين وبعض طرائق الكشف عن وجود هذه المشكلة علاوة عن تقدير معلمات أنموذج الانحدار بالطريقتين الحصينتين المربعات الصغرى المشذبة و (M).

2-1 أنموذج الانحدار الخطي البسيط Simple Linear Regression

إن دراسة العلاقة بين المتغيرات يتطلب تحديد المتغيرات المؤثرة في تلك العلاقة ومن ابسط وأفضل أنواع العلاقات في التقدير والتحليل الاحصائي هي العلاقة بين متغيرين أحدهما يسمى بمتغير الاستجابة والأخر هو المتغير التوضيحي, وتسمى العلاقة الخطية بين هذين المتغيرين بالانحدار الخطي البسيط simple linear regression , الصيغة الرياضية التي تأخذه هذه العلاقة تسمى أنموذج الانحدار الخطي البسيط، ويمكن كتابتها بالصيغة الآتية,[2],[6],[11]:

$$Y_i = \alpha + \beta X_i + \epsilon_i \qquad , \ i = 1,2, \dots, n \qquad \dots (1)$$

اذ إن, Y_i يمثل متغير الاستجابة, X_i يمثل المتغير التوضيحي, ϵ_i تمثل الأخطاء العشوائية للأنموذج, α, β هي معاملي الانموذج , n تمثل حجم العينة.

يستند أنموذج الانحدار الخطي البسيط على عدة افتراضات لغرض الحصول على تقديرات كفؤة وامكانية اجراء اختبارات المعنوية وبناء حدود الثقة لمعاملات الانموذج والقيم التنبؤية لمتغير الاستجابة وهذه الافتراضات هي:

1- متغير الاستجابة (Y) متغير عشوائي يتبع التوزيع الطبيعي بوسط حسابي (μ) وتباين (σ^2)، وقيمه مستقلة إحصائيا الواحدة عن الأخرى.
2- العلاقة بين متغير الاستجابة والمتغير التوضحيحي هي علاقة خطية.
3- قيم المتغير التوضيحي (X) ثابتة للعينات المتكررة وغير مرتبطة بحدود الخطأ العشوائي، أي ان $E(\epsilon_i X_i) = 0$.
4- الاخطاء العشوائية (ϵ_i) هي متغير عشوائي ذات قيم مستقلة وتتبع التوزيع الطبيعي بوسط حسابي يساوي صفر وتباين ثابت يساوي (σ^2) أي ان $\epsilon_i \sim IN(0, \sigma^2)$. وهذه الحالة تسمى بتجانس تباين حدود الخطأ العشوائي (Homogeneity).

2-2 مشكلة عدم تجانس تباين الاخطاء العشوائية Heteroskedastic

دقة المعلمات المقدرة باستخدام طريقة المربعات الصغرى الاعتيادية يعتمد اساسا على تحقق الفرضيات الخاصة بأنموذج الانحدار، وان عدم تحقق كل او بعض من هذه الفرضيات يؤثر على نتائج التقدير التي لا يمكن الاعتماد عليها في هذه الحالة لغرض اختبار الفرضيات وبناء فترات الثقة واجراء عملية التنبؤ،[6][2]، وان احدى الفرضيات التي اعتمدت في تقدير معاملات انموذج الانحدار هي فرضية تجانس التباين للأخطاء العشوائية. ولكن في الواقع العملي لا توجد ضمانات لتحقق هذه الفرضية، مما يؤدي الى حدوث مشكلة عدم تجانس تباين الاخطاء العشوائية (Heteroskedastic). هذه الفرضية قد لا تمثل مشكلة بالنسبة للنماذج التي تعتمد على بيانات السلاسل الزمنية لان قيم المتغير التوضيحي تكون عادةً ذات رتبة مماثلة في الحجم عند جميع نقاط المشاهدة، والشي نفسه ينطبق على قيم المتغير المعتمد. فعلى سبيل المثال لدالة الاستهلاك يكون مستوى الاستهلاك في السنوات الاخيرة ذو رتبة مماثلة لمستوى الاستهلاك قبل عشرون سنة، والشيء نفسه يكون في الدخل. فما لم تكن هناك بعض الظروف

الخاصة فإن فرضية تجانس التباين في هذه النماذج يبدو امرا مقبولا. ولكن عندما نتعامل مع بيانات الاقتصاد الجزئي فقد تحتوي على مشاهدات مختلفة جوهريا في الحجم. مثال على ذلك بيانات الدخل والاستهلاك لعدة عوائل, ففرضية تجانس التباين هنا ليس امرا مقبولا جدا على اسس مسبقة نظرا لاننا نتوقع تباينا اقل في مستوى الاستهلاك للعوائل ذات الدخل المنخفض مقارنة بالعوائل ذات الدخل المرتفع. ففي هذه الحالة والحالات المشابهة لها فإن النموذج المناسب لها هو عدم تجانس تباين حدود الخطأ العشوائي. وان فرضية عدم تجانس تباين الاخطاء العشوائية تكون كالاتي:

$$E(\epsilon'\epsilon) = \sigma_i^2 = \begin{bmatrix} \sigma_1^2 & 0 & \dots & 0 \\ 0 & \sigma_2^2 & \dots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \dots & \sigma_n^2 \end{bmatrix} \qquad \dots (2)$$

