استخدام انحدار الحرف الموزون كأسلوب مقترح لمعالجة الشواذ والتداخل الخطي في التحليل التوسطي

ا.م. د. بشار عبد العزيز الطالب م. صالح مؤيد شاكر البقال
جامعة الموصل / كلية علوم الحاسبات والرياضيات

Using Weighted Ridge Regression as A Proposed Approach to Remedy Outliyngness and Multi-Collinearity in Mediation Analysis

Lec. Salih M. S. Al.b. Assist.Prof. Dr. Bashar A.Al_Talib
University of mosul / College of computer sciences and

تاريخ استلام البحث ٢٠١٥/١١/٢٢ تاريخ قبول النشر ٢٠١٦/٢/٢٢

المستخلص:

يتناول هذا البحث اقتراح استخدام أسلوب الحرف الحصين (الموزون) لمعالجة بعض المشاكل التي يعاني منها التحليل التوسطي (Mediation Analysis) والتي تتضمن وجود بعض القيم الشاذة في المتغير التوضيحي أو المتغير التابع أو كلاهما معا. فضلاً عن وجود مشكلة التداخل الخطي بين المتغير التوضيحي والمتغير التوسطي والذي يؤدي بدوره إلى ظهور تقديرات غير دقيقة. لذا فقد تم اقتراح استخدام أسلوب الحرف الحصين من اجل معالجة التطرف الموجود في المشاهدات فضلاً عن معالجة التداخل الخطي الموجود بين المتغير التوسطي والمتغير المستقل والتوصل إلى صيغة فعالة لمعالجة هذه المشاكل والحصول على حدود ثقة أكثر دقة مقارنة بحدود الثقة التي تم الحصول عليها باستخدام المربعات الصغرى الاعتيادية. وقد تم تطبيق الأسلوب المقترح من خلال دراسة محاكاة فضلاً عن تحليل بيانات حقيقية.

الكلمات المفتاحية: تحليل الوساطة، الانحدار الحصين، أسلوب الحرف الحصين.

Abstract:

In this paper we propose to use the Robust Ridge Regression procedure to deal with some problems that the Mediation analysis suffer from, these problems include the existence of some outlier points in the independent variable or dependent variable or both of them. As well as the existence of linear dependency between independent and Mediated variables which produce in turn the existence of in-accurate estimators. So the use of Robust Ridge Regression procedure is proposed to remedy the extremeness exist in the points, in addition to solve the Multi-Collinearity problem exists between the mediated and independent variable, and to obtain an active formula that address these problems. Furthermore getting more accurate confidence intervals compared with that of ordinary least squares method. The proposed procedure implemented through simulation study and real data analysis.

Keywords: Mediation Analysis, Robust Regression, Robust Ridge procedure.

المقدمة:

في كثير من الظواهر الموجودة في الطبيعة تكون هناك علاقات تبادلية (تأثير وتأثر) بين المتغيرات، وفي موضوع الانحدار يسمى المتغير المؤثر بالمتغير المستقل والذي نرمز له بـ X ، أما المتغير المتأثر (من قبل المتغير المستقل X) فيدعى بالمتغير التابع ويرمز له بـy . ولكن السؤال الذي يتبادر إلى الذهن هو كيف يرتبط هذان المتغيران ولماذا؟ فمثلا إذا كان هناك متغير مستقل يمثل ظاهرة الفقر مع متغير تابع يمثل ظاهرة جنوح الإحداث، أو دراسة كيف ان العلاج النفسي يقلل حالة الاكتئاب. فمن خلال هذه الأمثلة يمكننا أن نفهم لماذا أن X و y مرتبطان من

خلال متغير ثالث وهو المتغير التوسطي (Mediator) والذي عادة ما يرمز له بالرمز M، وبشكل أكثر دقة فإنَّ المتغير التوسطي هو المتغير الذي يتوسط السلسلة السببية التي تربط المتغير المستقل بالمتغير التابع أي أن المتغير المستقل يؤثر في المتغير التوسطي والذي بدوره يؤثر في المتغير التابع (Guralnike,1970).

عملية الوساطة Mediation Process

يمكن وصف عملية الوساطة بأنها سلسلة سببية تتضمن فكرة أن هناك متغيراً واحداً يؤثر على متغير ثانٍ والذي بدوره يؤثر على متغير ثالث. أن المتغير التوسطي والذي يرمز له بالمتغير M هو المتغير الذي يتوسط هذه العلاقة بين المتغير التنبؤي والمتغير التابع أو متغير المخرجات (Mackinnon & Fairchild,2009).

وبشكل بياني يمكن توضيح عملية الوساطة بالرسم التوضيحي التالي:

$$X \xrightarrow{a} M \xrightarrow{b} Y$$

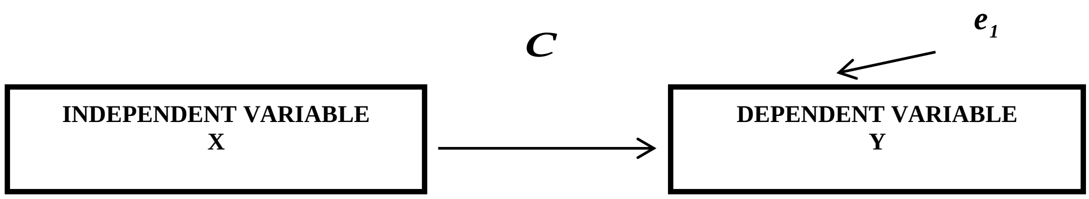
شكل رقم (١)

يوضح عملية الوساطة

مخطط نموذج الوساطة The Mediation Model Diagram

لإنجاز عملية تحليل الوساطة ينبغي توضيح النموذج التوسطي والذي يجب تحديدهُ بشكل صحيح مسبقا من حيث اتجاه العلاقة بين المتغيرات الثلاثة X و M و y. إذ يتم انجاز عملية الوساطة من خلال استخدام مقدرات المربعات الصغرى للنماذج وكما هو موضح في المخطط التالي مع المعادلات المستخدمة للتقدير (Emily et al.,2015) (Mackinnon,2008):

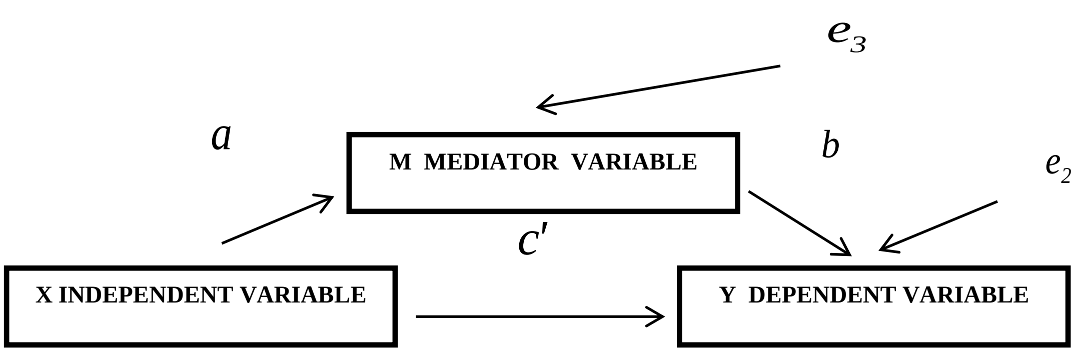
من المعادلة الرابطة للمتغير X مع المتغير y والموضحة في الشكل (٢)، إذ إنَّ e_1 يمثل الجزء غير مفسر من y من خلال علاقتهِ مع المتغير X، كما انهُ يمثل نموذج التأثير الكلي، وذلك لأنه يمثل العلاقة الكلية بين المتغير X والمتغير y من دون المتغيرات الأخرى.



$$y = i_1 + cX + e_1$$

شكل رقم (٢): الشكل يوضح مخطط مسار معادلة نموذج الانحدار

أما الشكل (٣) فهو يمثل أنموذج الوساطة، إذ إنَّ المتغير المستقل X يرتبط بالمتغير التوسطي M والذي بدوره يرتبط بالمتغير التابع y. كما انهُ يمثل أنموذج المتغير الثالث إذ هناك علاقة توسطية أساسية للمتغير X إلى المتغير M ثم إلى المتغير y. كما أن هناك علاقة للمتغير X مع المتغير y ولكن ليس من خلال المتغير M بل هناك تأثير مباشر من المتغير X إلى المتغير y (c'). كذلك يمكن ملاحظة أن هناك رموزاً فوق كل مسار يقابل كل علاقة، فمثلاً لعلاقة المتغير X مع المتغير M قد تم إعطاءه الرمز a، والعلاقة للمتغير M مع المتغير y بالرمز b، وعلاقة للمتغير X مع المتغير y بالرمز c'.



$$y = i_2 + c'X + bM + e_2$$

الشكل رقم(٣) شكل يوضح مخطط مسار المعادلة للنموذج التوسطي

كما يمكن ملاحظة أن علاقة المتغير X مع المتغير y لها الرمز العلوي (') والذي يطلق عليه بـ Prime (العلاقة c') كي تعكس التعديل للمتغير التوسطي في الشكل (٣) بينما لا يوجد هذا الرمز فوق المسار c في الشكل (٢) لأنه لم يعدل بالمتغير التوسطي M. اما الحد e_2 يمثل الجزء العائد للمتغير y الذي لم يتم تفسيره من قبل X و M. والحد e_3 تمثل الجزء العائد للمتغير M الذي لم يوضح من خلال علاقته مع المتغير X.

معادلات الانحدار المستخدمة لتقييم الوساطة

في تقييم عملية الوساطة لنموذج وساطة بسيط يتم استخدام ثلاث معادلات انحدار موضحة بالصيغ التالية (Mackinnon et al.,2002):

$$y = i_1 + cX + e_1 \qquad \dots(1)$$

$$y = i_2 + c'X + bM + e_2 \qquad \dots \quad (2)$$

$$M = i_3 + aX + e_3 \qquad \dots \quad (3)$$

إذ إنَّ y هو متغير تابع، و X متغير مستقل و M هو متغير توسطي، كما ان c تمثل العلاقة بين المتغير المستقل والمتغير التابع في المعادلة (1)، و c'تمثل المعلمة الرابطة للمتغير المستقل مع المتغير التابع والمكيفة بتأثير المتغير التوسطي M، أما b فهي المعلمة الرابطة للمتغير التوسطي مع المتغير التابع المكيف بتأثير المتغير المستقل، أما المعلمة a فهي المعلمة الرابطة للمتغير المستقل مع المتغير التوسطي، و e_1 و e_2 و e_3 تمثل الأجزاء غير المفسرة في المعادلات أعلاه على التوالي، وان i_1و i_2 و i_3 تمثل المقاطع (Intercepts).

