تحليل قياسي تطبيقي لآثار الازدواج الخطى المتعدد على قيم ومعنوية معاملات نموذج الانحدار الخطى المتعدد(١)

An Applied Measurable Analysis for the consequences of multicollinearity On Multivariate Regression coefficients

أ.د/كمال سلطان محمد سالم
عميد المعهد العالي للحاسب الآلي
كينج مريوط- الإسكندرية

د/محمد خلف عبد العال رفاعي
مدرس الإحصاء التطبيقي
معهد العجمي العالي للعلوم الإدارية- الإسكندرية

(١)يقدم الباحثان خالص التقدير للاساتذة المحكمين على كافة الملاحظات التى أشارو إليهاوالتى ساهمت فى أثراء هذا البحث

تحليل قياسي تطبيقي لآثار الازدواج الخطى

المتعدد على قيم ومعنوية معاملات نموذج الانحدار الخطى المتعدد(٢)

An Applied Measurable Analysis for the consequences of multicollinearity On Multivariate Regression coefficients

أ.د/كمال سلطان محمد سالم — د/محمد خلف عبد العال رفاعى

عميد المعهد العالى للحاسب الآلى — مدرس الاحصاء التطبيقى

كينج مريوط- الاسكندرية — معهد العجمى العالى للعلوم الادارية- الاسكندرية

ملخص البحث:

يتناول البحث موضوع تحليل قياسى تطبيقى لآثار الازدواج الخطى المتعدد على قيم ومعنوية معاملات نموذج الانحدار الخطى المتعدد.وتتمثل مشكلة البحث فى قلة الدراسات القياسية التطبيقية فى مجال أثر الازدواج الخطى المتعدد على قيم ومعنوية وإشارات معاملات الانحدار وكذا فيما يتعلق بأثر استخدام طريقتى (PCR , RR).

والهدف من البحث يتمثل فى تحليل ودراسةآثار وجود مشكلة الازدواج الخطى المتعدد على قيم وإشارات معنوية معاملات الانحدار الخطى المقدرة بطريقة (O.L.S)،وتكمن **أهمية البحث** فى تحديد : أفضل الطرق القياسية للآثار السلبية لمشكلةالازدواج الخطى على المعاملات أما منهجية البحث تتمثل فى النقاط التالية:

-متغيرات البحث هى: المتغيرات المستقلةو هى { متوسط أصول الأسرة بالدولار (X_1) – العمر (X_2) – عدد سنوات التعليم (X_3) – (X_4) دخول مستحقة للأسرة غير محصلة – (X_5) متوسط أجر الساعة بالدولار) والمتغير التابع يتمثل فى متوسط عدد ساعات العمل خلال السنة(y)}.

يتمثل الفرض الرئيسى في صورته الصفريه بأن وجود الازدواج الخطى لا يؤثر على قيم ومعنوية وأشارة المتغيرات

التفسيرية فى علاقتها مع المتغير التابع.

أساليب علاج مشكلة الازدواج الخطى (١) طريقة ريدج (RR) (٢) طريقة المكونات الرئيسية للانحدار (PCR)و بعد إثبات صحة الفروض من عدمها قد توصل البحث للنتائج التالية:

---

(١)يقدم الباحثان خالص التقدير للاساتذة المحكمين على كافة الملاحظات التى أشارو إليهاوالتى ساهمت فى أثراء هذا البحث

-وجد أن المتغيرات (X_1, X_3, X_2) غير معنوية فى نموذج الانحدار الخطى المتعدد المقدر بـ (LS)وفى النموذج الفردى يوجد متغيرات غير معنوية أهمها (X_4 , X_3)

-يفضل النموذج (PCR) عن نظيره (RR) **كلمات مفتاحية:** الإزدواج الخطى- طريقة المربعات الصغرى- معامل تضخم التباين- طريقة ريدج-طريقة انحدار المكونات الرئيسية

Abstract:

In this The research .An applied measurable Analysis for the consequences of multicollinearity On Multivariate Regression coefficients has been studied . The research problem represents the Few studies of applied econometrics in the field of multicollinearity on values, significant and signs of regression coefficient. Also the relation to The effect of method PCR, RR.

Objective: The research aims to the following:

- The consequences of multicollinearity on values, signs and significant of coefficients of linear regression estimated by(O.L.S.)

- The best econometrics methods using to solve the negative effects for the problem of multicollinearity.

Research Methodology:

The research variables are (independent and depends variables).

- the average of family assets in dollars (X_1), age (X_2), years of education (X_3), accrued incomes (X_4), average per hour (X_5) and dependent variable is the average labor hours during the year(y).

The Hypothesis of research:

The Null Hypothesis H_0: multicollinearity does not affect the significant, sign and value of the regression coefficients (B_i) for independent variables.

The Alternative Hypothesis H_1: multicollinearity affects the significant, sign and value of the regression coefficient(B_i) for independent variables.

The Remedies For multicollinearity are the method of (RR) and the method of (PCR).

The method of(PCR), the researchers reach the conclusion after proving the right of hypothesis or net.

- Variables (X_2, X_3, X_4) are non–significant in the model estimated by(O.L.S)

- There are non-significant variables such as (X_3, X_4).

- Model(PCR) is better than (RR).

Keywords:Multicollinearity-(O.L.S)-(VIF)-Ridge Regression(RR)-Principal component Regression(PCR)

١.مقدمة Introduction

من فروض assumptions طريقة (LS) المستخدمة فى تقديرمعالم نموذج الانحدار الخطى العادى فرض عدم وجود ارتباط (كامل) أو قريب من الكامل بين المتغيرات التفسيرية فى هذا النموذج وعند عدم توفر هذا الفرض تكون هناك مشكلة الازدواج الخطى المتعدد ويعزى تعريف هذه المشكلة إلى Ragnar Frisch حيث عرفها بوجود علاقة خطية كاملة أو تامة بين بعض أو كل المتغيرات "التفسيرية " فى نموذج الانحدار ومن المنطقى أن يقتصر هذا التعريف على العلاقات الخطية بين المتغيرات X^S ولكنه لا يمتد إلى العلاقات غير الخطية كما فى دالة التكاليف من الدرجة الثالثة والنتائج الأساسية لتلك المشكلة هى إما عدم إمكانية تقدير معاملات الانحدار فى حالة وجود علاقات خطية كاملة بين المتغيرات التفسيرية وتكون الأخطاء المعيارية Standard errors لا نهائية، أما إذا كانت العلاقات غير كامله فإنه يمكن حساب معاملات الانحدار لها ولكن تكون الأخطاء القياسية لها كبيرة وبالتالى تكون معاملات الانحدار غير معنوية وعليه تفقد معناها أو دقتها وفى التطبيق الفعلى فى مجال تقدير دالات الإنتاج أو الواردات فإنه عادة ما تظهر نتائج تسبب كثيراً من سوء الاستنتاج وعدم المنطقية ومخالفة المنطق العلمى من تلك النتائج مثل:

أ. عدم معنوية معاملات الانحدار لغالبية المتغيرات التفسيرية فى هذه النماذج.