وتحت صحة الفرضية اعلاه فان مقدرات المربعات الصغرى الاعتيادية تبقى غير متحيزة لكنها لا تمتلك اقل تباين ممكن، اي انها لن تكون افضل تقدير خطي غير متحيز (BLUE).
ان حدوث مشكلة عدم تجانس تباين حدود الخطأ العشوائي يؤدي الى نتائج تقدير مضللة وذلك بسبب كون تباين معاملات الانحدار المقدر يكون تقديرا متحيزا للتباين الحقيقي لتلك المقدرات مما يؤثر على اختبار المعنوية وبناء حدود الثقة وكذلك عمل التنبؤات المستقبلية. وهذا يعني اذا قمنا بتحليل الانحدار في ظل الاعتقاد الخاطئ –أي انه لا توجد مشكلة عدم تجانس التباين- فإن الاستدلال حول معاملات المجتمع يكون غير صحيح، كذلك قد تكون حدود الثقة او منطقة القبول اضيق و أوسع مما هي عليه فعلا. لذا يجب اختبار وجود هذه المشكلة من عدمها عند تقدير معاملات أنموذج الانحدار, هناك العديد من الاختبارات المعتمدة في الكشف عن هذه المشكلة والتي سنتطرق الى قسما من تلك الاختبارات.

2-4 الاختبارات المعتمدة في الكشف عن مشكلة عدم تجانس التباين

هناك عدد كثير من الاختبارات التي تعتمد في الكشف عن مشكلة عدم تجانس تباين حدود الخطأ العشوائي نذكر منها الاتي:

1- اختبار بارك (Park test)
2- اختبار غولدفيلد-كوانت (Goldfeld-Quandt test)
3- اختبار كوينكر-باسيت (Koenker-Bassett test)
4- اختبار ارتباط الرتب لسبيرمان (Spearman Rank Correlation)
5- اختبار بروش- باغان (Breusch-Pagan-Godfrey test)
6- اختبار وايت العام (White's general test)
7- اختبار غليجسر (Glejser test)

وسنتطرق بشيءٍ من التفصيل لكل من اختبار بروش-باغان واختبار وايت العام.

2-4-1 اختبار بروش- باغان

وضع هذا الاختبار من قبل الباحثين (Breusch و Pagan) عام 1979, [10], [12], وتتلخص فكرة هذا الاختبار بافتراض وجود أنموذج الانحدار الخطي العام الاتي:

$$y_i = \alpha + \beta_1 X_{1i} + \cdots\cdots + \beta_n X_{ni} + \epsilon_i \qquad \ldots (3)$$

بافتراض ان تباين الاخطاء العشوائية (σ_i^2) يمكن وصفها بشكل دالة خطية بمجموعة من المتغيرات غير العشوائية ولتكن(Zj) اذ ان (j=1,2,...,m) وكالاتي:

$$\sigma_i^2 = \delta_0 + \delta_1 Z_{1i} + \cdots\cdots + \delta_m Z_{mi} \qquad \ldots (4)$$

المتغيرات العشوائية (Zj) قد تكون جميع او بعض من المتغيرات التوضيحية. فإذا كانت $(\delta_1 = \delta_2 = \cdots = \delta_m = 0)$ فإن (σ_i^2) سوف يساوي القيمة الثابتة (δ_0) بمعنى ان تباين الاخطاء العشوائية يكون ثابتا، على هذا الاساس يمكن اختبار كون تباين الاخطاء العشوائية ثابتة من عدمها باختبار الفرضية الاتية:

$$H_0: \delta_1 = \delta_2 = \cdots\cdots = \delta_m = 0$$

$$H_1: \delta_1 \neq \delta_2 \neq \ \ldots\ldots \neq \delta_m \neq 0$$

وتتلخص خطوات الاختبار بالاتي:

يتم تقدير معاملات انموذج الانحدار الخطي المتعدد المبين بالعلاقة (3) باستخدام طريقة المربعات الصغرى الاعتيادية (OLS) ثم حساب قيم البواقي المقدرة (e_i) .