التأثير الموسط Mediated Effect:

هناك أسلوبان لقياس التأثيرات التوسطية من نماذج الانحدار مستندة على الاستخدامات المختلفة للمعلمات a و b و c و c'. ويمثل حاصل الضرب(ab)مقدار التأثير التوسطي. ولكون المتغير X يؤثر على المتغير yبشكل غير مباشر من خلال المتغير M بذلك فان التأثير التوسطي معروف أيضا كتأثير غير مباشر (indirect effect) (Peng & Tyler,2016). اما تأثير المتغير X على المتغير y بعد تكييفه (adjustment) بالمتغير Mمن خلال المسار c' فيعرف على انهُ التأثير المباشر. كما أن التأثير التوسطي أيضا يساوي الفرق بين المعلمات c و c' $(c-c')$. ونتيجة لذلك فإنَّ التأثير الكلي يمكن أن يحلل إلى تأثير مباشر c' وتأثير غير مباشر $c-c'=ab$. إنَّ التبرير المنطقي وراء مقدار الوساطة ab هي انها وساطة تعتمد على مدى كون المتغير المستقل يؤثر على المتغير التوسطي (المعلمة a) والمدى للمتغير التوسطي الذي يؤثر بدوره على المتغير المعتمد (المعلمة b). و تعكس الكمية ab كيف ان تغير وحدة واحدة في المتغير X يؤثر على المتغير y بشكل غير مباشر من خلال المتغير M.

الوساطة الكاملة والوساطة الجزئية: Complete and Partial Mediation

هناك نوعين من عملية الوساطة وهما الوساطة الكاملة والوساطة الجزئية. والمقصود بالوساطة الكاملة ان العلاقة التوسطية تكون علاقة ذات وساطة كاملة أي أن المعلمة c' غير معنوية عندما يكون المتغير التوسطي M ضمن النموذج كما في المعادلة (2).

اما الوساطة الجزئية فتحدث عندما تكون a و b معنويين، ففي هذه الحالة فان الوساطة تكون موجودة فضلاً عن ذلك إذا كانت c' أيضا معنوية عندئذ يقال عنها وساطة جزئية (Fairchild & Mackinnon.,2009).

التأثير غير المباشر Indirect Effect

يعرف التأثير الغير المباشر بأنه قوة العلاقة بين المتغير X والمتغير y والذي يتوسطهما المتغير الثالث M. إنَّ كمية أو مقدار الوساطة تدعى بالتأثير غير مباشر والذي يمكن وصفهُ من خلال علاقتهِ مع التأثير الكلي والتأثير المباشر بالصيغة التالية:

التأثير الكلي التأثير المباشر التأثير غير المباشر

$$Total\ Effect = Direct\ Effect + Indirect\ Effect$$

$$c = c' + ab$$

اختبار النماذج التوسطية Test of Mediation Models

بشكل عام هناك ثلاثة أساليب تم تقديمها لاختبار النماذج المتضمنة متغيرات توسطية. يمثل الأسلوب الأول بأسلوب الخطوات السببية والذي اقترح من قبل (Judd & Kenny,1981a;1981b) وهو الأسلوب الأكثر استخداما في أدبيات علم النفس، إذ يحدد هذا الأسلوب سلسلة من الاختبارات للروابط في السلسلة السببية. اما الأسلوب الثاني العام فهو الأسلوب الذي يتمثل بالفرق بين المعاملات وإذ يتمثل بالفرق بين معامل الانحدار قبل وبعد تكييفه بالمتغير التوسطي (Freedman & Sckatzkin,1992). وأخيرا يستند الأسلوب الثالث العام على حاصل الضرب بين المعاملات المتضمنة في مسارات النموذج التوسطي (Alwin & Hauser,1975).

إنَّ الهدف الرئيسي من هذه الأساليب هو معرفة آلية أو ميكانيكية انتقال التأثير من المتغير التوضيحي إلى المتغير التابع، وتكمن هذه الآلية من خلال المتغير التوسطي الذي يتوسط العلاقة بين هذين المتغيرين. وهو ما يعرف بالتأثير غير المباشر (Indirect Effect). وسوف يتم في هذا البحث اعتماد الأسلوب الثالث الذي يتمثل بحاصل الضرب بين المعاملات.

حدود الثقة لتأثير الوساطة: Confidence Intervals for the Mediated Effect

إنَّ تقدير تأثير الوساطة والخطأ القياسي لهُ يمكن أن يستخدم لتكوين حدود الثقة لتأثير الوساطة في المجتمع. وتستخدم حدود الثقة بشكل واسع كونها تتضمن الخطأ في المقدر، وبالتالي توفر مدى من القيم الممكنة للتأثير بدلا من قيمة واحدة. وهناك توجه كبير جدا نحو إيجاد حدود الثقة (Harlow et al.,1997) والسبب في ذلك هو أن الباحثين يفترضون انهم يلاحظون قيمة التأثير فضلاً عن المعنوية الإحصائية. إنَّ حدود الثقة الواسعة تعني عدم الدقة في قياس قيمة التأثير في المجتمع (Krantz,1999). وقد بينا سابقاً أن المقدر $\hat{a}\hat{b}$ يعطي تقدير تأثير الوساطة. فضلا عن ذلك فان هناك عدة صيغ بديلة للخطأ القياسي لـ $\hat{a}\hat{b}$ التي يمكن استخدامها لتكوين حدود الثقة للتأثير غير المباشر ab. ويدعى الخطأ القياسي المستند على $\hat{a}\hat{b}$ بحاصل ضرب الخطأ

القياسي للمعاملات. وان هذه الصيغ يمكن أن تستخدم لتكوين حدود الثقة لتأثير الوساطة المستند على أساس المعاملات بأسلوب التوزيع المحاذي للتوزيع الطبيعي (Asymptotic Normal Distribution) والذي وضحهُ (Sobel,1982) والتي يأخذ الصيغ التالية:

$$Lower\ Confidence Limit(LCL) = \hat{a}\hat{b} - critical\ Value * (S_{\hat{a}\hat{b}}) \quad \dots \quad (4)$$

$$Upper\ Confidence\ Limit(UCL) = \hat{a}\hat{b} + critical\ Value * (S_{\hat{a}\hat{b}}) \quad \dots \quad (5)$$

إذ إنَّ مقدر التأثير التوسطي هو $\hat{a}\hat{b}$. وان $critical\ Value$ هي القيمة الحرجة المطلوبة لحساب حدود الثقة، وان $S_{\hat{a}\hat{b}}$ هو مقدر للخطأ القياسي لتأثير الوساطة. إنَّ الخطأ القياسي الأكثر استخداما لـ $\hat{a}\hat{b}$ $(S_{\hat{a}\hat{b}})$ هي صيغة تم اشتقاقها من قبل الباحث (Sobel,1982) ومستندة على المشتقات الأولى باستخدام طريقة دلتا لمتعدد المتغيرات (Folmer,1981). والصيغة الناتجة موضحة في أدناه إذ إنَّ $S^2_{\hat{a}}$ و $S^2_{\hat{b}}$ هي مربع الخطأ القياسي لـ $\hat{a}$ و $\hat{b}$ على التوالي:

يمكن تطبيق أسلوب دلتا من خلال استخدام الصيغة التالية:

$$V(\hat{a}\hat{b}) = D'VD \quad \dots \quad (6)$$

حيث أن D تمثل متجه المشتقات الجزئية بالنسبة إلى المعلمتين $\hat{a}$ و $\hat{b}$. اما V فتمثل مصفوفة التباين والتباين المشترك بين المعلمات $\hat{a}$ و $\hat{b}$. بذلك ومن خلال تطبيق المعادلة (6) نحصل على:

$$Q = V(\hat{a}\hat{b}) = \begin{bmatrix} \frac{\partial Q}{\partial \hat{a}} & \frac{\partial Q}{\partial \hat{b}} \end{bmatrix} \begin{bmatrix} S^2_{\hat{a}} & Cov(\hat{a},\hat{b}) \\ Cov(\hat{a},\hat{b}) & S^2_{\hat{b}} \end{bmatrix} \begin{bmatrix} \frac{\partial Q}{\partial \hat{a}} \\ \frac{\partial Q}{\partial \hat{b}} \end{bmatrix}$$

علماً أن التباين المشترك بين $\hat{a}$ و $\hat{b}$ يساوي صفر (Tofighi et al.,2009).

بذلك فإنَّ صيغة الخطأ القياسي للتأثير الغير مباشر تتمثل بالمعادلة التالية:

$$S_{(\hat{a}\hat{b})} = \sqrt{\hat{a}^2 S^2_{\hat{b}} + \hat{b}^2 S^2_{\hat{b}}} \quad \dots \quad (7)$$

وتوضح المعادلة (7) الصيغة المستخدمة لحساب التباين المشترك في كثير من برامج الحاسوب مثل برنامج EQS (Bentler,1997) و MPLUS (Mackinnon et al.,2008) وبرنامج LISREL (Joreskog & Sorbom,2001) لحساب مقدرات الخطأ القياسي لتأثيرات الوساطة.

مشاكل التحليل التوسطي Mediation Analysis Problems

إن المشاكل التي تعترض التحليل التوسطي تتشابه مع المشاكل التي تعترض فروض الانحدار الخطي العام، إذ قد يحصل انتهاك لقسم من هذه الفروض من خلال وجود قيم شاذة في

المتغير المستقل أو المتغير التابع أو كلاهما معاً والذي بدوره ينعكس على متجه قيم البواقي وبالتالي الحصول على مقدرات متحيزة عن المعلمات الحقيقية للمجتمع. فضلاً عن ذلك قد يكون هناك انتهاك في فرض آخر وهو الاستقلالية بين المتغيرات التوضيحية والذي لا يمكن تلافيه خاصة في النموذج التوسطي، والسبب في ذلك أنه بعد إجراء انحدار المتغير M على المتغير X يتم استخدام هذين المتغيرين كمتغيرين توضيحيين للتنبؤ بالمتغير y مما يؤدي إلى ظهور اعتمادية خطية بين المتغيرين X و M. ومثل هذه الاعتمادية تؤدي إلى نتائج غير دقيقة للمقدرات. وقد تم إجراء العديد من الدراسات حول هذه المشكلة والمسماة بتعدد العلاقة الخطية (Multi-collinearity) مثل (Fisher & Mason,1981).

وقد تناولت البحوث أساليب متعددة لمعالجة هذه المشكلة ومنها طريقة المكونات الرئيسية (Principal Component) (Mansfeld, Webster & Gunst,1977) ومقدرات الحرف (Ridge Regression) (Hoerl & Kennard,1970a).

إنَّ التداخل الخطي أو الاعتمادية الخطية بين المتغير X والمتغير التوسطي M يؤدي إلى انخفاض القوة لتحليلات الوساطة، إذ تعتمد قوة الاختبار على قوة الارتباط بين المتغير التوسطي والمتغير التابع فضلاً عن ذلك فأنها تعتمد أيضا على الارتباط بين المتغير المستقل والمتغير التوسطي (Beasly,2012).

وقد ناقش الباحثان (Hoyle & Kenny,1999) قوة الاختبار للفرضيات بشكل مختصر كدالة للتداخل الخطي بين المتغير X والمتغير M. إذ بينا أن تغيرات أكثر في المتغير التوسطي M من قبل المتغير التوضيحي X سوف يؤدي إلى تغاير اقل من المتغير M إلى المتغير التابع y. أي عندما يكون المعامل a كبير في نقطة معينة فان b يمكن أن يقل. وبتعبير آخر وعلى فرض أن هناك مثالاً نادر الحدوث بأن يكون فيه المعامل $a=1$ في أنموذج قياسي للمعادلة (3) إذ تكون المتغيرات في هذه الحالة قياسية، وعندها فإنَّ المتغير X سوف تفسر المتغير M بنسبة ١٠٠% وبذلك فإنَّ المتغير M لا يستطيع تفسير التغيرات في المتغير y بشكل واضح، أي أن b سوف تساوي صفر والسبب يعود للاعتمادية الخطية بين المتغيرين X و M.