ب. مخالفه إشارات المعاملات المقدرة للمنطق النظرى.

٢.مشكلة البحث:Problem of Research

عنداستخدام طريقة المربعات الدنيا العاديه(O.L.S)فى تقدير معاملات المتغيرات التفسيرية لنموذج الانحدار

الخطى المتعدد فإن كل معامل يعكس العلاقة أو التأثير بين هذا المتغير والمتغير التابع وذلك عند تحقق عدم وجود ارتباط شبه تام بين هذا المتغير والمتغيرات التفسيرية الأخرى[3-2]سالم .

أما فى وجود هذا الارتباط المرتفع(أى وجود ظاهرة الازدواج الخطى) فإن ذلك يؤثر بشكل معنوى على تقديرات المعاملات المقدرة وبالتالى يصعب التمييز بين تأثير متغير معين وتأثير المتغيرات الأخرى ، فضلاً عن ذلك فإن الازدواج الخطى يرفع قيم تباين المعاملات وبالتالى قد يؤدى إلى إشارات غير متوقعة للمعاملات ويعزى ذلك إلى أن توزيع المعاملات (β^{s}) مع وجودالازدواج الخطى يكون اكثر

اتساعاً مقارنة مع عدم وجوده وبالتالى يزيد من إمكانية أن تقع قيمة معينه ($\hat{\beta}$) المقدرة فى الجانب الآخرمن القيمة الحقيقية لـ β وبالتالى تظهر إشارات مخالفة[33] Studenmund .

و خلاصة القول أن الازدواج الخطى يؤدى إلى:

أ-حدوث تغيير فى قيم ومعنوية بعض المعاملات المقدرة.

ب- حدوث إشارات غير متوقعة لبعض المعاملات.

وعند استخدام هذه المعاملات فى تفسير العلاقات بين كل متغير تفسيرى والمتغير التابع فإنه يؤدى إلى نتائج مضلله وغير متوافقة مع المنطق العلمى من ناحية ولا تعكس الواقع الحقيقى من ناحية أخرى.

وتكثر البحوث فى تعريف وقياس ونتائج وعلاج مشكلة الازدواج الخطى المتعدد وخاصة استخدام عدة طرق قياسية لعلاج مشكلة الازدواج الخطى المتعدد ومنها:

أ.طريقة انحدار ريدج Ridge Regression Method (RR) .

ب. طريقة انحدار المكونات الرئيسية Principal Component Regression (PCR)

ج. طريقة الانحدار الجزئى Partial Least Squares Regression (PLRS)

وتكون المفاضلة بينها فى ضوء عدة معايير منها:

قيم R^2 و RMSR (Root Mean Squares Error) وتكون المفاضلة على أساس هذه القيم فقط بدون تحليل أثر استخدام هذه الطرق على قيم وإشارات ومعنوية معاملات الانحدار فى الطرق المختلفة.

ولذلك فإن مشكلة البحث تتركز فى ندرة الدراسات القياسية التطبيقية فى مجال أثر الازدواج الخطى المتعدد على قيم ومعنوية وإشارات معاملات الانحدار وكذا فيما يتعلق بأثر استخدام طريقتى(PCR, RR)على هذه النتائج.

٣. هدف البحث: Objective of Research

ينصب هدف البحث فى الإجابة على سؤالين هما

أ.ما هى آثار ونتائج وجود مشكلة الازدواج الخطى المتعدد على قيم وإشارات ومعنوية معاملات الانحدار الخطى المقدرة بطريقة (O.L.S).

ب. ما هى أفضل الطرق القياسية لعلاج أوالتغلب أوتقليل الآثار السلبية لمشكلة الازدواج الخطى على هذه المعاملات؟

٤. أهمية البحث: Importance of Research

تعتبر مشكلة الازدواج الخطى من المشكلات التى تناولها العديد من الباحثين ولم يتم التوصل لحل نهائى حتى الآن، حيث توجد العديد من الطرق والمقدرات والتى يتفاوت أداؤها فى التعامل مع

المشكلة وتزداد عدم القدرة على التوصل إلى تقديرات دقيقة فى حالة كون تباينات وتغايرات مقدرات (O.L.S)كبيرتوفقاً لدرجة الازدواج الخطى كماتظهر حساسية الأخطاء المعيارية لمقدرات (O.L.S) لأى تغير بسيط فى البيانات مما يؤدى إلى نتائج مضللة عند تجاهل وجود مشكلة الازدواج الخطى أو عدم معالجتها بأسلوب فعال لذلك من الأهمية التعرض لهذه المشكلة وتقديم حلول وعلاج لها باستخدام طريقتى انحدار ريدج (RR) وانحدار المكونات الرئيسية (PCR)

من ثم تبرز أهمية الدراسة من خلال كونها تتبنى مدخلاً يستخدم لتقدير معالم النموذج بصورة أكثر دقة وبالتالى زيادة القدرة على التنبؤ و أجراء اختبارات الفروض بصورة سليمة.

٥.منهجية البحث:Methodology of Research

تتكون منهجية البحث من النقاط التالية:

٥-١):النموذج الرياضى الخطى المستخدم فى البحث: Multiple Linear Regression Model

هو نموذج الانحدار الخطى المتقدم وصورته العامة هى:

تأخذ الصيغة الدالية العامة لانحدار Y على X

$$Y_i = f(x_i) + \varepsilon_i , \qquad i = 1,2,\ldots,p$$

حيث ε_i هو حد الخطأ العشوائى، $f(x_i)$ تعنى أن Y دالة فى $(x_1, x_2, \ldots x_p)$ وسيتم افتراض أن الدالة خطية فى المعالم ،والمتغيرات المستقلة لا يشترط أن تكون خطية وإذا حصلنا على عينة من الحجم n، وكان المتغير التابع Y دالة فى المتغيرات المستقلة($x_1, x_2, \ldots x_p$) واعتبار أن مشاهدات x أى العمود الأول فى مصفوفة التصميم كلها تساوى واحد حتى يسمح بوجود جزء ثابت intercept فإنه يمكن كتابه نموذج الانحدار الخطى المتعدد على النحو التالى:

$$Y = \underbrace{\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \ldots\ldots \beta_p x_p}_{\text{متوسط Y عند وجود قيم معنية للمتغيرات } x_1, x_2, x_3, \ldots\ldots x_n} + U \quad \text{(حد الخطأ العشوائى)} \qquad (1)$$