حساب القيمة التقديرية لتباين الاخطاء بالاعتماد على قيم البواقي المقدرة وفق الصيغة التالية والتي تمثل الصيغة التقديرية وفق طريقة الامكان الاعظم:

$$\tilde{\sigma}^2 = \frac{\sum_{i=1}^{n} \hat{e}_i^2}{n} \qquad \ldots (5)$$

1. حساب قيم المتغير (P_i) الذي يمثل نسبة مربع قيم البواقي المقدرة الى تباين الاخطاء المقدر بالخطوة الثانية، وفق الصيغة الاتية:

$$P_i = \frac{e_i^2}{\tilde{\sigma}^2} \qquad \ldots (6)$$

2. بناء انموذج انحدار باعتبار المتغير (P_i) متغير استجابة على (m) من المتغيرات التوضيحية التي تمثل المتغير (Zj) والتي تم تعريفها سابقا وكالاتي:

$$P_i = \delta_0 + \delta_1 Z_{1i} + \cdots\cdots + \delta_m Z_{mi} + v_i \qquad \ldots (7)$$

وان (v_i) حد الخطأ العشوائي للأنموذج (7)

3. حساب مجموع مربعات الانحدار (SSR) للأنموذج (7)، ثم حساب المقدار (Θ) المعرف بالاتي:

$$\Theta = \frac{1}{2} SSR \qquad \ldots (8)$$

وعلى افتراض إن الأخطاء العشوائية تتبع التوزيع الطبيعي واذا كانت الاخطاء العشوائية تتصف بتجانس التباين ومع ازدياد حجم العينة فإن التوزيع الاحتمالي التقريبي للمقدار (Θ) هو توزيع مربع كاي (χ^2) بدرجة حرية $(m-1)$ اي ان

$$\Theta \underset{asy}{\sim} \chi^2_{(m-1)}$$

لذا يمكن رفض فرضية العدم إذا كانت القيمة المحسوبة (Θ) اكبر من القيمة الجدولية لـ (χ^2) بمعنى ان تباين حدود الخطأ العشوائي تتصف بعدم التجانس,[11].

2-4-2 اختبار وايت العام

اقترح هذا الاختبار من قبل الباحث (White) عام (1980)،[11], ويعد هذا الاختبار من الاختبارات المهمة للكشف عن مشكلة عدم تجانس التباين. والذي لا يعتمد على افتراض ان الاخطاء العشوائية تتبع التوزيع الطبيعي. لتوضيح فكرة هذا الاختبار نفترض توفر أنموذج انحدار خطي متعدد بمتغيرين توضيحيين وفق الصيغة التالية، ويمكن تعميم ذلك لانموذج انحدار متعدد بعدد (k) من المتغيرات التوضيحية:

$$y_i = \alpha + \beta_1 x_{i1} + \beta_2 x_{i2} + \epsilon_i \qquad \dots (9)$$

حيث يتم اجراء الاختبار وفق الخطوات الاتية:

1- حساب البواقي (e_i) لانموذج الانحدار في المعادلة (9) بطريقة المربعات الصغرى الاعتيادية (OLS).

2- صياغة أنموذج انحدار مساعد بمتغير استجابة يمثل مربعات البواقي المقدرة والمتغيرات التوضيحية لهذا الانموذج عبارة عن المتغيرات التوضيحية لانموذج الانحدار الاصلي (9) ومربعاتها وحاصل ضربهما، وكما مبين بالصيغة الاتية:

$$e_i^2 = \gamma_0 + \gamma_1 x_{i1} + \gamma_2 x_{i2} + \gamma_3 x_{i1}^2 + \gamma_4 x_{i2}^2 + \gamma_5 x_{i1} x_{i2} + v_i \quad \dots (10)$$

انموذج الانحدار وفق الصيغة (10) يمكن صياغته في حالة انموذج الانحدار الخطي البسيط بالشكل الاتي:

$$e_i^2 = \gamma_0 + \gamma_1 X_i + \gamma_2 X_i^2 + v_i \qquad \dots (11)$$

ومن الجدير بالملاحظة ان أنموذج الانحدار المساعد يتضمن على الحد الثابت (γ_0)، بغض النظر عما إذا كان الانموذج الأصلي يتضمن معلمة الحد الثابت او لا.

3- الفرضية المراد اختبارها يمكن صياغتها بالشكل الاتي:

تجانس تباين حدود الخطأ العشوائي :H_0

عدم تجانس تباين حدود الخطأ العشوائي :H_1

4- يتم حساب معامل التحديد (R^2) لأنموذج الانحدار المساعد، تحت صحة فرضية العدم فان حاصل ضرب حجم العينة (n) في معامل التحديد (R^2) يتبع بشكل تقاربي توزيع مربع كاي بدرجة حرية مساوية لعدد المتغيرات التوضيحية في انموذج الانحدار المساعد أي أن:

$$n.R^2 \sim \chi^2_{(k)} \qquad \dots (12)$$

بمقارنة القيمة المحسوبة لاحصاءة الاختبار ($n.R^2$) مع القيمة الجدولية لتوزيع مربع كاي (χ^2) عند درجة حرية (k)، فاذا كانت القيمة المحسوبة اكبر من الجدولية يتم رفض فرضية العدم, [4], [11].