إنَّ الطريقة الشائعة التي تستخدم في تحليل الوساطة هي الطريقة الاعتيادية المتمثلة بالمربعات الصغرى. فضلاً عن ذلك فإنَّ الاختبارات المتعلقة بالتأثيرات لتحليل الوساطة والمتضمنة باختبار التأثير غير المباشر وحدود الثقة تستند جميعها على نتائج المربعات الصغرى. وكما هو معلوم فإنَّ أسلوب المربعات الصغرى هو من الأساليب الدقيقة في حالة تحقق الفروض الأساسية للنموذج. إلا أنهُ عند عدم تحقق هذه الفروض فإنَّ أسلوب المربعات الصغرى سوف يعاني من مشاكل متعددة تختلف في تأثيرها حسب الفرض الذي تم انتهاكهُ. ومن المشاكل الشائعة في معظم البيانات هي مشكلة التداخل الخطي بين المتغيرات التوضيحية والتي ربما تكون أكثر خطورة في

حالة وجود ارتباط خطي ضعيف مقارنة بالحالة التي يكون فيها الارتباط تام، والسبب في ذلك هو عدم الاهتمام من قبل الباحث بهذه المشكلة كون الارتباط ضعيف، مما يؤدي إلى التأثير على المقدرات وتباينات هذه المقدرات والذي يؤدي إلى التأثير على الاختبارات المعنوية للمقدرات وكذلك حدود الثقة لها. كما أن هناك نوعاً آخر من المشاكل التي قد تؤدي إلى انتهاك فرض آخر من فروض النموذج والتي هي مشكلة القيم الشاذة في البيانات (outliers) والذي يؤدي بالنتيجة إلى سلوك متجه الأخطاء إلى توزيع غير التوزيع الطبيعي الذي يفترض ان $e_i \sim N(0,\sigma^2)$ (Rousseeuw & Leroy,1987)، وهذه المشكلة تؤثر سلبا على المقدرات من حيث اجتذابها نحو القيم الشاذة والتي تؤثر بدورها على معنوية الاختبارات للمقدرات وكذلك حدود الثقة لها.

أسلوب انحدار الحرف الحصين: Robust Ridge Regression

تم في هذا البحث استخدام أسلوب الحرف الحصين المقدم من قبل (Askin and Montgomery,1980) والمتمثل بالمعادلة التالية:

$$\hat{\beta}_{WR} = (X'WX + kI_k)^{-1} X'Wy \qquad \dots \qquad (8)$$

إلا أن استخدام هذا الأسلوب سيكون بشكل مختلف عن الأسلوب الذي استخدمه (Askin and Montgomery,1980). إذ استخدم Askin مقدرات M الحصينة لمعالجة التطرف والتداخل الخطي، فضلاً عن ذلك فقد استخدم كلٌ من (Pfaffenberger & Dielman,1990) أسلوب الحرف الحصين من خلال الدمج بين خصائص LAV(least Absolute value) ومقدر الحرف. كما قدم الباحثان (Samkar & Alpu,2010) أسلوب الحرف الحصين باستخدام مقدرات أسلوب M العام المستند على مقدرات ابتدائية مثل مقدر S ومقدرات MM وطبقاها على الانحدار الخطي. فضلاً عن ذلك فقد استخدم (Zahari et al.,2012) مقدر MM في انحدار الحرف ولكن على النموذج الخطي العام، كما استخدم (Pati et al.,2014) أسلوب انحدار الحرف الحصين ولكن من خلال استخدام Least Median Squares وقارنها مع المربعات الصغرى وطريقة الانحراف المطلق الأصغر. وفي هذا البحث تم استخدام أسلوب الحرف الحصين المستند على مقدرات MM من حيث الأوزان ومن حيث تقدير معلمة الحرف k، مع استخدام مقدارين للكفاءة النسبية لمقدرات MM كتحسين لمقدر الحرف الحصين ولكن على تحليل الوساطة من أجل تقدير التأثير التوسطي وإيجاد حدود الثقة لهُ ومقارنتها مع أسلوب المربعات الصغرى وبيان مدى تفوق الطريقة المقترحة على نتائج التحليل التوسطي. وبما أن الأسلوب هو دمج بين أسلوب الحرف ومقدرات MM فقد تم إعطاءه رمز RMM (Robust MM).

إنَّ تقدير التأثير غير المباشر أو التأثير التوسطي يتمثل بالمقدر $\hat{a}\hat{b}$ وهما المعلمتان المقدرتان بالمعادلتين (2) و (3) على التوالي. لقد تم إتباع أسلوب المقدر الحصين MM من اجل

معالجة القيم الشاذة الموجودة في المتغير التوضيحي X، إذ تم الحصول على المعلمة $\hat{a}$المقدرة بأسلوب حصين اي الحصول على المعلمة المقدرة $\hat{a}_{MM}$ مع تباين هذا المقدر $S^2_{\hat{a}_{MM}}$. أما المعادلة (2) فقد تم تقدير معلماتها من خلال معالجة القيم الشاذة الموجودة في المتغير X وكذلك القيم الشاذة الموجودة في المتغير M (والتي جاءت أصلا من القيم الشاذة الموجودة في المتغير X نتيجة انحدار المتغير Mعلى المتغير Xفي المعادلة (3)) باستخدام أوزان تم الحصول عليها أيضا من مقدرات MM، أما معلمة الحرف k فقد تم تقديرها باستخدام أسلوب معدل لما اقترحهُ (Hoerl & Baldwin.,1975) وذلك باستخدام مقدرات MM وليس مقدرات المربعات الصغرى التقليدية وبذلك يتم الحصول على المعلمة المقدرة $\hat{b}_{R+MM}$ مع تباين هذا المقدر S^2_{R+MM}.

وبذلك يتم الحصول على المقدرات $\hat{a}$ و $\hat{b}$ وتباينهما ولكن بشكل أكثر دقة من خلال معالجة الشواذ والتداخل الخطي الموجود بين المتغيرات من خلال معادلة واحدة وبالتالي الحصول على مقدر التأثير التوسطي ولكن بأسلوب الحرف الحصين $\hat{ab}_{RMM}$ مع تباين المقدر $S^2_{\hat{ab}_{RMM}}$.

تقدير الحرف الحصين: Robust Ridge Estimation

يمكن توضيح مقدر الحرف الحصين لمعالجة الشواذ والتداخل الخطي الموجود في المتغيرات Xو M لمعادلة الانحدار (2). إذ انهُ من خلال النموذج الخطي العام والمتمثل بالصيغة التالية:

$$y = X\beta + \varepsilon \qquad \dots \quad (9)$$

وعند وجود مشكلة عدم تحقق فرض تجانس تباين الخطأ عادة ما يتم استخدام أسلوب المربعات الصغرى الموزونة. إذ يتلخص هذا الأسلوب باستخدام أوزان محددة من اجل الحصول على مقدرات كفؤة وغير متحيزة وتمتلك خاصية كونها أفضل مقدرات خطية غير متحيزة(Best Linear Unbiased Estimate)(BLUE). إنَّ أساس هذا الأسلوب يتضمن ضرب طرفي النموذج الخطي البسيط المتمثل بالمعادلة (9) بمصفوفة قطرية تتضمن في عناصر قطرها على الجذر التربيعي لهذه الأوزان وكما يلي (كاظم،٢٠٠٢):

$$P^{-1}y = P^{-1}X\beta + P^{-1}\varepsilon \qquad \dots \qquad (10)$$

علماً أن

$$P^{-1}=\begin{bmatrix}\sqrt{w_1} & & & \\ & \sqrt{w_2} & & \\ & & \cdot & \\ & & & \cdot \\ & & & & \sqrt{w_n}\end{bmatrix}, W^{-1}=(PP')^{-1}=\begin{bmatrix}w_1 & & & \\ & w_2 & & \\ & & \cdot & \\ & & & w_n\end{bmatrix}$$

بذلك يمكن كتابة المعادلة (10) بالشكل التالي:

$$y^*=X^*\beta+\varepsilon^* \qquad \ldots \quad (11)$$

حيث أن

$$y^*=\sqrt{w}\,y\ ,\ x^*=\sqrt{w}\,X\ ,\ \varepsilon^*=\sqrt{w}\,\varepsilon \qquad \ldots \quad (12)$$

إن الفرض الأساسي المعتمد في إتباع طريقة المربعات الصغرى الموزونة (WLS) هو:

$$E(e_i^2)=\sigma^2\ e_i^*=\sigma_{e_i}^2\ W^{-1}$$

علما بأنَّ $E(S_e^2)=\sigma_e^2$ يمثل تباين العينة، اما (w_i) فما هي إلا عبارة عن أوزان (Weights) وبمعرفتها تكون مسالة التقدير والاختبار والتنبؤ بسيطة وممكنة.

إن النموذج الخطي العام الموضح في المعادلة (11) أعلاه يحقق الفروض الأساسية اللازمة لتطبيق طريقة المربعات الصغرى وذلك لان:

$$E(e^*e^{*\prime})=\sigma_{e^*}^2\ I_n \qquad \ldots \quad (13)$$

بذلك فإنَّ

$$e^*\sim N(0,\sigma_e^2) \qquad \ldots \quad (14)$$

من خلال المعادلة (13) تتحقق فرضيتي تجانس تباين الخطأ وانعدام الارتباط الذاتي بين الأخطاء، وعليه فإنَّ الفرضيات الأساسية الخاصة بنموذج الانحدار متحققة وبالتالي يمكن إتباع أسلوب المربعات الصغرى لتقدير النموذج (11) للحصول على متجه مقدرات المربعات الصغرى الموزونة. إذ بعد التفاضل الجزئي بالنسبة إلى β ومساواتها بالصفر يتم الحصول على:

$$X^{*\prime}X^*\hat{\beta}_{WLS}=X^{*\prime}y^* \qquad \ldots \quad (15)$$

وبضرب طرفي المعادلة (15) بالمصفوفة $(X^{*\prime}X^*)^{-1}$ يتم الحصول على:

$$\hat{\beta}_{WLS}=(X^{*\prime}X^*)^{-1}X^{*\prime}y^* \qquad \ldots \quad (16)$$

إنَّ المقدر أعلاه هو مقدر المربعات الصغرى الموزونة، وهو مقدر غير متحيز للمعلمة الأصلية. أي إن:

$$E(\hat{\beta}_{WLS})=\beta \qquad \ldots \quad (17)$$

بذلك فإنَّ مقدر المربعات الصغرى الموزونة هو مقدر غير متحيز للمعلمة الأصلية.