وهذه المعادلة هى عبارة عن نظام به n معادلة يمكن التعبير عنها باستخدام صيغة المصفوفات والمتجهات Marix form كما يلى:

$$Y_{(n\times 1)} = X_{[n\times(p+1)]} + \beta_{[(P+1)\times 1]} + \varepsilon_{(n\times 1)} \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (2)$$

حيث y :متجه المتغير التابع وهو متجه عمود أبعاده $(n\times 1)$، x مصفوفة التصميم design matrixذات أبعاد

$[n\times(p+1)]$ ، β هو متجة المعاملات ذو أبعاد$[(p+1)\times 1]$

وأخيراً ε: هو متجه الخطأ العشوائى حيث أبعادة $(n\times1)$ ويتم تقدير معلمات النموذج بأستخدام طريقة المربعات الصغرى العامة (O.L.S) بافتراض توافر بعض الافتراضات المتعلقة بالنموذج وهى: العلاقة بين المتغير التابع والمتغيرات المستقلة يفترض أن تكون خطية, والنموذج تم توصيفه بشكل سليم,والمتغيرات التفسيرية مستقلة، ومتوسط الخطأ العشوائى صفر, وتباينه ثابت ويساوى $\sigma 2$ والمتغيرات التفسيرية مستقلة بين الخطأ العشوائى وبافتراض أن المتغيرات التفسيرية غير مرتبطة خطياً بأن

$$\hat{\beta} = (x'x)^{-1}\ x/y \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(3)$$

حيث $\hat{\beta}$ هو متجه مقدرات معالم النموذج باستخدام طريقة (O.L.S) وهو أفضل مقدر خطى غير متحير (Best Liner Unbiased Estimator) BLUE والنموذج المقدر يأخذ الشكل التالى:

$$\hat{Y} = X\ \hat{\beta}$$

وبافتراض اعتدالية الأخطاء فإن $\hat{\beta}$ لها توزيع طبيعى

بالخصائص التالية:

$$\hat{\beta} \approx N(\beta, {}^{2}\sigma\ (x'x)^{-1}) \ldots\ldots\ldots\ldots\ldots\ldots\ldots (4)$$

-(٥-٢):البيانات المستخدمة فى البحث: Data of Research

وقد أستند البحث على بيانات سنوية لعدد٣٥ مجموعة عمل(وهى عدد الملاحظات والتى تم إعدادهامن قبل مؤسسة

راندالامريكيةلدراسة العلاقة بين عرض العمل (والذى يمثلة متوسط عدد ساعات العمل) وزيادة فى الأجوروقد تم نشر هذة البيانات فى كتاب الأقتصادالقياسى(Basic Econometrics) فى صفحة ٣٨٥ وهو وارد فى المرجع رقم [16] Gujarati وتتضمن عدة متغيرات تم اختيار خمسة متغيرات مستقلة منها وذلك(لمحدودية عدد الملاحظات وهى ٣٥ ملاحظه) ومتغير تابع وهى:[متوسط أصول الأسرة (X_1) – متوسط عمر العامل (X_2) – متوسط عدد سنوات الدراسة المنتهية (X_3) – متوسط الدخول المستحقة للأسرة والغير محصلة (X_4) – متوسط معدل الأجر بالدولار (X_5) – متوسط عدد ساعات العمل خلال السنة (Y)] .

-(٥-٣):فروض البحث Hypothesis of Research

يقوم هذا البحث على فرض رئيسى مؤادة فى صورته الصفرية لاتؤثروجود مشكلة الازدواج الخطى المتعدد على معنوية وقيم وإشارة معاملات هذه المتغيرات الخمسة التفسيرية وتأثيرهاعلى المتغير التابع

٦. الدراسات السابقة Lecture Review

يمكن تصنيف الدراسات السابقة إلى عدة مجموعات منها فيما يتعلق

بأساليب اكتشاف درجة الازدواج الخطى:Detection of Multicollinearity

كلما اقترب معامل التحديد R^2 من الواحد الصحيح للنموذج ككل فإن هذا يدل على وجود مشكلة ازدواج خطى فعال يؤدى إلى تضخم تباين مقدرات المربعات الصغرى بشكل غير مبرر حتى أنها تؤول إلى ما لا نهاية فى حالة ما إذا كان $R^2 = 1$ كما وردت أشارة فى البحوث Judge Woold [38] Hayashi [18] -[16] Gujarati-[15]

أو إذا كانت معاملات الارتباط البسيطة بين كل متغيرين من متغيرات الانحدار Xi , Xj أكبر من0.8 وقد أوضح [17] (Haistorisky) أن الازدواج الخطى يكون فعال عندما يكون الارتباط $r_{ij} > 0.8$

حيث أن:

ارتفاع قيم معاملات R^2بين كل متغير تفسيرى وباقى المتغيرات التفسيرية الأخرى.

بينماقد اقترح [22](Lawrence) أن الازدواج الخطى يكون فعال عندما $1)\frac{r_{ij}}{R}$ حيث أن r_{ij} يمثل معامل الارتباط البسيط بين المتغيرين x_i, x_j ويمثل R معامل الارتباط المتعدد بين المتغير التابع Y والمتغيرين المستقلينx_j، x_i .

وقد أشار [19] (Hocking) أن استخدام القيم والمتجهات المميزة Eigen values & Eigen vectors لمعرفة مدى فعالية الازدواج الخطى فيعتبر الازدواج الخطى فعال إذا كانت أحد الجذور الكامنة latent root للمصفوفة $(x'x)$ يقترب من الصفر.

$$\lambda_i = y_i' x' x y_i \approx 0$$

حيث أن λ_i : أحد الجذور الكامنة Latent root

Y_i : أحد المتجهات الكامنة latent vector

بينما بين [7](Belsley) أن رقم الحالة condition number ودليل الحالة index condition يستخدم كلاً منهما لتحديد درجة الازدواج الخطى حيث أن:

$$\text{رقم الحالة (K)} = \frac{\text{أقصى قيمة للقيم المميزة}}{\text{أدنى قيمة للقيم المميزة}}$$

$$\text{دليل الحالة (CI)} = \sqrt{K}$$

فإذا اتضح أن:

$100 \le K \le 1000$فتوجد مشكلة ازدواج خطى من معتدل إلى قوى$K \ge 1000$ فتوجد مشكلة ازدواج خطى حادة

$10 \le CI \le 30$فتوجد مشكلة ازدواج خطى من معتدل إلى قوى $CI \ge 30$ فتوجد مشكلة ازدواج خطى حادة

وقدم (Schaefer et al)[32] مقياس يسمى معامل تضخم التباين Variance Inflation Factor(VIF)

$$VIF = \frac{1}{(1-R_i)^2} \quad \ldots\ldots\ldots\ldots\ldots (6)$$

$$\text{Tolerance (TOL)} = 1 - R_i^2$$

حيث: R_i^2 معامل التحديد لانحدار المتغير المستقل X_i على باقى المتغيرات المستقلة.