5-2 طريقة المربعات الصغرى المشذبة Trimmed Least Squares Method (TLS)

تم اقتراح هذه الطريقة من قبل الباحث (Rousseeuw) في عام (1984) والتي تعد من طرائق التقدير الحصينة لتقدير معلمات أنموذج الانحدار وتتميز هذه الطريقة بكفاءة افضل من طريقة المربعات الصغرى الاعتيادية (OLS) في حال عدم تجانس تباين الخطأ العشوائي وكذلك عند وجود القيم الشاذة،[9][3]، وتتلخص هذه الطريقة بالخطوات الاتية:

1- ايجاد مربعات البواقي (e_i^2) لأنموذج الانحدار الخطي البسيط بطريقة المربعات الصغرى الاعتيادية (OLS).
2- يتم ترتيب تلك البواقي تصاعديا $e_1^2 \le e_2^2 \le \cdots\cdots \le e_n^2$.
3- ايجاد حجم المجموعة الجزئية (k) بالاعتماد على نسبة التشذيب (λ) ، وفق الصيغة الاتية, [7], [13]:

$$k = n(1-\lambda) + \lambda(p+1) \qquad 13)$$

اذ أن n يمثل حجم العينة, p يمثل عدد المتغيرات التوضيحية.

وفي حال كانت نسبة التشذيب (λ) مساوية الى (50%) فان حجم المجموعة الجزئية (k) يحسب وفق الصيغة التالية حيث تبلغ نقطة الانهيار اقصى حد لها عند هذه النسبة:

$$k = \frac{n+p+1}{2} \qquad \dots (14)$$

4- يتم حساب مجموع مربعات البواقي بعد استبعاد مربعات البواقي الكبيرة من المجموع للحصول على تقدير لمعاملات أنموذج الانحدار من خلال تصغير دالة الهدف الاتية:

$$\hat{\beta}_{LTS} = \arg min_{\beta} \sum_{i=1}^{k} e_i^2 \qquad \dots (15)$$

ان طريقة عمل (TLS) تشبه طريقة عمل (OLS) لكن الاختلاف يتمثل بإن طريقة (TLS) تقوم باستبعاد تأثير القيم ذات البواقي الكبيرة قبل البدء بالتقدير. طريقة (TLS) يمكن ان تكون كفوءة جدا اذا تم تشذيب البيانات بطريقة صحيحة, حيث تكون مقدرات هذه الطريقة مكافئة حسابيا لطريقة (OLS)، وعلى العكس من ذلك اذا كان التشذيب اكبر من المطلوب فقد يتم استبعاد بعض البيانات الجيدة من الحساب وهذا قد يؤثر على كفاءة مقدرات هذه الطريقة.

6-2 طريقة التقدير M Estimation Method (M)

قدم الباحث (Huber) عام (1964), [12], مجموعة من المقدرات تسمى مقدرات الإمكان الأعظم النوعية (Maximum Likelihood Type Estimator) ويرمز لها بالمختصر (M)، والتي تعتمد فكرتها على تقدير معاملات انموذج الانحدار من خلال تصغير دالة في البواقي بدلا من تصغير دالة البواقي نفسها اي ان دالة الهدف المعتمدة تكون وفق الصيغة الاتية:

$$\hat{\beta}_M = \arg min_\beta \sum_{i=1}^{n} \rho(u_i) \quad \dots (16)$$

إذ ان $[\rho(u_i)]$ هي دالة بالبواقي تتصف بانها متماثلة مع الحد الادنى عند الصفر. تعتمد طريقة التقدير (M) على طريقة المربعات الصغرى معادة الوزن التكرارية لتقدير معاملات أنموذج الانحدار. وبحل دالة الهدف (16) بالتعويض عن الاخطاء العشوائية (u_i) بالبواقي القياسية, [5], [8], [12]:

$$\hat{\beta}_M = \arg min_\beta \sum_{i=1}^{n} \rho\left(\frac{e_i}{\hat{\sigma}_{MAD}}\right) =$$
$$\arg min_\beta \sum_{i=1}^{n} \rho\left(\frac{Y_i - \sum_{j=0}^{k} x'_{ij}\hat{\beta}_j}{\hat{\sigma}_{MAD}}\right) \quad \dots (17)$$

إذ ان $(\hat{\sigma}_{MAD})$ تمثل وسيط مطلق انحرافات البواقي (e_i) عن وسيطها وتحسب وفق الصيغة الاتية:

$$\hat{\sigma}_{MAD} = \frac{median\ |e_i - median(e_i)|}{0.6745} \quad \dots (18)$$

ان $(\hat{\sigma}_{MAD})$ هو غالبا ما يكون تقدير القياس من التركيب الخطي للبواقي، والثابت (0.6745) يجعل (S) تقدير غير متحيز تقاربيا لـ(σ) اذا كان حجم العينة كبيرا والبيانات تتبع التوزيع الطبيعي. الدالة (ρ) تأخذ عدة صيغ والجدول (2-1) يبين بعض من تلك الصيغ، وباشتقاق دالة الهدف في المعادلة (17) ومساواتها للصفر نحصل على:

$$\sum_{i=1}^{n} x_{ij}\psi\left(\frac{y_i - \sum_{j=0}^{k} x_{ij}\beta_j}{\hat{\sigma}_{MAD}}\right) = 0 \quad , \quad j = 0,1, \dots, k \quad \dots (19)$$