ومن أجل تقدير معلمات الحرف الحصين يتم إتباع الخطوات التالية:

من خلال المعادلة (15) فإنَّه في هذه المرحلة يتطلب الآمر كما هو معروف إضافة مصفوفة قطرية kI_k ذات بعد $p \times p$ إلى المصفوفة $(X^{*\prime}X^{*})$ قبل أخذ المعكوس لها وكما يلي (Hoerl & Kennard,1970a):

$$(X^{*\prime}X^{*}+kI_k)\hat{\beta}=X^{*\prime}y^{*} \quad \dots (18)$$

وبضرب طرفي المعادلة (18) بالمصفوفة $(X^{*\prime}X^{*}+kI_k)^{-1}$ نحصل على:

$$\hat{\beta}_{WR}=(X^{*\prime}X^{*}+kI_k)^{-1}X^{*\prime}y^{*} \quad \dots (19)$$

وبإرجاع كل من X^{*} و y^{*} إلى صيغتهم الأصلية وحسب المعادلة (12) يتم الحصول على:

$$\hat{\beta}_{WR}=(X'\sqrt{w}'\sqrt{w}X+kI_k)^{-1}X'\sqrt{w}'\sqrt{w}y)$$

$$\hat{\beta}_{WR}=(X'WX+kI_k)^{-1}X'Wy \quad \dots (20)$$

بذلك فإنَّ المعادلة (20) تمثل مقدر الحرف الحصين بالأوزان الترجيحية.

خواص مقدر الحرف الحصين: Properties of Robust Ridge estimator

١- التحيز Biasedness

إنَّ مقدر الحرف الحصين هو عبارة عن تحويل خطي لمقدر المربعات الصغرى الموزونة، وهذا التحويل يعتمد على X و y و W. أي أنهُ يمكن تعويض الحد الأيسر $X^{*\prime}X^{*}\beta_{WLS}$ من المعادلة (15) بدلا من $X^{*\prime}y^{*}$ في المعادلة (20) لنحصل على:

$$\hat{\beta}_{WR}=(X^{*\prime}X^{*}+kI_k)^{-1}X^{*\prime}X^{*}\hat{\beta}_{WLS}$$

$$\hat{\beta}_{WR}=\tau\hat{\beta}_{WLS} \quad \dots (21)$$

حيث أن

$$\tau=(X^{*\prime}X^{*}+kI_k)^{-1}(X^{*\prime}X^{*}) \quad \dots (22)$$

وبأخذ التوقع لطرفي المعادلة (21) نحصل على:

$$E(\hat{\beta}_{WR})=\tau E(\hat{\beta}_{WLS})$$

ومن المعادلة (17) فان:

$$E(\hat{\beta}_{WR})=\tau\beta \quad \dots (23)$$

أي أن مقدر الحرف الحصين متحيز للمعلمة الأصلية، وان مقدار التحيز هو:

$$\tau=(X'WX+KI_n)^{-1}(X'WX)$$

٢- التباين Variance

ومن اجل إيجاد تباين مقدر الحرف الحصين، يتم أخذ التباين لطرفي المعادلة (21) أي أن:

$$Var(\hat{\beta}_{WR}) = Var(\tau\hat{\beta}_{WLS})$$

$$Var(\hat{\beta}_{WR}) = Var(\tau(X^{*\prime}X^{*})^{-1}X^{*\prime}y^{*})$$

$$Var(\hat{\beta}_{WR}) = Var(y^{*})\tau(X^{*\prime}X^{*})^{-1}(X^{*\prime}X^{*})(X^{*\prime}X^{*})^{-1}\tau'$$

$$\therefore Var(\hat{\beta}_{WR}) = \sigma^2_{WR}\tau(X^{*\prime}X^{*})^{-1}\tau' \qquad \ldots \quad (24)$$

٣- تقدير تباين العينة :Sample Variance Estimation

$$S_e^2 = \frac{1}{n-k-1}e^{*\prime}e^{*}$$

$$S_e^2 = \frac{1}{n-k-1}(y^{*} - X^{*}\hat{\beta}_{WR})'(y^{*} - X^{*}\hat{\beta}_{WR})$$

ومن خلال المعادلة (12) نحصل على:

$$S_e^2 = \frac{1}{n-k-1}(\sqrt{W}y - \sqrt{W}X\hat{\beta}_{WR})'(\sqrt{W}y - \sqrt{W}X\hat{\beta}_{WR})$$

$$S_e^2 = \frac{1}{n-k-1}e'_{B_{WR}}We_{B_{WR}}$$

$$S_e^2 = \sigma^2_{WR} \qquad \ldots \quad (25)$$

بذلك ومن خلال المعادلة (20) يمكن حل مشكلتي التداخل الخطي ومشكلة القيم الشاذة في نفس الوقت. وقد تم تطبيق ذلك على أنموذج الوساطة، إذ تم تعويض المصفوفة W بأوزان تم أخذها من مقدرات MM الحصينة والتي تتميز بامتلاكها نقطة انهيار عالية تصل إلى 50% وكفاءة نسبية عالية في الوقت نفسهُ. أما قيمة معلمة الحرف k فقد تم استخدام اسلون الذي اقترحهُ (Hoerl & Baldwin.,1975) والذي يأخذ الصيغة التالية:

$$k = \frac{pS^2}{\hat{\beta}'_{LS}\hat{\beta}_{LS}}, \quad S^2 = \frac{(y - X\hat{\beta}_{LS})'(y - X\hat{\beta}_{LS})}{n-p-1}$$

إذ إنَّ p هي عدد المتغيرات المستقلة. وقد تم في هذا البحث استخدام أسلوب MM بدلاً من أسلوب المربعات الصغرى في تقدير معلمة الحرف k باستخدام الصيغة التالية:

$$k = \frac{pS^2_{MM}}{\hat{\beta}'_{MM}\hat{\beta}_{MM}}, \; S^2_{MM} = \frac{(y - X\hat{\beta}_{MM})'(y - X\hat{\beta}_{MM})}{n-p-1}$$

وبذلك ومن خلال المعادلات (9-20) ستتم معالجة التداخل الخطي الموجود بين المتغير X والمتغير M، وكذلك معالجة القيم الشاذة الموجودة في المتغيرات وباستخدام معادلة واحدة.

مقدر MM (MM-Estimator)

قدم الباحث (Yohai,1987) مقدر MM كمقدر محسن عن المقدر S. أي أن مقدر MM يعتمد على مقدر S. لذا في بادئ الأمر يجب توضيح مقدر S. أن المقدر S يمتلك نقطة انهيار ((Breakedown point (BDP)) عالية تبلغ 50% ولكن على حساب الكفاءة النسبية. وقد وصف كلٌ من (Olive & Hawkins,2008) خاصية نقطة الانهيار العالية على انها أصبحت تقليداً متبعاً وانها لا تعد شرطاً كافياً للمقدر الحصين الجيد. وحقيقة الأمر انهُ يجب ملاحظة قيمة نقطة الانهيار من حيث إنهُ إذا كان تأثير القيم الشاذة يمكن أن يكون غير محدد وعندها يعد ذلك انهياراً في مقدار مقاومة المقدر لتلك القيم. لذلك فإنَّ كلاً من (Rousseew & Leroy,1987) قَدَّموا ثلاثة شروط يجب أن تحققها دالة الهدف لمقدر S. والشرط المهم هو الشرط الثالث والذي كان يتضمن أن نقطة انهيار هذا المقدر تكون 50%. أي أن:

$$\frac{Q}{\rho(a)} = 0.5 \qquad \dots \qquad (26)$$

إذ إنَّ $\rho(a) = a^2/6$. أن دالة الهدف المستخدمة في مقدر S هي واحدة من الدوال المرتبطة مع دالة الوزن Bisquare لـ Tukey. اذ فرض كلٌ من (Rousseew & Leroy,1987) أن Q هي $Q = E\Phi(\rho(e))$. والتي هي القيمة المتوقعة لدالة الهدف إذا تم افتراض أن e تمتلك توزيع طبيعي قياسي. لقد ذكر كلٌ من (Rousseew & Yohai,1984) عندما يكون ثابت التوليف $a = 1.547$ فعندها سوف يحقق ذلك الشرط الثالث للمقدر S والذي هو أن نقطة انهيار S هي 50% (BDP). بذلك فإنَّ:

$$Q = E\phi(\rho(e)) = \int_{-\infty}^{\infty} \rho(e)\, f(e)\, dx$$
$$= 2\int_{0}^{\infty} \rho(e)\, f(e)\, dx \Rightarrow 2\int_{0}^{1.547} \rho(e)\, f(e)\, dx + 2\int_{1.547}^{\infty} \rho(e)\, f(e)\, dx$$

علماً أن
$$\rho(e) = \begin{bmatrix} \frac{e^2}{2} - \frac{e^4}{2a^2} + \frac{e^2}{6a^2} & if\,|e| \le a \\ \frac{a^2}{6} & if\,|e| > a \end{bmatrix}$$

إذ إنَّ
$$f(e) = \frac{1}{\sqrt{2\pi}} e^{-\frac{e^2}{2}} \quad , \therefore Q = 0.199$$

وحسب المعادلة (26) فإنَّ $(0.199/\rho(1.547)) = 0.498 \approx 0.50$. وإذ إنَّ قيمة ثابت التوليف a في دالة Bisquare لـ Tukey يمثل الثابت الذي بزيادتهُ سوف تزداد الكفاءة النسبية للمقدر. ولكن إذا تم زيادة الثابت a سوف يؤدي إلى زيادة قيمة Q وبالتالي عند تطبيق المعادلة (26) التابعة لنقطة الانهيار سوف يكون الناتج اقل من 50%. بذلك فإنَّ زيادة الكفاءة

النسبية في مقدر S سوف يؤدي إلى النقصان في نقطة انهيار المقدر. ومن هذا المنطلق تم اقتراح مقدر MM من قبل (Yohai,1987). إذ إنَّ هذا المقدر يتم حسابهُ من خلال ثلاث مراحل وهي:

المرحلة الأولى: يتم استخدام مقدر ابتدائي ذو نقطة انهيار عالية والذي يرمز لهُ بـ $\tilde{\beta}$ (أي مثل مقدر S). علماً أن هذا المقدر غير كفوء بشكل كافٍ. باستخدام هذا المقدر فإنَّ الأخطاء التابعة لهُ سوف تحسب من خلال $e_i(\tilde{\beta}) = y_i - x_i'\tilde{\beta}$.

المرحلة الثانية: يتم استخدام أخطاء المقدر $\tilde{\beta}$ والتي هي$S(e_1(\tilde{\beta}),...,e_n(\tilde{\beta}))$ والتي يرمز لها بالرمز S_n. وأن دالة الهدف المستخدمة في هذه المرحلة يتم الرمز لها بـ ρ_0.

المرحلة الثالثة: في هذه المرحلة يتم إيجاد مقدر MM من خلال

$$\sum_{i=1}^{n} x_{ij}\psi_1\left(\frac{y_1 - x_i'\beta}{S_n}\right) = 0 \qquad j = 1,\dots,p$$

إذ إنَّ

$$\psi_1(e) = \frac{\partial \rho_1(e)}{\partial e}$$

أن دالة الهدف ρ_1في هذه المرحلة يجب أن لا تكون نفس دالة ρ_0، ولكن يجب أن تحقق الشروط التالية:

١- ρ تكون متماثلة وهي دالة قابلة للاشتقاق ومستمرة، وان $\rho(0) = 0$.