وقد حدد (David et al) [11] أنه يوجد ازدواج خطى إذا كان VIF> 10 أى أن R_i^2 تزيد من 0.9 .

أما الدراسات الخاصة بأثار ونتائج الازدواج الخطى: The consequences of Multicollinearity

فقد أشارت الدراسات التاليه [10] Chatelaine, [11] David, [16] (Gujarati), [17] Mishra [26] , Haistovisky Studenmund [33], Pasha [29],

إلى الآثار والنتائج التالية للإزدواج الخطى:

١. ستظل مقدرات المربعات الصغرى هى أفضل مقدرات خطية غير متحيزة (BLUE) إلا أن تبايناتها ستكون متضخمة.

٢. اتساع فترات الثقة.

٣. يزداد الخطأ المعيارى (SE) للمتغيرات التفسيرية ومن ثم تقل قيمة t المحسوبة وتزداد احتمالات قبول الفرض العدم $H0: \beta = 0$.

٤. تتغيرالمعالم المقدرة معنوياً بأية تغيرات طفيفة فى العينة.

٥. تأثر الخطأ المعيارى (SE) بأى تغيرات طفيفة فى البيانات مثل حذف أو إضافة مشاهدة وكذلك استقلال مشاهدة بمشاهدة أخرى.

٦. قد تختلف إشارة معاملات الانحدار على ما هو مقبول أو ثابت عليه فى النظرية الاقتصادية أو غيرها من الثوابت العملية الأمر الذى يسبب صعوبة فى تفسير معاملات الانحدار $(\beta's)$.

وفيمايتعلق بالدراسات الخاصة بالأساليب القياسية المستخدمة لعلاج مشكلة الازدواج الخطى Econometrics Methods used to Detection Multicollinearity

فقد قدم El–Dereny [14] , (Gujarati) [16] , Marquardt [25] , RobertLaurence[30], Studenmund[33].

أسلوب إسقاط المتغيرات Dropping avariables فعندما تواجه بمشكلة ازدواج خطى حاد فنجد أن أسهل الأساليب هو إسقاط المتغير أو المتغيرات التى تعانى من مشكلة الازدواج الخطى الحاد إلا أن إسقاط المتغير الذى يعانى من مشكلة الازدواج الخطى سيؤدى ذلك إلى الحصول على مقدرات متحيزة للمعالم وخطأ توصيف النموذج Specification error وهو أمر أشد خطورة من الازدواج الخطى حيث أن مقدرات(O.L.S) لا تزال (BLUE) حتى فى ظل الازدواج الخطى شبة التام ومن ثم تضعف قدرة النموذج فى عمليات التنبؤ فى المستقبل.

واقترح [15](GG. Judge.et.al)طريقة الحصول على أو إضافة بيانات جديدة Additional or new data

تعتبر من أسهل طرق معالجة الازدواج الخطى إلا أن قيد التكلفة يجعل من غير السهل الحصول على أو إضافة بيانات جديدة ويمكن توضيح هذه الطريقة على النحو التالى بشكل مبسط

$$Vor(\hat{\beta}_2)=\frac{\sigma^2}{\Sigma X_{2.}^2(1-r_{23}^2)} \quad \ldots\ldots\ldots\ldots (8)$$

وكما هو واضح فإنه كلما زاد حجم العينة فإن $\Sigma X_{2.}^2$ سوف تزداد بالطبع الأمر الذى يؤدى على تقليل تباين $\hat{\beta}_2$.

١. استخدام مقدرات بديلة لمقدرات كل من المربعات الصغرى والإمكان الأعظم أو هى مقدر المربعات الصغرى المقيدة، مقدر المكونات الرئيسية، مقدر انحدار ريدج ومقدر الانحدار الجزئى وفى هذا البحث سوف يتم الاعتماد على طريقتى (PCR)-(RR) لعلاج مشكلة الازدواج الخطى.

وهى: مقدر ريدج(RR)(Ridge Estimator) |20| (Hoerl and kennard)

وضح كل من [1] الصفير, [4] Ahmed, [6] Batah, [9] Bjorkstrom, [13] Draper, [20] Hoerl, [23] Le Cessie, [25] Marquardt, [28] statistical Software, [35] Webster, Ncss

مقدر ريدج والذى يتبلور فى إضافة مقادير طفيفة موجبة لعناصر القطر الرئيسى لمصفوفة حواصل الضرب التباديلى $(x'x)$ قبل إيجاد معكوسها وذلك فى حالة وجود ازدواج خطى فى النموذج، حيث تمثل (λ) مقدار طفيف موجب يسمى بمعلمة ريدج ومن خلال دراسة خصائص مقدر انحدار ريدج فإن مقدر النتائج $\hat{\beta}_R$ يكون به بعض التحيز لكنه يقلل التباين بمقدار كبير وذلك بدورة يؤدى إلى تحسين متوسط مربعات الخطأ والقدرة على التنبؤ بصورة أفضل بالإضافة إلى الحصول على مقدر مستقر لمعالم نموذج الانحدار.

$$\hat{\beta}=(x'x+\lambda)^{-1}xy$$

والفكرة الأساسية التى يبنى عليها انحدار ريدج هى أن مصفوفة التباين والتغاير للمتغيرات المستقلة تكون غير منفردة فى حالة وجود ازدواج خطى فى النموذج فيتم إضافة مقادير طفيفة موجبة لعناصر القسط الرئيسى.