وإن (ψ) هي المشتقة الاولى للدالة (ρ)، (x_{ij}) هي المشاهدة (i) للمتغير التوضيحي (j) وان $(x_{i0} = 1)$ لتضمين الحد الثابت. وتحل المعادلة (19) بالاعتماد على الاوزان المعرفة بالاتي, [1], [12]:

جدول (1) : بعض صيغ الدالة (ρ)

Method	Objective Functions	Weight Function	قيمة aالثابت
Huber	$\begin{cases} \frac{1}{2}e^2 & , \quad if\ \lvert e\rvert \le a \\ a\lvert e\rvert - \frac{1}{2}a^2 & , \quad if\ \lvert e\rvert > a \end{cases}$	$\begin{cases} 1 & , \quad if\ \lvert e\rvert \le a \\ \frac{a}{\lvert e\rvert} & , \quad if\ \lvert e\rvert > a \end{cases}$	1.345
Tukey's Bisquare	$\begin{cases} \frac{a^2}{6}\left(1 - \left(\frac{e}{a}\right)^2\right)^3 & , \quad \lvert e\rvert \le a \\ \frac{1}{6}a^2 & , \quad \lvert e\rvert > a \end{cases}$	$\begin{cases} \left(1 - \left(\frac{e}{a}\right)^2\right)^2 & , \quad \lvert e\rvert \le a \\ 0 & , \quad \lvert e\rvert > a \end{cases}$	4.685

ونستطيع تقدير الاوزان وفق الصيغة الاتية:

$$w_i = w(e_i) = \frac{\psi\left(\frac{Y_i - \sum_{j=1}^{k} x'_{ij}\hat{\beta}}{\hat{\sigma}_{MAD}}\right)}{\left(\frac{Y_i - \sum_{j=1}^{k} x'_{ij}\hat{\beta}}{\hat{\sigma}_{MAD}}\right)} \quad \dots (20)$$

تصبح بالشكل الاتي:*(19)*وبهذا فإن صيغة المعادلة

$$\sum_{i=1}^{n} x_{ij} w_i \left(y_i - \sum_{j=0}^{k} x_{ij}\hat{\beta}\right) = 0 \ , \ j = 0,1, \dots, k \quad \dots (21)$$

وباستعمال طريقة *(IRLS)* يمكن حل المعادلة *(21)*. بالاعتماد على قيم معاملات الانحدار الابتدائية $\left(\hat{\beta}^0\right)$، وبإعادة كتابة المعادلة *(21)* باستخدام المصفوفات لنحصل على معادلة المربعات الصغرى الموزونة الاتية:

$$X'WX\hat{\beta} = X'WY \quad \dots (22)$$

وبحل المعادلة (22) نحصل على تقدير معاملات الانحدار وفق طريقة (M) وكالاتي:

وبحل المعادلة (22) نحصل على تقدير معاملات الانحدار وفق طريقة (M) وكالاتي: $\hat{\beta}_M =$

$$(X'WX)^{-1}X'WY \quad \dots (23)$$

و(W) هي مصفوفة قطرية ذات سعة $(n * n)$ وعناصر قطرها الرئيسي تمثل الاوزان $w(e_i)$. وتتلخص خطوات طريقة (M) بالاتي:

1- نقوم بتقدير معاملات الانحدار باستعمال طريقة المربعات الصغرى الاعتيادية لنحصل على تقدير اولي للمعلمات ونرمز لها بـ$\left(\hat{\beta}^0\right)$.
2- حساب قيم البواقي $\left(e_i = Y_i - \hat{Y}_i\right)$.
3- حساب قيم الاوزان $\left(w(e_i)\right)$.
4- نحسب $\left(\hat{\beta}_M\right)$ باستعمال المربعات الصغرى الموزونة (WLS) باستعمال الاوزان $\left(w(e_i)\right)$ ووفق المعادلة (23).
5- تكرار الخطوات(4-2) حتى نحصل على قيم مقدرات متقاربة جدا.

إن طريقة تقدير(M) ذات كفاءة عالية، إلا إن هذه الطريقة تكون حساسة بوجود القيم الشاذة في المتغيرات التوضيحية وفي هذه الحالة فإن نقطة الانهيار لها تساوي $\left(\frac{1}{n}\right)$.

3. المبحث الثالث:

3-1 الجانب التطبيقي

يتضمن هذا المبحث اجراء تطبيق عملي على بيانات حقيقية خاصة بالكميات المنتجة والكميات المصدرة من النفط الخام في العراق، إذ سيتم الاعتماد على ثلاثة حجوم عينات مختلفة لتقدير معاملات انموذج الانحدار الذي يصف العلاقة بين الكميات المصدرة للنفط كمتغير استجابة والكميات المنتجة منه كمتغير توضيحي, باستعمال طريقة المربعات الصغرى المشذبة باربعة

نسب للتشذيب وطريقة تقدير (M) بالاعتماد على دالتي (Tukey's Bisquare و Huber). ومن ثم المقارنة بين طريقتي التقدير لإيجاد افضل طريقة بينهما باستخدام معيار المقارنة متوسط مربعات الخطأ (MSE).