٢- إنَّ $a > 0$ إذ إنَّ ρ تزداد تماماً عند الفترة $[0,a]$وثابتة عند الفترة $[a,\infty)$.

٣- $\rho_1(e) \le \rho_0(e)$.

إنَّ المرحلتين الأولى والثانية تكون مخصصة لنقطة انهيار مقدر MM والتي تبلغ 50%، بينما المرحلة الثالثة تهدف إلى كفاءة نسبية عالية للمقدر. لقد اثبت كلٌ من (Hadi and Simonoff,1993; Yohai,1987) أن نقطة انهيار مقدر MM تعتمد على اختيار ثابت التوليف في أولى مرحلتين، بينما الكفاءة النسبية للمقدر تعتمد على ثابت التوليف في المرحلة الثالثة. لذلك وخلافاً لمقدر M فإنَّ نقطة الانهيار والكفاءة النسبية لمقدرات MM تكون مستقلة الواحدة عن الأخرى. إذ بينما نقطة الانهيار تبقى ثابتة عند 50% فإنَّ الكفاءة النسبية يمكن أن تقرب إلى الواحد الصحيح وهي الحالة المثالية (Bianco et al.,2005).

وقد ناقش الباحث (Yohai, 1987) خطورة اختيار ثابت التوليف العالي في المرحلة الثالثة من اجل الحصول على كفاءة نسبية عالية، والذي عادة يستخدم ثابت التوليف بمقدار $a = 4.24$ للحصول على كفاءة نسبية تبلغ 95%. في حين أن مقدر MM سوف تبقى نقطة انهيارهُ ثابتة كما حددت في أولى مرحلتين. ولذلك فان مقدر MM سوف ينهار فعلاً ويكون أكثر حساسية للشواذ واقل حصانة مقارنة بحالة كونهِ يمتلك كفاءة نسبية أقل. كما اثبت الباحث (Stuart ,2011) أن مقدرات MM تكون أكثر حصانة مع نقطة انهيار 50% في حالة جعل الكفاءة النسبية اقل من 95%، إذ يتم تقليل الكفاءة النسبية من خلال تقليل ثابت التوليف لدالة

الهدف في المرحلة الثالثة. وسيتم في هذا البحث اختيار ثابت التوليف $a = 3.42$ وكذلك $a = 4.24$ للحصول على كفاءة نسبية 84.7% و 95% على التوالي والمقارنة بينهما.

مقاييس التقييم Evaluation Measurement:

تم تقييم المقدرات من خلال ما يلي:

١- **سعة حدود الثقة Width of confidence intervals:** ويقصد به المسافة بين الحد الأعلى والحد الأدنى لحدي الثقة، إذ كلما كان السعة قليلة كلما كان ذلك أفضل (Preacher and Selig,2012).

٢- **التغطية Coverage:** المقصود بالتغطية نسبة المقدرات التي تقع ضمن الحدود، وعندما تكون نسبة الخطأ من النوع الأول 0.05 فهذا يعني ان المعلمة الحقيقية يجب أن تكون ضمن حدود الثقة 95 مرة من أصل 100 حد ثقة حتى تكون التغطية جيدة. علماً أن التغطية والسعة لحدود الثقة يكونان متعارضين، وذلك لأنَّ كلما كانت حدود الثقة ضيقة كلما كان ذلك أفضل ولكن هذا الضيق سوف ينعكس سلباً على التغطية؛ لأنَّ في هذه الحالة ستكون التغطية اقل لضيق هذه الحدود (Preacher and Selig,2012).

٣- **القوة Power:** المقصود بالقوة هي قوة الاختبار Power of the test أي هي احتمالية رفض فرضية العدم عندما تكون خاطئة، إذ كلما كانت هذه الاحتمالية عالية كلما كان ذلك أفضل(Beasley,2012) .

٤- **التحيز للمعلمة المقدرة عن المعلمة الحقيقية Bias:** المقصود بالتحيز هو الفرق بين المعلمة المقدرة والمعلمة الحقيقية (Pati et al.,2014).

٥- **الجذر التربيعي للانحراف القياسي للتحيز Root Standard deviation of Bias:** هو مربع الفرق بين المعلمة المقدرة من كل عينة من عينات مونتي كارلو والمعلمة الحقيقية ثم جمع هذه الفروق المربعة واخذ الجذر التربيعي لها. وكلما كان هذا المقياس أقل كلما كان ذلك أفضل (Pati et al.,2014).

٦- **الجذر التربيعي لمتوسط مربعات الخطأ Root Mean Squears error:** هو الجذر التربيعي لمتوسط مربعات الأخطاء لنموذج الانحدار الكلي. وكلما كان هذا المقياس اقل كلما كان ذلك أفضل (Lawrence et al.1990).

٧- **الكفاءة Efficiency:** يقصد بالكفاءة قسمة الجذر التربيعي للانحراف القياسي للتحيز المستخرج بالخطوة (٥) لأسلوب المربعات الصغرى على الجذر التربيعي للانحراف القياسي للتحيز لأسلوب الحرف الحصين. وكلما كان هذا المقياس أكبر من الواحد الصحيح كلما دل ذلك على كفاءة الأسلوب المقترح مقارنةً بالمربعات الصغرى (Pati et al.,2014).

أساليب الكشف عن القيم الشاذة:

يوجد العديد من أساليب الكشف عن وجود القيم الشاذة في البيانات نذكر منها:

١ - استخدام أسلوب الرسم الصندوقي Boxplot للكشف عن القيم الشاذة الموجودة في المتغير التابع y .

يتألف هذا الرسم من خمسة مكونات وهي,(Carter et al.2009):

- الوسيط (Median)
- الربيع الأول والربيع الثالث واللذان يقعان على نهايتي الصندوق من الأعلى والأسفل على التوالي.
- قيم نقاط البيانات التي تقع مجاور الأسيجة (fences) العليا والدنيا.
- الشاربان (whiskers) واللذان يوصلان بين النهايات للصندوق وبين الأسيجة.
- القيم الشاذة التي تقع خارج حدود الأسيجة.

وسيتم توضيح الرسم بالجانب العملي وكيفية الكشف عن القيم الشاذة.

٢- أسلوب MVE (Minimum Volume Ellipsoid)

قدم الباحث (Rousseew,1983) أسلوب حجم القطع الناقص الأصغر MVE (Minimum Volume Ellipsoid) لتحديد القيم الشاذة لمصفوفة من المتغيرات. ففي التحليل الحصين فإنَّ مقدر أصغر قطع ناقص (MVE) غالباً ما يستخدم لتقدير الموقع والتشتت في متعدد المتغيرات. إذ يعرف مقدر MVE لمصفوفة التباين المشترك على انهُ اصغر قطع بيضوي يحوي نصف المشاهدات بينما تقدير الموقع في (MVE) فهو نقطة الوسط لذلك الشكل البيضوي. إنَّ العمل بخوارزمية مقدر MVE تتلخص من خلال اختيار عدد من العينات الجزئية مساوي إلى C^{n}_{p+1} من مصفوفة البيانات الأصلية. بعد ذلك يتم حساب قيمة الأوساط الحسابية $\mu_{(j)}$ ومصفوفة التباين المشترك $\sum_{(j)}$ لكل عينة جزئية (j) بحجم $p+1$. ثم يتم حساب المسافات التربيعية وفق المعادلة التالية:

$$D_j^2=(x_i-\mu_J)'\sum\nolimits_J^{-1}(x_i-\mu_J)$$

يتم اختيار أفضل عينة جزئية (j) من خلال العينة التي تمتلك اصغر دالة هدف والتي تتمثل بالمعادلة التالية:

$$\tilde{J}=arg\ min\ det\sum\nolimits_J D^2_{(hp)(J)}$$

العينة التي تمتلك أصغر دالة هدف يتم حساب μ و $\sum$ لها. تعاد الخطوات أعلاه وبزيادة مشاهدة واحدة لكل مرحلة إلى أن نصل إلى حد التوقف وهو حد الوصول إلى حجم العينة الجزئية (np) إذ تكون هذه العينة متجانسة ومتسقة وتتجه نحو المركز. يجري استخراج متجه الأوساط ومصفوفة التشتت والتي من خلالها نستخرج المسافات

التربيعية الحصينة وتشخص من خلالها المشاهدات الشاذة للعينة المستخرجة (وهي المشاهدة التي تزيد قيمة المسافة الحصينة لها عن قيمة $\chi^2_{n,1-\alpha/2}$) (٢,١٠، لقاء وأفراح). كما يمكن أن يعطي هذا الأسلوب توضيح رسومي حول القيم المتطرفة في البيانات من خلال تقسيم مربع الرسم إلى أربع أجزاء. فالجزء الأيمن العلوي يظهر بهِ القيم الشاذة مع رقم الصف الذي توجد بهِ هذه القيم في المتغيرات. وسيتم توضيح ذلك بشكل أكثر في الجانب العملي.

أساليب الكشف عن التداخل الخطي:

هناك عدة أساليب للكشف عن التداخل الخطي سيتم استخدام البعض منها.

١- **أسلوب الارتباط الثنائي بين المتغيرات:** فمن خلال مصفوفة الارتباط بعد تحويل المتغيرات إلى الصورة القياسية. إذ يتم تحديد قيمة معامل الارتباط التي تدل على وجود ارتباط معنوي بين متغيرين من خلال اختبار t (المشهداني وهرمز،١٩٨٩).

٢- **أسلوب العدد الشرطي (C.N.) Condition Number:** إذ يتمثل هذا الأسلوب بإيجاد الجذور المميزة لمصفوفة الارتباط بين المتغيرات، ومن ثم تقسم أكبر جذر مميز على اصغر جذر مميزة (Belsley et al.,1980). أي أن

$$C.N. = \frac{Max\ \ell_i}{Min\ \ell_i} \qquad \dots \qquad (27)$$

فإذا كانت قيمة العدد الشرطي تساوي الواحد الصحيح فهذا يدل على استقلالية المتغيرات فيما بينها، أما إذا كانت اكبر فهذا يدل على وجود تعدد علاقة خطية بين المتغيرات. وقد وضح Belsley أن التداخل الخطي الضعيف تكون عندهُ قيمة العدد الشرطي من 5 إلى 10، أما التداخل الخطي المعتدل إلى القوي تكون قيمة العدد الشرطي من 30 إلى 100 (Lawrence & Arthur,1990).

٣- **اختبار Farrar-Glauber:** إذ يستند هذا الاختبار على الإحصاءة χ^2 اذ يتم اختبار الفرضية التالية (٢٠٠٢، كاظم):

$$H_0 : (x_j)\ Orthogonal\ ,\ H_0 : (x_j)\ Not\ \ Orthogonal$$

$$\chi^2_0 = -[n - 1 - \frac{1}{6}(2 * p + 5)] * ln(D) \qquad \dots \qquad (28)$$

n هو حجم العينة و p تمثل عدد المتغيرات المستقلة. وان $ln(D)$ تمثل اللوغاريتم الطبيعي لمحدد مصفوفة معاملات الارتباط بين المتغيرات ثم يتم مقارنة χ^2_0 المحسوبة مع الجدولية بدرجة حرية $p(p-1)/2$ ومستوى معنوية معين. فإذا ظهر أن القيمة المحسوبة أكبر من الجدولية فهذا يعني أن هناك تداخلاً خطياً بين المتغيرات.