مقدر ريج (Ridge Estimator) $\hat{\beta}_{R}$

$$\hat{\beta}_{R} = [x'x + KIP]^{-1} x'\bar{y}\hat{\beta} - [Ip - \lambda x'x]^{-1} \hat{\beta} \ldots\ldots\ldots\ldots\ldots\ldots (11)$$

حيث: تمثل (Ip): مصفوفة الوحدة من الرتبة p, (λ):تمثل مقدار طفيف موجب يسمى بمعلمة ريدج

$$\lambda = \frac{1}{K}, K)0 \ldots\ldots\ldots (12)$$

بينماقدم كلاً من Wold [37] , Roso [31] ,Maitra [24] , Janathan [21] , Ana [5]

طريقة مقدر المكونات الرئيسية Principle Component Regression (PCR)

والتى يمكن إثباتها فيمايلى:

نفرض أن لدينا K من المتغيرات المستقلة، فإنه يمكن إيجاد مجموعة من الدوال الخطية للمتغيرات المستقلة وذلك كما يلى Jonathan [21]:

$$Z_1 = a_1 \times_1 + a_2 \times_2 + \ldots\ldots + a_k \times_k \ (a_1^2 + a_2^2 + \ldots. + a_k^2 = 1) \ \ldots\ldots\ldots (13)$$

$$Z_2 = b_1 \times_1 + b_2 \times_2 + \ldots\ldots + b_k \times_k \ (b_1^2 + b_2^2 + \ldots. + b_k^2 = 1) \ \ldots\ldots\ldots (14)$$

حيث أن:

Z_1 : تمثل المكونات الرئيسية الأولى وهى عبارة عن دالة خطية للمتغيرات المستقلة والتى تمتلك أكبر تباين.

Z_2 : تمثل المكونات الرئيسية الثانية وهى عبارة عن دالة خطية للمتغيرات المستقلة والتى تمتلك أقل من تباين المكونات الرئيسية الأولى.

تستخدم طريقة المكونات الرئيسية فى حالة وجود عدد كبير من المتغيرات يتم وضعها فى صورة توليفات خطية حتى تتمكن من تخفيض عدد المتغيرات مع إمكانية تبسيط وفهم وتفسير الظاهرة محل البحث والاحتفاظ بأكبر نسبة ممكنة من التباين ويعاب على هذا الأسلوب:

١. قد يؤدى إلى صعوبات فى الفهم والتفسير وخصوصاً إذا اختلفت وحدات قياس المغيرات الأصلية.

٢. لا يلتفت أسلوب المكونات الرئيسية على علاقات التغايربين المتغيرات الأصلية وإنما يهتم فقط بتبايناتها حتى يتم الاحتفاظ بأكبر نسبة من هذه التباينات.

ويمتاز هذا الأسلوب بالآتى:

تتميز المكونات الرئيسية بأنها غير مرتبطة خطياً ببعضها البعض مهما كان شكل العلاقة الخطية بين المتغيرات الأصلية ولن يكون أسلوب المكونات الرئيسية مفيداً إذا كانت المتغيرات الأصلية غير مرتبطة خطياً ففى هذه الحالة ستكون المكونات الرئيسية هى نفسها المتغيرات الأصلية.

ولاستخراج المقدرات باستخدام طريقة المكونات الرئيسية يتم ذلك بإجراء انحدار التغير (Y) المعتمد على المكونات الرئيسية التى تمتلك أعلى تباين بحيث نجد أن:

$$var(z_1) \rangle var(z_2) \ldots \rangle var(z_k) \ldots \ldots \ldots \ldots \ldots (12)$$

$$var(z_1) + var(z_2) \ldots + var(z_k) = var(x_1) + var(x_2) + \ldots var(x_k) \ldots \ldots \ldots (13)$$

ولقد تم استخدام برنامج SPSS – v.r22 لتقدير معاملات (LS) وبرنامج NCSS لتقدير معاملات(PCR ,RR).

٧. مؤشرات وجود الازدواج الخطى المتعدد: Indicators of Multicollinearity

يتم استعراض مؤشرات وجود الازدواج الخطى المتعدد وهى:

١. ارتباط قيم معاملات الارتباط البسيط بين متغيرات البحث ببعضها البعض ويتضح ذلك من مصفوفة الارتباط كما هو موضح فى الجدول رقم (٢).

٢. أرتفاع قيم معاملات R^2 بين كل X وباقى الـ $X^{'s}$ فهى تزيد من ٠,٩٣ لأربع متغيرات هى X_1, X_2, X_3, X_4, X_5, وتبلغ ٠,٨٦.

٣. تزيد قيمة **VIF** عن ١٠ لكل المتغيرات X_1, X_3, X_4, X_5 وتبلغ فقط ٧,٢ للمتغير X_2.

وتقترب قيمة معامل التباين المسموح به **Tolerance** من الصفر بالنسبة لكل المتغيرات فيما عدا المتغير X_2 فتكون ٠,١٤ وأن قيمة الرقم الشرطى **Condition number** تساوى ١٤٦,٥٠ وهو يمثل أكبر قيمة لـ**Condition Index** وذلك للمتغير (x_5) و كعرف إحصائى اذا زادت قيمة الرقم الشرطى عن ٣٠ فهذا يدل على وجود مشكلة الازدواج الخطى والجدول رقم (٣) يوضح لنا نتائج معامل تضخم التباين (**Variance Inflation Factor**) ومعامل التباين المسموح به (**Tolerance**) والدليل الشرطى (**Condition Index**).

جدول رقم (٢)

مصفوفة معاملات الارتباط بين المتغيرات المستقلة والمتغير التابع

المتغيرات (المستقلة والتابع)	متوسط عدد الساعات العمل خلال السنة (Y)	متوسط أصول الأسرة بالدولار (X_1)	متوسط عمر العامل (X_2)	متوسط عدد السنوات الدراسية المنتهية (X_3)	متوسط الدخول المستحقة للأسرة غير محصلة (X4)	متوسط معدل الأجر بالدولار (X_5)
متوسط عدد الساعات العمل خلال السنة (Y)	١					
متوسط أصول الأسرة بالدولار (X_1)	٠,٧٢ (صفر)	١				
متوسط عمر العامل (X_2)	-٠,٠٨ (٠,٣٣)	٠,٤٢ (٠,٠٠٦)	١			
متوسط عدد سنوات الدراسة المنتهية (X_3)	٠,٦٨ (صفر)	٠,٦٣ (صفر)	-٠,٣٣ (٠,٠٣)	١		
متوسط الدخول المستحقة للأسرة غير محصلة (X4)	٠,٦٨ (صفر)	٠,٩٩ (صفر)	٠,٥٠ (٠,٠٠١)	٠,٥٤ (صفر)	١	
متوسط معدل الأجر بالدولار (X_5)	٠,٦٢ (صفر)	٠,٧٨ (صفر)	٠,٠٤ (٠,٤٠)	٠,٨٨ (صفر)	٠,٧٠ (صفر)	١