2-3 وصف عينات البحث

يمثل التطبيق العملي لهذا البحث تحليل العلاقة بين الكميات المنتجة من النفط الخام متمثلا بالمتغير (P) والكميات المصدرة من النفط الخام متمثلا بالمتغير (E)، وقد تم تسجيل القيم الشهرية لكلا المتغيرين. إذ سيتم استعمال ثلاثة حجوم عينات هي $(n_1 = 72,\ n_2 = 120, n_3 = 144)$. البيانات تم الحصول عليها من المجموعة الاحصائية السنوية الصادرة من الجهاز المركزي للاحصاء/ وزارة التخطيط / جمهورية العراق.

ولغرض تحليل هذه العلاقة تم نمذجتها بأنموذج الانحدار الخطي البسيط وكالاتي:

$$E_t = \alpha + \beta P_t + \epsilon_t \quad \ldots (24)$$

إذ إن E_t الكميات المصدرة من النفط الخام ويمثل متغير الاستجابة, P_t الكميات المنتجة من النفط الخام وهو المتغير التوضيح, ϵ_t يمثل حد الخطأ العشوائي، بافتراض ان تباينات حد الخطأ العشوائي تكون غير متجانسة, α و β معاملات أنموذج الانحدار المطلوب تقديرها.

3-3 الاختبارات المعتمدة في الكشف عن مشكلة عدم تجانس التباين

قبل اجراء عملية تقدير معاملات أنموذج الانحدار الذي يمثل العلاقة بين متغيري التطبيق موضوع البحث تم اجراء اختبار فرضية عدم تجانس تباين حدود الخطأ العشوائي لأنموذج الانحدار باستعمال الاختبارين الاحصائيين (بروش-باغان ووايت العام). إذ تم اختبار فرضية عدم تجانس التباين لكلا الاختبارين وكانت النتائج كما موضحة في الجدول (2) ولحجوم العينات الثلاثة.

جدول(2): نتائج اختبارات الكشف عن مشكلة عدم تجانس التباين لحدود الخطأ العشوائي.

حجم العينة	Breusch-Pagan		White	
	Θ	P-value	χ^2	P-value
$n_1 = 72$	9	0.0027	11.15	0.0038
$n_2 = 120$	43.24	0.0000	49.77	0.0000
$n_3 = 144$	63.26	0.0000	72.48	0.0000

من الجدول اعلاه وبملاحظة قيم احصاءة الاختبار لكل من اختبار (بروش-باغان و وايت العام) إذ كانت ذات قيمة معنوية مما يشير الى عدم تجانس تباين حدود الخطأ العشوائي ولحجوم العينات الثلاث. وبالنظر لعدم تحقق فرضية تجانس التباين تم الاعتماد على طريقتي التقديرالحصينتين موضوع البحث اللتان تم توضيحهما في الجانب النظري لتقدير معاملات أنموذج الانحدار.

4-3 التقدير بطريقة المربعات الصغرى المشذبة

باستعمال طريقة المربعات الصغرى المشذبة تم تقدير معاملات أنموذج الانحدار الخطي البسيط، إذ تم الاعتماد على اربعة نسب للتشذيب $(\lambda_1 = 5\%, \lambda_2 = 10\%, \lambda_3 = 20\%, \lambda_4 = 50\%)$. والجدول (3) يعرض نتائج التقدير لحجوم العينات الثلاثة. اشارت

النتائج الى معنوية تأثير الكميات المنتجة من النفط الخام لجميع حجوم العينات، وان نسبة التشذيب (20%) عند حجم عينة (72) كانت هي الافضل إذ حققت اقل قيمة لمعيار المفاضلة متوسط مربعات الخطأ (MSE) وكانت قيمتها تساوي (8.95) وكانت قيمة المؤشر الاحصائي (R^2) لها تساوي (0.958). اما عند حجم عينة (120) فكانت نسبة التشذيب (50%) هي المثلى لتحقيقها اقل قيمة لمتوسط مربعات الخطأ وتساوي (1.59) وبقدرة تفسيرية لأنموذج الانحدار المقدر تساوي (0.991). وفيما يخص العينة الثالثة ذات حجم (144) فقد كانت نسبة التشذيب المثلى (50%) قد حققت قيمة لمعامل التحديد تساوي (0.994) كما كانت قيمة معيار المفاضلة متوسط مربعات الخطأ تساوي (1.21) مما يشير الى افضلية هذه النسبة.