الجانب العملي:

لقد تم في هذا الجانب إجراء تطبيقين، الأول هو أسلوب محاكاة والثاني هو التطبيق على بيانات حقيقية.

أ- أسلوب المحاكاة:

تم تطبيق أسلوب المحاكاة على بيانات تم توليدها عشوائياً. إذ تم توليد المتغير المستقل X عشوائياً باستخدام التوزيع الطبيعي القياسي، كما تم توليد متجه الأخطاء التابع إلى معادلات الانحدار (2) و(3) وفق التوزيع الطبيعي القياسي أيضا. وقد تم استخدام أربعة أحجام من حجم التأثير وهي الصفر والحجم الصغير والمتوسط والكبير(Zhang & wang,2007). أي أن

$$Zero : a=b=0, Small : a=b=0.14, Medium : a=b=0.39, Large : a=b=0.59$$

وتم توليد أربعة أحجام من العينات وهي (25,50,100,500) وبنسب تلويث (0%,5%,20%,40%) لكل من المتغير المستقل والتوسطي وبوساطة كاملة أي أن $c'=0$، إذ إنَّ النتائج لن تختلف عن حالة تعويض $c'\neq 0$ كما بين (Mackinnon et al.,2004). وتم استخدام مستويين من الكفاءة وهي 95% و 87.4%. وبذلك سيكون هناك 128 تركيبة من المقدرات (4 أحجام عينات ×4 مستويات من التلويث × 4 مستويات من حجم التأثير × 2 مستوى من الكفاءة). تم استخدام أسلوب مونتي كارلو بـ 1000 تكرار لكل تركيبة.

النتائج:

نظراً لكثرة الجداول والبالغ عددها 218، فقد تم عرض الجداول التي توضح أعلى نسبة تلويث والبالغة 40% لكل عينة من اجل توضيح مدى مقاومة الأسلوب المقترح لظاهرة الشواذ العالية، وكذلك عرض جداول توضح أعلى مستوى للتأثير من أجل بيان مدى معالجة الأسلوب المقترح لظاهرة التداخل الخطي بين المتغير التنبؤي والمتغير التوسطي. كما تم عرض جدول لمتوسط جميع الجداول ولجميع التراكيب. والجداول المعروضة تضمنت عرض النتائج في الحالتين التي يكون فيها ثابت التوليف $a=4.24$ ومرة أخرى $a=3.42$ من أجل توضيح زيادة الكفاءة النسبية للمقدر المقترح عند تقليل ثابت التوليف a.

في بادئ الأمر سيتم ملاحظة الجدول (١) أدناه والذي يتضمن النتائج التابعة لحجم العينة n=25. إذ نلاحظ إنَّ مقدر RMM قد أظهر تفوقهُ على مقدر OLS ولجميع المقاييس. إذ إنَّ مقياس السعة للمقدر RMM كانت أقل من OLS لحدود الثقة التابعة لهما. كما أن حدود الثقة لـ RMM ظهرت معنوية لعدم احتوائها على الصفر في حدودها، بينما مقدر OLS كان يحتوي على الصفر في حدودهِ مما يدل على عدم معنوية التأثير التوسطي، مع أن في عملية توليد البيانات تم جعل المتغير Mمتغير توسطي من خلال إدخاله في السلسلة السببية بين X و y، وبوساطة كاملـة. أما في مقياس قوة الاختبـار فقد امتلك مقدر RMM قـوة أكثر من OLS وفي

كلتا الحالتين من نسبة التلويث ومستوى التأثير وفي كلا المستويين من الكفاءة النسبية.
أما مقياس التغطية فقد أظهر مقدر OLS تغطية أكثر من RMM ولكن هذا لا يُعَدُّ قصوراً على المقدر RMM إذ إنَّ وكما تم توضيح ذلك في تقييم حدود الثقة، أن مقياس السعة ومقياس التغطية هما مقياسان متعارضان ولكن من المفضل هو أن تكون سعة حدود الثقة أقل سعة من أجل الحصول على مقدر أكثر دقة (Krantz,1999). أما مقاييس التحيز والجذر التربيعي للتحيز والجذر التربيعي لمتوسط مربعات الخطأ فقد اظهر مقدر RMM أيضا تفوقه على OLS. اما مقياس الكفاءة فقد كان مقدر RMM أكثر كفاءة من مقدر OLS، كما أن الكفاءة زادت بشكل أكثر عند استخدام ثابت التوليف $a = 3.42$ بالنسبة إلى المقدر RMM الحصين.

اما الجدول (٢) أدناه وعند حجم العينة n=50 فان مقدر الحرف الحصين كان الأفضل مقارنة بالمربعات الصغرى، والنتائج في الجدول أدناه مماثلة إلى النتائج في الجول (١) مع فارق واحد في الجدول (١) إذ إنَّ الكفاءة النسبية لمقدر الحرف الحصين كانت اعتلى من المربعات الصغرى عند مستوى حجم التأثير 0.59، ولكن عند ثابت التوليف $a = 3.42$ كانت اقل من ثابت التوليف $a = 4.24$ ولكن هذا لا يعني أن مقدر OLS أفضل من مقدر RMM وذلك لان في جميع الحالات كانت الكفاءة النسبية للمقدر RMM أعلى من المقدر OLS.

جدول رقم (١)

حجم العينة n=25

	على مستوى التلويث 40% عند ثابت التوليف Tuning constant=4.24				على مستوى التلويث 40% عند ثابت التوليف Tuning constant=3.42			
Method	OLS		RMM		OLS		RMM	
Boundary	Lower	Upper	Lower	Upper	Lower	Upper	Lower	Upper
Conf.Interv.(CI)	-332.8	304.35	0.136	0.511	-553.1	500.20	0.110	0.539
Width of CI	637.2		0.375		1053.3		0.428	
Coverage	0.994		0.403		0.9945		0.394	
Power	0.005		0.730		0.0055		0.747	
Bias	14.386		-0.194		26.628		-0.195	
RootMSE($\hat{\beta}$)	419.39		0.2991		1448.81		0.301	
RootMSE	3612.78		0.888		13251.8		4.714	
Efficient	1378.108				4931.295			
	على مستوى التأثير 0.59 عند ثابت التوليف Tuning constant=4.24				على مستوى التأثير 0.59 عند ثابت التوليف Tuning constant=3.42			
Method	OLS		RMM		OLS		RMM	
Boundary	Lower	Upper	Lower	Upper	Lower	Upper	Lower	Upper
Conf.Interv.(CI)	-260.4	254.2	0.206	0.797	-246.6	257.8	0.219	0.777
Width of CI	514.741		0.591		504.49		0.558	
Coverage	0.971		0.628		0.974		0.609	
Power	0.097		0.772		0.102		0.783	
Bias	3.422		-0.153		-5.281		-0.150	
RootMSE($\hat{\beta}$)	616.40		0.279		544.95		0.282	
RootMSE	7909.80		1.3431		7399.9		0.830	
Efficient	1930.367				1706.489			

جدول رقم (٢)

حجم العينة n=50

	على مستوى التلويث 40% عند ثابت التوليف Tuning constant=4.24				على مستوى التلويث 40% عند ثابت التوليف Tuning constant=3.42			
Method	OLS		RMM		OLS		RMM	
Boundary	Lower	Upper	Lower	Upper	Lower	Upper	Lower	Upper
Conf. Interv.(CI)	-1168.4	924.3	0.168	0.485	-6197.8	3186.8	0.181	0.481
Width of CI	2092.815		0.354		9384.69		0.299	
Coverage	0.997		0.354		0.997		0.320	
Power	0.002		0.766		0.002		0.782	
Bias	122.14		-0.197		1505.5		-0.201	
RootMSE($\hat{\beta}$)	5698.35		0.267		53827.3		0.268	
RootMSE	89522.57		0.855		408384.6		0.808	
Efficient	21673.98				168418.7			
	على مستوى التأثير 0.59 عند ثابت التوليف Tuning constant=4.24				على مستوى التأثير 0.59 عند ثابت التوليف Tuning constant=3.42			
Method	OLS		RMM		OLS		RMM	
Boundary	Lower	Upper	Lower	Upper	Lower	Upper	Lower	Upper
Conf. Interv.(CI)	-131.6	114.7	0.298	0.715	-169.2	144.78	0.304	0.708
Width of CI	246.360		0.417		313.99		0.404	
Coverage	0.977		0.550		0.976		0.522	
Power	0.231		0.956		0.237		0.944	
Bias	8.796		-0.158		12.55		-0.158	
RootMSE($\hat{\beta}$)	221.82		0.240		379.0		0.243	
RootMSE	1933.6		0.838		3379.6		0.7802	
Efficient	767.5192				1277.914			

جدول رقم (٣)

حجم العينة n=100

	على مستوى التلويث 40% عند ثابت التوليف Tuning constant=4.24				على مستوى التلويث 40% عند ثابت التوليف Tuning constant=3.42			
Method	OLS		RMM		OLS		RMM	
Boundary	Lower	Upper	Lower	Upper	Lower	Upper	Lower	Upper
Conf. Interv.(CI)	-156.87	155.72	0.238	0.372	-151.36	136.05	0.240	0.3708
Width of CI	312.60		0.134		287.42		0.130	
Coverage	0.9435		0.209		0.943		0.198	
Power	0.076		0.824		0.072		0.832	
Bias	0.705		-0.175		7.783		-0.175	
RootMSE($\hat{\beta}$)	270.58		0.244		336.78		0.245	
RootMSE	26132.2		0.842		12188.26		0.796	
Efficient	1460.536				1998.047			
	على مستوى التأثير 0.59 عند ثابت التوليف Tuning constant=4.24				على مستوى التأثير 0.59 عند ثابت التوليف Tuning constant=3.42			
Method	OLS		RMM		OLS		RMM	
Boundary	Lower	Upper	Lower	Upper	Lower	Upper	Lower	Upper
Conf. Interv.(CI)	-67.47	72.94	0.368	0.639	-67.47	72.94	0.369	0.634
Width of CI	140.42		0.270		140.42		0.265	
Coverage	0.957		0.448		0.957		0.436	
Power	0.287		0.999		0.287		0.998	
Bias	-2.384		-0.155		-2.38		-0.154	
RootMSE($\hat{\beta}$)	103.55		0.218		103.55		0.219	
RootMSE	3460.22		0.835		3460.2		0.778	
Efficient	395.62				395.40			

جدول رقم (٤) حجم العينة n=100								
Method	OLS		RMM		OLS		RMM	
Boundary	Lower	Upper	Lower	Upper	Lower	Upper	Lower	Upper
Conf. Interv.(CI)	-200.39	178.25	0.2233	0.360	-246.52	293.02	0.229	0.3591
Width of CI	378.65		0.137		539.54		0.129	
Coverage	0.978		0.169		0.978		0.165	
Power	0.022		0.844		0.021		0.848	
Bias	11.19		-0.161		-23.12		-0.164	
RootMSE($\hat{\beta}$)	271.07		0.186		713.63		0.189	
RootMSE	37585.8		0.849		53609.7		0.794	
Efficient	1123.60				2791.3			
	على مستوى التأثير 0.59 عند ثابت التوليف Tuning constant=4.24				على مستوى التأثير 0.59 عند ثابت التوليف Tuning constant=3.42			
Method	OLS		RMM		OLS		RMM	
Boundary	Lower	Upper	Lower	Upper	Lower	Upper	Lower	Upper
Conf. Enter.(CI)	728.89	1015.9	1737.5	1743.9	733.80	1072.9	1737.4	1744.1
Width of CI	287.04		6.458		339.166		6.661	
Coverage	0.737		0.434		0.737		0.444	
Power	0.502		1		0.501		1	
Bias	868.260		-0.091		837.28		-0.098	
RootMSE($\hat{\beta}$)	1115.69		2,151		1365,69		2,152	
RootMSE	9769,15		0,839		18719,5		0,838	
Efficient	1219,023				2155,613			

أما الجدول (٣) أعلاه وعند حجم العينة n=100 كانت نتائجه مماثلة إلى نتائج الجدول (١) وحتى من حيث الكفاءة النسبية للمقدرين مع تساوي نسبة الكفاءة تقريباً عند مستوى حجم التأثير عند ثابتي التوليف $a=3.42$ و $a=4.24$.