من الجدول السابق يتضح أن تأثير متوسط معدل الأسرة بالدولار (X_1) على متوسط عدد ساعات العمل خلال السنة (Y) أظهروجود علاقة طردية قوية ٠,٧٢ ووجود علاقة معنوية تبلغ صفر ،كما تبين أن تأثير متوسط عمر العامل(x_2) على متوسط ساعات العمل خلال السنه أظهروجودعلاقة عكسية ضعيفة جداً -٠,٣٣ ووجود علاقة غير معنوية ٠,٣٣ والعلاقة بين متوسط عمر العامل (X_2)، ومتوسط أصول الأسرة بالدولار (X_1) طردية ضعيفة ٠,٤٢ وهى علاقة معنوية تبلغ ٠,٠٠٦ ، وأن تأثير عدد سنوات الدراسة المنتهية (X_3) على متوسط عدد ساعات العمل خلال السنة (y) أظهار علاقة طردية قوية ٠,٦٨ ووجود علاقة معنوية تبلغ صفر وأن العلاقة بين متوسط عدد سنوات الدراسة المنتهية (X_3) وبين متوسط أصول الأسرة بالدولار (X_1) ومتوسط عمر العامل (X_2) طردية قوية ٠,٦٣ وعكسية ضعيفة -٣٣, وكلاهم معنوى وتبلغ معنوية كلاً منها صفر، ٠,٠٣ على التوالى وأن تأثير متوسط الدخول (X_4) على متوسط عدد ساعات العمل خلال السنة (Y) أظهر وجود علاقة طردية قوية ٠,٦٨ ووجود علاقة معنوية تبلغ صفر وأن العلاقة بين متوسط الدخول (X_4) وكلاً من متوسط أصول الأسرة بالدولار (X_1)، متوسط عمر العامل (X_2)، ومتوسط عدد سنوات الدراسة المنتهية (X_4) هى على التوالى طردية قوية جداً، طردية قوية، طردية قوية وتبلغ ٠,٩٩ ، ٠,٥٠، ٠,٥٤ على التوالى وأن تأثير متوسط معدل الأجر بالدولار (X_5) على متوسط عدد ساعات العمل خلال السنة (Y) أظهروجود علاقة طردية قوية ٠,٦٢ ووجود علاقة معنوية تبلغ صفر، وأن العلاقة بين متوسط معدل الأجر بالدولار (X_5) ومتوسط أصول الأسرة بالدولار (X_1)، ومتوسط عمر العامل (X_2) ومتوسط عدد سنوات الدراسة المنتهية (X_3) ومتوسط الدخول (X_4) هى على التوالى طردية قوية ٠,٧٨ وطردية ضعيفة جداً ٠,٠٤ وطردية قوية جداً ٠,٨٨و طردية قوية ٠,٧٠وتبين أن المتغيرات (X1) ، (X3) ، (X4) معنوية وتبلغ المعنوية على التوالى صفر أما عن المتغير (X5) فهو غير معنوى بنسبة ٠,٤٠ ومن الجدول السابق يتضح ارتفاع قيم معاملات الارتباط بين غالبية المتغيرات التفسيرية وكذا معنويتها.

جدول رقم (٣)

نتائج معامل تضخم التباين ومعامل التباين المسموح به وقيمة الدليل الشرطى

Variance	Tolerance	VIF	Condition index
X_1	0.009	111.986	5.799
X_2	0.137	7.289	15.653
X_3	0.054	18.663	42.405
X_4	0.010	99.531	89.100
X_5	0.073	13.769	146.504

٨.النماذج المقدرة Estimated Models

تم تقدير أربعة نماذج لكل المتغيرات وهى:

١-استخدام طريقة(LS) فى تقدير معاملات الخمسة متغيرات التفسيرية.

٢-العلاقات الفردية بين كل متغير تفسيرى والمتغير التابع.

٣-استخدام طريقة (RR) فى تقدير معاملات الخمسة متغيرات التفسيرية.

٤-استخدام طريقة (PCR) فى تقدير معاملات الخمسة متغيرات التفسيرية.

جدول رقم (٤)

معاملات الانحدار للمتغيرات المستقلة الخمسة بطريقة (L.S)

المتغيرات المستقلة	معاملات انحدار(LS)			القيمة الاحتمالية P- Value
	المعامل (β)	الخطأ القياسى (S.E)	نسبة t	
(X_1) متوسط أصول الأسرة بالدولار	٠,٠٤٣٥٨	٠,٠٢٠٦	٢,١٢ *	٠,٠٤
(X_2) متوسط عمر العامل	٢,٥٦٩٨ –	٣,٤٢٠٩	٠,٧٥١ –	٠,٤٧
(X_3) متوسط عدد سنوات الدراسة المنتهية	٤١,١٦٢	٢٠,٦٤٩٩	١,٩٩٣	٠,٠٥
(X_4) متوسط الدخول المستحقة للأسرة غير مرحلة	٠,٣٩٤٥ –	٠,٤٠٨٢	٠,٩٦٦	٠,٣٣
(X_5) متوسط معدل الأجر بالدولار	١٤٢,٥٣٨ –	٤٥,١٣١٣	٣,١٥٨ *	٠,٠٠٣
الجزء الثابت	٢٠٨٣,٧٧			
$ر^2$ = ٠,٧٧٩٣				
σ = ٣٢,٦١١٤				

وعند تحليل موجز لتلك النتائج اتضح بالنسبة لنموذج(LS)مع الخمس متغيرات المستقلة

- وجود علاقة موجبة ومعنوية بين (X_1) متوسط أصول الأسرة بالدولار ومتوسط عدد ساعات العمل خلال السنة (Y).
- وجود علاقة سالبة وغير معنوية بين (X_2) متوسط عمر العامل ومتوسط عدد ساعات العمل خلال السنة (Y).
- وجود علاقة موجبة وغير معنوية بين (X_3) متوسط عدد سنوات الدراسة المنتهية ومتوسط عدد ساعات العمل خلال السنة (Y).
- وجود علاقة سالبة وغير معنوية بين (X_4) متوسط الدخول ومتوسط عدد ساعات العمل خلال السنة (Y).
- وجود علاقة سالبة ومعنوية بين (X_5) متوسط معدل الأجر بالدولار ومتوسط عدد ساعات العمل خلال السنة (Y).

ولقد تم تقدير المعاملات الفردية بين كل X والمتغير التابع فكانت النتائج موضحة فى الجدول رقم (٥) وهى معاملات الانحدار للمتغيرات الفردية بطريقة (LS).