جدول (3): نتائج تقدير طريقة التقدير الحصينة المربعات الصغرى المشذبة

نسبة التشذيب	المعلمات	تقدير المعلمات	S.E.	t-value	P-value	R^2	MSE
$n_1 = 72$							
$\lambda_1 = 5\%$	α	1.60	5.15	0.31	0.7566	0.842	23.12
	β	0.78	0.04	18.78	0.0000		
$\lambda_2 = 10\%$	α	0.76	4.74	0.16	0.8726	0.871	17.46
	β	0.79	0.04	20.61	0.0000		
$\lambda_3 = 20\%$	α	1.10	3.92	0.28	0.7801	0.958	8.95
	β	0.78	0.03	24.44	0.0000		
$\lambda_4 = 50\%$	α	0.84	3.36	0.25	0.8044	0.914	10.60
	β	0.79	0.03	28.07	0.0000		
$n_2 = 120$							
$\lambda_1 = 5\%$	α	1.81	1.51	1.20	0.2325	0.964	12.98
	β	0.78	0.01	55.10	0.0000		
$\lambda_2 = 10\%$	α	2.38	1.32	1.80	0.0741	0.973	8.52
	β	0.77	0.01	61.82	0.0000		
$\lambda_3 = 20\%$	α	2.69	1.10	2.46	0.0159	0.982	3.86
	β	0.77	0.01	71.56	0.0000		
$\lambda_4 = 50\%$	α	0.29	0.92	0.31	0.7573	0.991	1.59
	β	0.80	0.01	19.63	0.0000		
$n_3 = 144$							
$\lambda_1 = 5\%$	α	0.77	1.06	0.72	0.4707	0.976	12.30
	β	0.79	0.01	74.02	0.0000		
$\lambda_2 = 10\%$	α	1.21	0.93	1.30	0.1961	0.982	8.56
	β	0.78	0.01	82.35	0.0000		
$\lambda_3 = 20\%$	α	0.68	0.85	0.80	0.4271	0.985	5.66
	β	0.79	0.01	86.97	0.0000		
$\lambda_4 = 50\%$	α	-0.83	0.66	-1.26	0.2119	0.994	1.21
	β	0.81	0.01	105.85	0.0000		

5-3 التقدير بطريقة (M)

تم تقدير معاملات أنموذج الانحدار الخطي البسيط باستعمال طريقة (M) وبالاعتماد على دالتي الوزن (Tukey's Bisquare و Huber)، والنتائج تم عرضها في الجدول (4). إذ اشارت الى معنوية تأثير الكميات المنتجة من النفط الخام لكلتا الدالتين من خلال الاعتماد على قيمة (t) وبمستوى معنوية (0.05) ولحجوم العينات الثلاثة، وللمقارنة بين دالتي الوزن تم الاعتماد على قيمة معيار متوسط مربعات الخطأ، اذ اشارت قيمة هذا المعيار الى افضلية دالة الوزن

(Huber) اذ بلغت قيمة المعيار MSE (56.78 , 37.14 , 31.89) لحجوم العينات الثلاثة وعلى التتابع.

جدول (4): نتائج تقدير طريقة التقدير الحصينة (M)

دالة الوزن	المعلمات	تقدير المعلمات	S.E.	t-value	P-value	R^2	MSE
$n_1 = 72$							
M (Tukey)	α	3.47	5.25	0.66	0.5086	0.545	60.02
	β	0.76	0.04	18.07	0.0000		
M (Huber)	α	-0.09	5.49	-0.02	0.9872	0.659	56.78
	β	0.80	0.04	18.07	0.0000		
$n_2 = 120$							
M (Tukey)	α	3.27	1.42	2.31	0.0209	0.774	39.10
	β	0.76	0.01	57.99	0.0000		
M (Huber)	α	2.11	1.51	1.40	0.1608	0.876	37.14
	β	0.78	0.01	55.64	0.0000		
$n_3 = 144$							
M (Tukey)	α	1.94	1.03	1.88	0.0596	0.780	33.91
	β	0.77	0.01	76.14	0.0000		
M (Huber)	α	0.76	1.15	0.66	0.5069	0.896	31.89
	β	0.79	0.01	69.56	0.0000		

3-6: اختيار افضل طريقة تقدير لمعاملات أنموذج الانحدار:

تم اجراء مقارنة بين نتائج التقدير المثلى لطريقتي التقدير الحصينة موضوع البحث بالاعتماد على معيار المفاضلة متوسط مربعات الخطأ MSE علاوة عن المؤشر الاحصائي معامل التفسير R^2, الجدول (5) يبين تلك النتائج لحجوم العينات الثلاث. واشارت الى ان طريقة المربعات الصغرى المشذبة بنسبة تشذيب ($\lambda_3 = 20\%$) هي الافضل بالنسبة لحجم العينة (n_1=72) لتحقيقها اقل قيمة لمعيار المفاضلة متوسط مربعات الخطأ (MSE) واكبر قيمة لمعامل التحديد, اما بالنسبة لحجمي العينة (n_3 =144,n_2=120) فقد كانت طريقة المربعات الصغرى المشذبة بنسبة تشذيب ($\lambda_4 = 50\%$) هي الافضل.