والجدول (٤) أعلاه والذي يمثل مقارنة المقدرين عند حجم العينة n=500 فقد كانت النتائج تبين تفوق مقدر RMM على مقدر OLS ولجميع مقاييس التقييم. كما يوضح الجدول زيادة الكفاءة النسبية عند ثابت التوليف $a=3.42$ أكثر مما هو الحال عند ثابت التوليف $a=4.24$

جدول رقم (٥) متوسط جداول دراسة المحاكاة								
	على مستوى التلويث 40% عند ثابت التوليف Tuning constant=4.24				على مستوى التلويث 40% عند ثابت التوليف Tuning constant=3.42			
Method	OLS		RMM		OLS		RMM	
Boundary	Lower	Upper	Lower	Upper	Lower	Upper	Lower	Upper
Conf. Interv.(CI)	-111.61	204.38	108.67	109.3	-461.89	379.83	108.6	109.32
Width of CI	315.997		0.647564		841.735		0.657979	
Coverage	0.960375		0.617669		0.960094		0.604016	
Power	0.163083		0.722529		0.164813		0.726136	
Bias	62.51465		-0.09435		149.9294		-0.09459	
RootMSE($\hat{\beta}$)	574.308		0.28905		3872.049		0.290767	
RootMSE	12274.18		0.8734		34695.82		1.041216	
Efficient	2878.192				13328.83			

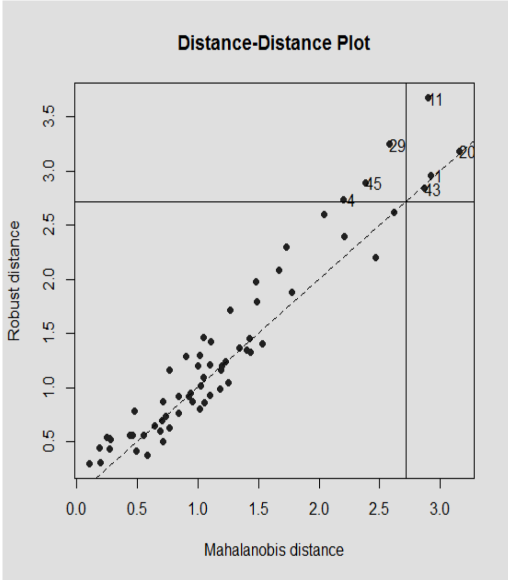
ولتوضيح دراسة المحاكاة هذه بشكل إجمالي يمكن ملاحظة الجدول(٥) أدناه، إذ إنَّ هذا الجدول يمثل متوسط الجداول ولجميع التراكيب ويُعد بمثابة توضيح إجمالي لعملية المحاكاة ونتائجها. إذ يلاحظ من هذا الجدول أن مقدر RMM أفضل من مقدر OLS ولجميع مقاييس التقييم وحتى بالنسبة إلى ثابت التوليف $a = 3.42$ مقارنةً بثابت التوليف $a = 4.24$ مما يدل على زيادة كفاءة المقدر المقترح بشكل اكبر. كما أن الجدول يوضح حدود الثقة لمقدر الحرف الحصين كانت لا تتضمن الصفر على عكس المربعات الصغرى وعند كلا ثابتي التوليف $a = 3.42$ و $a = 4.24$ مع أن في عملية توليد البيانات تم جعل المتغير M متغير توسطي من خلال إدخاله في السلسلة السببية بين X و y، وبوساطة كاملة. فضلاً عن ذلك أظهر مقدر المربعات الصغرى انهُ يمتلك تغطية اكبر في حدود الثقة مقارنةً بمقدر الحرف إلا أن هذا وكما تم ذكر ذلك مسبقاً لا يعتبر قصوراً لمقدر RMM إذ إنَّ مقياس السعة ومقياس التغطية هما مقياسان متعارضان ولكن من المفضل هو أن تكون سعة حدود الثقة أقل سعة من أجل الحصول على مقدر أكثر دقة (Krantz, 1999).

يتضح مما سبق أن مقدر الحرف الحصين قد اثبتت كفاءتهُ مقارنةً بالمربعات عند استخدامهُ في تحليل الوساطة ولكافة مقاييس التقييم، إذ إنَّ هذا المقدر عالج القيم الشاذة الموجودة في المتغيرات من خلال استخدام مقدرات MM الحصينة، كما انهُ عالج مشكلة التداخل الخطي الموجود بين المتغيرات والتي تكون هذه المشكلة موجودة ضمنياً في نماذج تحليل الوساطة والسبب يعود إلى تصميم النماذج التي تعتمد عليها نماذج الوساطة في تقدير التأثير التوسطي.

ب- تحليل بيانات الدراسة الحقيقية

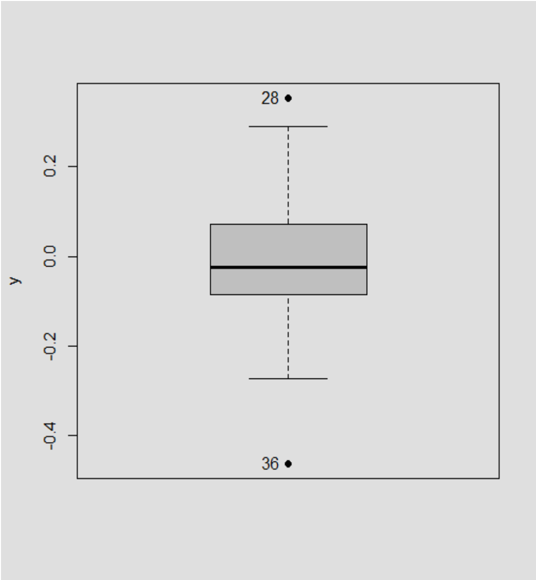
تمثل الجانب العملي الثاني في استخدام بيانات حقيقية تم جمعها من مستشفى السلام بالموصل. إذ تم جمع بيانات عن ثلاثة متغيرات، إذ إنَّ المتغير الأول والذي هو المتغير X كان يمثل العمر (Age) والمتغير الثاني والذي هو المتغير التوسطي M كان يمثل الوزن (weight) والمتغير الثالث والذي هو المتغير التابع y كان يمثل ضغط الدم الانقباضي systolic blood pressure (SBP). إذ تم جمع هذه البيانات من 71 شخص.

وقد تم الكشف عن القيم الشاذة باستخدام الرسم الصندوقي (Boxplot) للمتغير التابع y وقد ظهر أن هناك قيمتان متطرفتان وهما 28 و 36 في هذا المتغير وكما موضح بالشكل (٤). كما تم إجراء اختبار باستخدام MVE (Minimum Volume Ellipsoid estimator) والموضح بالشكل (٥) إذ نلاحظ أن هناك قيم قوة رافعة (Leverage Points) في مصفوفة المتغيرين المستقل والتوسطي وهي (4,11,20,29,43,45)، لاحظ الشكل(٥).



شكل رقم (٥)

يوضح قيم قوة الرفع في مصفوفة المتغيرات التوضيحية X



شكل رقم(٤)

رسم الـ boxplot للمتغير

أما التداخل الخطي فقد تم إجراء اختبار t على معامل الارتباط بين المتغيرين X و M وكانت قيمة $p = 0.000$ ، وبمقارنتها مع قيمة $\alpha = 0.05$ يلاحظ إنها أقل، أي إنَّ الاختبار معنوي مما يدل على إنَّ هناك ارتباط معنوي بين المتغيرين.

كما تم استخدام أسلوب العدد الشرطي الموضح في المعادلة (27) وكانت نتيجة العدد الشرطي $C.N. = 4.21$ وهو اكبر من الواحد الصحيح مما يدل على وجود تداخل خطي بين المتغيرين. كما انهُ يعتبر من التداخل الخطي الضعيف (Lawrence & Arthur, 1990)، وهذا يكون أكثر خطورة وذلك لضعف الارتباط بينما هو في الحقيقة يؤثر على المقدرات.

وأخيرا تم استخدام اختبار Farrar-Glauber الموضح في المعادلة (28) ووجد أن قيمة $p = 0.000$ ، وبمقارنتها مع قيمة $\alpha = 0.05$، يلاحظ إنها أقل مما يعني رفض العدم وقبول البديلة أي أن المتغيرات ليست متعامدة (غير مستقلة). وقد كانت الفرضية المستخدمة في اختبار معنوية التأثير التوسطي:

$$H_0 : ab = 0 \;,\; H_1 : ab \neq 0$$

وباستخدام أسلوب RMM وكذلك المربعات الصغرى في إيجاد حدود الثقة للتأثير التوسطي، فضلاً عن إيجاد سعة حدي الثقة للأسلوبين ومتوسط مربعات الخطأ والكفاءة من خلال قسمة جذر متوسط مربعات الخطأ للمربعات على جذر متوسط مربعات الخطأ لأسلوب RMM، فان النتائج الموضحة في الجدول (٦) كانت تشير إلى مدى تفوق أسلوب RMM عن أسلوب المربعات الصغرى في اختبار التأثير التوسطي. كما يمكن ملاحظة أن حدود الثقة لأسلوب RMM كانت لا تتضمن الصفر في حدودها مما يدل على معنوية التأثير الغير مباشر أي ان

المتغير التوسطي والمتمثل بالوزن هو يتوسط العلاقة فعلاً بين المتغير التنبؤي والمتمثل بالعمر والمتغير التابع والمتمثل بضغط الدم وهذا يتطابق مع الناحية الطبية، على عكس حدود الثقة لأسلوب المربعات الصغرى والذي تضمن الصفر مما يدلى على عدم معنوية التأثير التوسطي وهذا مقياس جيد على تفوق أسلوب RMM كونهُ قد فسر السلسلة السببية من المتغير التوضيحي إلى المتغير التوسطي إلى المتغير التابع.