جدول رقم (٥)

معاملات الانحدار للمتغيرات الفردية بطريقة L.S

المتغيرات المستقلة	معاملات انحدار (LS)			القيمة الاحتمالية (P- Value)
	المعامل (β)	الخطأ القياسى (S.E)	نسبة(t)	
(X_1) متوسط أصول الأسرة بالدولار	٠,٠١٦ ر٢ =	٠,٠٠٣ ٠,٥١٨	٥,٩٥*	صفر
(X_2) متوسط عمر العامل	- ١,١٢٤ ر٢ =	٢,٥٣ ٠,٠٠٦	- ٠,٤٤٥	٠,٦٦
(X_3) متوسط عدد سنوات الدراسة المنتهية	٣٧,٤٠٧ ر٢ =	٦,٩٩٥ ٠,٤٦٤	٥,٣٤٧*	صفر
(X_4) متوسط الدخول المستحقة للأسرة غير مرحلة	٠,٣١٩ ر٢ =	٠,٠٦ ٠,٤٦٢	٥,٣٢١*	صفر
(X_5) متوسط معدل الأجر بالدولار	٧٨,٣٤٢ ر٢ =	٢٠,٠٧٩ ٠,٣١٦	٣,٩٠٢*	صفر

وبالنسبة للمعاملات الفردية بطريقة L.S: فتشير التقديرات الواردة فى الجدول رقم (٥) على الآتى:

وجود علاقة موجبةومعنوية بين المتغيرات التفسيرية(x_1,x_3,x_4,x_5) والمتغير التابع(y) بينماهناك علاقة سالبة وغير معنوية بين(x_2) ،(y) ولحساب الأثر المتداخل للارتباط المرتفع بين المتغيرات التفسيرية على قيم وإشارة كل متغير تفسيرى فإنه يتم التمييز (التفرقة) بين التأثير الاجمالى لكل متغير مستقل منفرداً على المتغير التابع وهومعامل الانحدار البسيط بين كل متغير تفسيرى والمتغير التابع أماالتأثير الصافى فهويعنى معامل الانحدار المقدر لكل متغير تفسيرى على المتغير التابع فى وجود المتغيرات التفسيرية الأخرى(أى الأربعة متغيرات الأخرى فى هذا البحث)

وبالتالى فهناك ٣ متغيرات وهى (X_2)، (X_3)، (X_4) غير معنوية فى نموذج الانحدار الخطى المتعدد بينما فى نموذج الانحدار الخطى البسيط (النموذج الفردى) يوجد متغير واحد فقط وهو (X_2) وعليه فيرى الباحثان أن سبب ذلك يرجع إلى وجود مشكلة الازدواج الخطى المتعدد وللتأكيد على صحة ذلك فقد تم حساب العلاقة بين التأثير الفردى لمتغيرين[1]وتأثيره الصافى فى نموذج الانحدار ذلك وفقاً للمعادلة التالية: (راجع فى هذا كتاب الاقتصاد القياسى ورسالة الدكتوراه) فى حالة خمسة متغيرات كما فى البحث وترقم على أن يأخذ Y رقم (١) والمتغيرات التفسيرية من رقم (٢) حتى رقم (٦).

$$\beta_{12.3456} = \beta_{12} - [(\beta_{32})(\beta_{13.2456}) + (\beta_{42})(\beta_{14.2356}) + (\beta_{52})(\beta_{15.2346}) + (\beta_{62})(\beta_{16.2345})$$

$$\beta_{13.2456} = \beta_{13} - [(\beta_{23})(\beta_{12.3456}) + (\beta_{43})(\beta_{14.2356}) + (\beta_{53})(\beta_{15.2346}) + (\beta_{63})(\beta_{16.2345})$$

$$\beta_{14.2356} = \beta_{14} - [(\beta_{24})(\beta_{12.3456}) + (\beta_{34})(\beta_{13.2456}) + (\beta_{54})(\beta_{15.2346}) + (\beta_{64})(\beta_{16.2345})$$

$$\beta_{15.2346} = \beta_{15} - [(\beta_{25})(\beta_{12.3456}) + (\beta_{35})(\beta_{13.2456}) + (\beta_{45})(\beta_{14.2356}) + (\beta_{65})(\beta_{16.2345})$$

$$\beta_{16.2345} = \beta_{16} - [(\beta_{26})(\beta_{12.3456}) + (\beta_{36})(\beta_{13.2456}) + (\beta_{46})(\beta_{14.2356}) + (\beta_{56})(\beta_{15.2346})$$

حيث أن: β_{16} β_{15} β_{14} ، β_{13} ، β_{12}

هى التأثيرات الفردية بين المتغير التابع وكل متغير مستقل

أما المعاملات Bij مثل β_{26} β_{25} β_{24} ، β_{23} ، β_{32} بين كل متغيرين مستقلين

، $\beta_{15.2346}$ ، $\beta_{14.2356}$ ، $\beta_{13.2456}$ ، $\beta_{12.3456}$ فهى التأثيرات الخاصة بين كل متغير تفسيرى(x,s)

أما المعاملات $\beta_{16.2345}$

(١):التأثير الصافى لـ X على Y يعنى أخذ العلاقة بين X و Y على الأخذ فى الاعتبار العوامل الأخرى.

و (y) فى حالة وجود المتغيرات المستقلة الأخرى.

فهى تمثل العلاقة بين المتغير التابع والمتغيرات المستقلة الخمسة مع الأخذ فى الاعتبار المتغيرات الأخرى فى نموذج الانحدار الخطى المتعدد ولتحليل هذا الأثر فإنه سوف يتم أخذ معامل X_5 كنموذج حيث أنه فى النموذج المقدر بطريقة (O.L.S) **سالب ومعنوى** بينما التأثير الفردى **موجب ومعنوى** وذلك كالتالى وفقاً للمعادلة الأخيرة ويتم الآن عرض التأثير الصافى لـ X_5 على Y.

$$\beta_{16.2345} = 78.342 - [(4873.362)\ (0.044) + (.415)\ (-2.451) + (2.242)\ (41.747) + (209.31)\ (-0.401)$$

$$= 78.342 - [214.4279 - 1.01883 + 93.5968 - 83.9333] = -144.73$$

وهو يقترب من قيمةالمعامل الفردى والبالغ 142.536

ويعزى هذا الفرق البسيط نتيجة تقريب كل مكونات حسابه وبالتالى يتأكد تأثير الازدواج الخطى على

تحويل العلاقة من موجب معنوى إلى سالب معنوى وعليه يتم تأكيد عدم صحة الفرض العدم (H_0) ومن ثم تأثير الازدواج الخطى على قيمة وإشارة ومعنوية معاملات الانحدار للمتغيرات X_2 و X_3 و X_4 للفروض الثانى والثالث والرابع وتأكيد عدم صحة الفرض العدم (H_0) للفروض الأول والخامس على عدم تأثير الازدواج الخطى على قيمةواشارة ومعنوية معاملات الانحدار للمتغيرات X_5 و X_1.