جدول (5): نتائج المقارنة بين طريقتي التقدير الحصينتين (TLS) و (M)

حجم العينة	طريقة التقدير	R^2	MSE
$n_1 = 72$	M (Huber)	0.659	56.78
	TLS ($\lambda_3 = 20\%$)	0.958	8.95
$n_2 = 120$	M (Huber)	0.876	37.14
	TLS ($\lambda_4 = 50\%$)	0.989	1.89
$n_3 = 144$	M (Huber)	0.896	31.89
	TLS ($\lambda_4 = 50\%$)	0.994	1.21

4. المبحث الرابع : الاستنتاجات والتوصيات

1-4 الاستنتاجات

من خلال ما تم التوصل اليه من نتائج فيما يتعلق بتقدير انموذج الانحدار الذي يصف العلاقة بين الكميات المصدرة للنفط الخام والكميات المنتجة منه وبالاعتماد على بيانات عينات البحث، يمكن التوصل الى الاستنتاجات الاتية:

1- ان طريقة المربعات الصغرى المشذبة (TLS) افضل من طريقة التقدير (M) لحجوم العينات الثلاثة.

2- اختلاف نسب التشذيب المستعملة في طريقة المربعات الصغرى المشذبة بين حجوم العينات المعتمدة, اذ وجد ان افضل نسبة تشذيب هي 20% بالنسبة لحجم العينة المتوسط (n_1=72) في حين كانت نسبة التشذيب 50% هي الافضل بالنسبة لحجمي العينة الكبيرة (n_2 =120 ,n_3 =144) .

3- لم يتأثر نوع دالة الوزن المعتمدة في طريقة التقدير M بحجم العينة, اذ وجد ان دالة الوزن (Huber) كانت افضل من دالة الوزن (Tukey's Bisquare) ولكافة حجوم العينات.

2-4: التوصيات

1- استعمال طرائق تقدير حصينة اخرى لتقدير معاملات أنموذج الانحدار بوجود مشكلة عدم تجانس التباين.

2- استعمال طرائق التقدير اللامعلمية ومقارنتها مع طرائق التقدير الحصينة في حالة عدم تحقق فرضية تجانس تباين الاخطاء العشوائية.

المصادر

[1] - احمد، احمد ذياب. (2008). المقارنة بين بعض دوال طريقة M الحصينة في التوزيع اللوجستي في حالة أنموذج الانحدار الخطي المتعدد. مجلة ديالى، العدد30، الصفحات 288-296.

[2] - بخيت، حسين علي ، و سحر فتح الله. (2008). الاقتصاد القياسي (المجلد المكتبة الوطنية). عمان، الاردن: دار اليازوري العلمية للنشر والتوزيع.

[3] - حميد، زينب حامد. (2018). مقارنة بين الطرائق الاعتيادية (LS,4SIV) والطرائق الحصينة (LTS,2SLWLS,RA) لتقدير معلمات أنموذج ARX(1,1,1) للاحمال الكهربائية. مجلة العلوم الاقتصادية و الادارية، العدد 109، الصفحات 496-514, العراق.

[4] - عبدالله، رائد اسمر. (2018). إختيار أفضل طريقة اختبار لمشكلة عدم تجانس التباين في نموذج الانحدار المتعدد (مع تطبيق عملي). رسالة ماجستير، كلية الادارة والاقتصاد جامعة كربلاء، العراق.

[5] - علي، سيف حسين. (2019). اختيار افضل تقدير معالم إنموذج الانحدار الحصين مع تطبيق عملي. رسالة ماجستير، كلية الادارة والاقتصاد جامعة كربلاء، العراق .

[6] - كاظم، اموري هادي. (2009). مقدمة في القياس الاقتصادي. عمان ، الاردن: دار زهران للنشر والتوزيع.

[7] - كاظم، خالد جمال. (2018). المربعات الصغرى المشذبة الموزونة لتقدير تأثير مياه الصرف الصحي في تلوث مياه نهر دجلة/ محافظة واسط. مجلة العلوم الاقتصادية والادارية، العدد 109، الصفحات 486-495, العراق.

[8] - Almetwally, E. M., & Almongy, H. M. (2018). Comparison Between M-Estimation, S-Estimation, and MM-Estimation Methods of Robust Estimation with Application and Simulation. International Journal of Mathematical Archive, 9(11), pp. 55-63.

[9] - Anderson, C., & Schumacker, R. E. (2003). A Comparison of Five Robust Regression Methods with Ordinary Least Squares Regression: Relative Efficiency, Bias, and Test of the Null Hypothesis. Understanding Statistics , 2(2), pp. 79-103.

[10] - Breusch, T. S., & Pagan, A. R. (1979). A Simple Test for Heteroscedasticity and Random Coefficient Variation. Econometrica, Vol.47(No.5), pp. 1287-1294.

[11] - Gujarati, D. N. (2003). Basic Econometrics. New York: McGraw-Hill Companies,Inc.

[12] - Susanti, Y., Pratiwi, H., Sulistijowati, S., & Liana, T. (2014). M Estimation, S Estimation, and MM Estimation in Robust Regression. International Journal of Pure and Applied Mathematics, Vol. 91(No. 3), pp. 349-360. doi:http://dx.doi.org/10.12732/ijpam.v91i3.7

[13] - Zaman, A., Rousseeuw, P. J., & Orhan, M. (2001). Econometric applications of high-breakdown robust regression techniques. Economic Letters, 71, pp. 1-8. Retrieved from http://www.sciencedirect.com/science/article/pii/S0165176500004043