جدول رقم (٦)

جدول يوضح نتائج دراسة بيانات مستشفى السلام في الموصل

	على مستوى التلويث 40% عند ثابت التوليف Tuning constant=4.24				على مستوى التلويث 40% عند ثابت التوليف Tuning constant=3.42			
Method	OLS		RMM		OLS		RMM	
Boundary	Lower	Upper	Lower	Upper	Lower	Upper	Lower	Upper
Conf. Interv.(CI)	-0.068	0.277	0.019	0.209	-0.068	0.277	0.040	0.223
Width of CI	0.345		0.189		0.345		0.1827	
RootMSE	0.110		0.073		0.110		0.065	
Efficient	2.268				2.848			

الاستنتاجات والتوصيات

الاستنتاجات:

من خلال ما تقدم يلاحظ ما يلي:

١- إنَّ أسلوب RMM قد تفوق على أسلوب المربعات الصغرى وفي جميع حالات التلويث وجميع حالات أحجام التأثير ولمختلف أحجام العينات.

٢- إنَّ عملية تقليل ثابت التوليف a في المرحلة الثالثة للمقدر MM من 4.24 إلى 3.42 أدى إلى زيادة الكفاءة النسبية للأسلوب المقترح RMM والذي أدى بدورِهِ إلى تقدير التأثير التوسطي بشكل أدق في عملية الوساطة.

٣- إنَّ نتائج هذا البحث أثبتت أن أسلوب التوزيع المحاذي إلى التوزيع الطبيعي يمكن الاعتماد عليه في حالة معالجة مشاكل التحليل التوسطي والمتمثلة بالقيم الشاذة والتداخل الخطي، وليس كما تشير معظم البحوث ولاسيما البحوث المتعلقة بعلم النفس والتي تناولت تحليل الوساطة بشكل واسع، إذ تشير هذه البحوث إلى ضعف أسلوب التوزيع المحاذي المقترح من قبل (Sobel, 1982) كونه يتطلب أحجام عينات كبيرة، ولكن كما هو واضح

فان هذا الأسلوب المحاذي قد اثبت كفاءة حتى مع أحجام العينات الصغيرة؛ وذلك لأنه تم معالجة الانتهاكات التي حدثت في فروض النموذج الخطي.

التوصيات:

إنَّ التوصيات التي يوصى بها في هذا البحث هي كما يلي:

١- الاهتمام بمسألة وجود التداخل الخطي في نماذج التحليل التوسطي والتي تؤدي إلى نتائج غير دقيقة.

٢- الاهتمام بالتحقق من وجود القيم الشاذة في البيانات ومعالجة تأثيراتها بمقدرات حصينة ذات نقطة انهيار عالية وكفاءة نسبية عالية أيضا مثل مقدرات MM وبثابت توليف $a = 3.42$.

٣- استخدام أسلوب التوزيع المحاذي إلى التوزيع الطبيعي في تحليل الوساطة ولكن بعد معالجة المشاكل التي تم ذكرها في التوصيتين المذكورتين أعلاه.

المصـــادر

اولا: العربية

١- كاظم، اموري هادي.٢٠٠٢. القياس الاقتصادي المتقدم، النظرية والتطبيق. المكتبة الوطنية.

٢- المشهداني، محمود حسن و هرمز، امير حنا.(١٩٨٩)، الإحصاء. جامعة بغداد.

٣- لقاء علي محمد وأفراح كاظم جويد،٢٠١٠. مقارنة المقدرات الحصينة في أسلوب التحليل العاملي. المجلة العراقية للعلوم الإحصائية. [ص٢٠٧-٢٢٦].

ثانيا: الاجنبية

4- Alwin, D. F., & Hauser, R. M. (1975). The decomposition of effects in path analysis. American Sociological Review, 40, 37–47.

5- Askin R G and Montgomery D. C.(1980). Augmented robust estimators Technometrics.; 22: 333-341.

6- Beasley T. Mar.(2012). Power of Product Tests of ediation as a Function of Mediator Collinearity. Multiple Linear egression Viewpoints. Vol. 38(2).

7- Belsley, D., Kuh, E., and Welsh, R. E., (1980). Regression Diagnostics. Wiley, New york.

8- Bentler, P. M. (1997). EQS for Windows (Version 5.6) [Computer software]. Encino, CA: Multivariate Software.

9- Bianco, A. M., M. Garcia Ben and V. J. Yohai (2005). 'Robust estimation for linear regression with asymmetric errors.' The Canadian Journal of Statistics, Vol. 33, No. 4, pp. 511-528.

10- Carter, N. J., Schwertman, N. C., and T. L. Kiser.(2009) A comparison of two boxplot methods for detecting univariate outliers which adjust for sample size and asymmetry. *Statistical Methodology*, 6(6):604–621.
11- Emily E Ricotta, Marc Boulay, Robert Ainslie, Stella Babalola, Megan Fotheringham, Hannah Koenker and Matthew Lynch1.(2015). The use of mediation analysis to assess the effects of a behaviour change communication strategy on bed net ideation and household universal coverage in Tanzania .Malaria Journal 14:15.
12- Fairchild, A.J., & MacKinnon, D.P. (2009). A General Model for Testing Mediation and Moderation Effects. *Prevention Science, 10*, 87-99. doi:10.1007/s11121-008-0109-6; URL.
13- Fisher, J. C. and Mason, R. L.,(1981). The Analysis of Multicollinear data in criminology, Acadmic Press, Inc.
14- Folmer, H. (1981). Measurement of the effects of regional policy instruments by means of linear structural equation models and panel data. Environment and Planning A, 13, 1435–1448.
15- Freedman, L. S., & Schatzkin, A. (1992). Sample size for studying intermediate endpoints within intervention trials or observational studies. American Journal of Epidemiology, 136, 1148–1159.
16- Guralnik, D. B. (Ed.) (1970). Webster’s new world dictionary of the American language (2nd ed.). New York: The World Publishing Company.
17- Hadi, A. S. and J. S. Simonoff (1993). 'Procedures for the Identi_cation of Multiple Outliers in Linear Models.' Journal of the American Statistical Association, Vol. 88, No. 424, pp.1264-1272.
18- Harlow, L. L., Mulaik, S. A., & Steiger, J. H. (1997). What if there were no significance tests? Mahwah, NJ: Lawrence Erlbaum Associates.
19- Hoerl, A. E. & Kennard, R.W. (1970). Ridge regression: Biased estimation for nonorthogonal problems. Technometric. 12(1):55-67.
20- Hoerl, A. E. and K. F. Baldwin.(1975). Ridge Regression: Some Simulations. Communications in Statistics.; (4): 104-123.
21- Hoyle, R. H., & Kenny, D. A. (1999). Sample size, reliability, and tests of statistical mediation. In R. H. Hoyle (Ed.), Statistical strategies for small sample research . 195–222. Thousand Oaks, CA: Sage.
22- Joreskog, K. G., & Sorbom, D. (2001). LISREL (Version 8.5) [Computer software]. Chicago: Scientifi c Software International.
23- Judd, C. M., & Kenny, D. A. (1981a).Estimating the effects of social interventions. CAMBRIDGE UNIVERSITY PRESS.
24- Judd, C. M., & Kenny, D. A. (1981b). Process analysis: Estimating mediation in treatment evaluations. Evaluation Review, 5, 602–619.

25- Krantz, D. H. .(1999). The null hypothesis testing controversy in psychology. Journal of the American Statistical Association, 44, 1372–1381.
26- Lawrence D. Kenneth and Arthur L. Jeffrey.(1990).Robust Regression Analysis and Applications. Marcel Dekker, INC.
27- Mackinnon P. David and Fairchild J. Amanda.(2009). Current Directions in Mediation Analysis, Association for Psychological Science Volume 18-Number 1.
28- Mackinnon P. David.(2008). Introduction to Statistical Mediation Analysis. Taylor &Francis Group, LLC.
29- MacKinnon, D. P., Lockwood C. M., Hoffman, J. M., West, S. G., & Sheets, V.(2002). A comparison of methods to test mediation and other intervening variable effects. Psychological Methods, 7, 83–104.
30- MacKinnon, D. P., Lockwood, C. M., & Williams, J. (2004). Confidence limits for the indirect effect: Distribution of the product and resampling methods. *Multivariate Behavioral Research*, *39*, 99–128.
31- Mansfeld, E. R. ,Webster, J. T. and Gunst, R. F.,(1977). An Analytic Variable Selection Techniques for principle component Regression. Applied statistics, Vol. 26, No. 1. 34-40.
32- Olive, D. J. and D. M. Hawkins (2008). 'The Breakdown of Breakdown' [Online]. URL http://www.math.siu.edu/olive/ppbdbd.pdf .
33- Pati K. Dano, Robiah Adnan and Rasheed B. Abdulkadir .(2014). Using Ridge Least Median Squares to Estimate the Parameter by Solving Multicollinearity and Outliers Problems. Nature and Science;12(11).
34- Peng Ding and Tyler J. Vander Weele.(2016). Sharp sensitivity bounds for mediation under unmeasured mediator-outcome confounding.math.st.20. [Online]. http://arxiv.org/abs/1601.05155v1
35- Pfaffenberger, R.C. and Dielman, T.E. 1990. Acomparison of regression estimators when both multicollinearity and outliers are present. Edit by: Arthur, Marcel Dekker, Inc. New York and Basel.
36- Preacher J. Kristopher and James P. Selig.(2012). Advantages of Monte Carlo Confidence Intervals for Indirect Effects. Communication Methods and Measures, 6:77–98.
37- Rousseeuw, P. J. (1983), Multivariate Estimation With High Breakdown Point, paper presented at Fourth Pannonian Symposium on Mathematical Statistics and Probability.
38- Rousseeuw, P. J. and V. J. Yohai (1984). 'Robust regression by means of S-estimators.' Robust and Nonlinear Time Series Analysis, eds. J. Franke, W. Härdel, and D. Martin. New York: Springer-Verlag, pp. 256-272.
39- Rousseew P.J., and A.M. Leroy.(1987). Robust Regression and Outlier Detection. Wiley- Interscience, New York.

40- Samkar, Hatice and Alpu, Ozlem.2010.Ridge Regression Based on some Roubest Estimators. Journal of Modern Applied Statistical Methods. Vol. 9, No. 2, 495-501.
41- Sobel, M. E. (1982). Asymptotic confidence intervals for indirect effects in structural equation models. Sociological Methodology, 13, 290–312.
42- Stuart,Catherine.(2011).RobustRegressio.(htt://www.maths.dur.ac.uk/ug/projects/highlights/cm3/stuart Robust Regression report.pdf).
43- Tofughi, D. , Mackinnon, P. David and Yoon M.,(2009). Covariances between regression coefficient estimates in a single mediator model. British Journal of Mathematical and Statistical Psychology, 62, 457–484.
44- Yohai, V.J. (1987). .High Breakdown-point and High Efficiency Estimates for Regression., The Annals of Statistics 15, 642.65.
45- Zahari, Siti Meriam Mohammad said Zainol and Muhammad Iqbal al-Banna bin Ismail. (2012) Weighted Ridge MM-Estimator in Robust Ridge Regression with Multicollinearity. Mathematical Models and Methods in Modern Science; ISBN:978-1-61804-106-7.
46- Zhang, Z., & Wang, L. (2008). Methods for evaluating mediation effects: Rationale and comparison., New trends in psychometrics (pp. 595–604). Tokyo, Japan: Universal Academy Press.