٩.علاج مشكلة الازدواج الخطى بأسلوب (RR , PCR)

وكمحاولة لعلاج مشكلة الازدواج الخطى تم أستخدام أسلوبين قياسين هما (RR ، PCR) وثم تقديرهما فى الصورتين الخطية واللوغارتيمية المزدوجة وفى ضوء معايير R^2 ، وإشارة المتغيرات المستقلة حيث تم الأقتصار على نموذج الانحدار الخطى المتعدد وتشير النتائج الموجودة فى الجدول رقم (٦) على المقارنة بين معاملات الانحدار والأخطاء القياسية وقيم t وكذا قيم R^2 ، RMSR والقيمة الاحتمالية(**p-value**) لنماذج (RR)، (PCR).

جدول رقم (٦)

معاملات الانحدار لنموذج (RR ، PCR)

المتغيرات	معاملات انحدار Ridge (RR)				معاملات انحدار المكونات الرئيسية (PCR)			
	المعامل (β)	الخطأ القياسي (S.E)	نسبة (t)	P-(Value)	المعامل (β)	الخطأ القياسي (S.E)	نسبة (t)	P-(Value)
(X_1) متوسط أصول الأسرة بالدولار	٠,٠١٣٢٣	٠,٠٠٣٢	٤,١٣٤٣*	صفر	٠,٠١٢٤	٠,٠٠١٨٧٨	٦,٦٠٣*	صفر
ر٢=	٠,٠١٦٣١	٠,٠٠٤٨٦	٣,٣٥٦					
(X_2) متوسط عمر العامل	-٥,٠٦٨٤	٢,١١٦١	٢,٣٩٢*	صفر	-٨,٠٧٧٩٦	١,٧١١١٥	٤,٧٢٠*	صفر
ر٢=	-٤,٨٥٧٨	٢,٤٨١٩٩	١,٩٥٥-					
(X_3) متوسط عدد سنوات الدراسة المنتهية ر٢=	٢٢,٢٥٠١	١٠,٤٧٩١	٢,١٢*	صفر	٩,٩٧٦٢١	٢,٥٩٢٢	٣,٨٥*	صفر
	٢٦,٩٣٣٦	١٣,٧٧٧٤	١,٩٥٥					
(X_4) متوسط الدخول المستحقة للأسرة غير محصلة ر٢=	٠,١٧٩٠٦	٠,٠٦٥٨٢	٢,٧٢٩	٠,٠٦	٠,٢٩٥٢٥	٠,٠٤٨٧٩	٦,٠٥*	صفر
	٠,١٣٧٤٩	٠,٠٩٦٤٥	١,٤٢٦					
(X_5) متوسط معدل الأجر بالدولار ر٢=	-٦٨,٥٨٠٣	٢٤,٩٤٢٩	٢,٧٥*	صفر	-٦١,٢٧٥٣	٠,٠٤٨٧٩	٣,٠٩*	صفر
	٨٦,٤٠١٦	٢٩,٩٥٣١	٢,٨٨					
ر٢=٠,٧١١٣ عند k=0.04 RMSE=٣٧,٣٠١٧					ر٢=٠,٧٤٥٨ =35, 005 σ			
ر٢=٠,٧٣٤٦ عند k=0.02 RMSE=٣٥,٧٦٦٣								

وتشير النتائج فى جدول رقم (٦)إلى المعاملات المقدرة لهذة المتغيرات بأستخدام RR و PCR

ويمكن إيجاز هذة النتائج فيمايلى:

أ- بالنسبة لنتائج الـ RR

١- معنوية معاملات المتغيرات(x_1,x_2 ,x_3,x_5) عند (p-value) تساوى صفر بينما كانت معنوية المتغير(x_4) عند ٠,٠٦.

٢- إيجابية أشارات المتغيرات(x_1,x_3,x_4) وسالبية إشارة المتغيرين(x_3,x_5).

٣- تبلغ R^2 لهذا النموذج ٠,٧٣٥ عند ٠,٠٢=K و RMSE = ٣٥,٧٧.

ب- فيما يتعلق بنائج طريقةPCR فهى:

١- معنوية معاملات المتغيرات الخمس عند (p-value) تساوى صفر.

٢- إيجابية أشارات المتغيرات (x_1,x_3,x_4) وسالبية إشارة المتغيرين(x_3,x_5).

٣- تبلغ قيمة R^2 للنموذج ٠,٧٤٥٨ بينما قيمة٣٥=RMSE وبالتالى يمكن القول بصفة عامة انه يوجد تقارب بين نتائج النموذجين.

وعند مقارنة النتائج الواردة فى هذا الجدول مع تلك المبينة فى جدول رقم (٥) والذى يضم معاملات الانحدار للمتغيرات الفردية المقدرة بطريقة (L.S) يتضح مايلى:

١- معنوية المتغيرات (X_1,X_3,X_4,X_5) كل بصورة منفردة مع(Y) عند مستوى p-value) يساوى صفر وعدم معنوية المتغير (x_2) منفرداً مع (Y).

٢- إتفاق إشارات المتغيرات(X_1,X_2, X_3,X_4) فى صورتها الفردية مع أشارات نفس المعاملات فى النموذجين (PCR, RR) حيث أن الثلاثه الأولى موجبه بينما أشارة المتغير (X_2) سالبة فى الصورة الفردية وكلا النموذجين.

٣- إختلاف قيم المعاملات فى صورتها الفردية وخاصة لمتغيرات (X_2,X_3) مع نظيرتها فى النموذجين (RR,PCR).

٤- أما معامل المتغير(X_5) الفردى فقد كان موجباً وقيمته ٧٨,٣٤ ومعنوى عند p-value) .

تساوى صفر بينما كان المعامل سالباً و معنوياً فى كلا النموذجين RR,PCR

ويبلغ على الترتيب – ٦٨,٥٨ ، – ٦١,٢٧ ويتوافق هذا مع فروض البحث التى تشير إلى أن وجود الازدواج الخطى يغير من قيم وإشارات بعض المتغيرات ومنها(X_5).

١٠.النتائج والتوصيات

(١٠-١): النتائج Conclusions

فى ضوء النتائج سالفة الذكر فإن البحث توصل إلى عدة نتائج عامه منها:

١-تؤثر مشكلة الازدواج الخطى بشكل معنوى على قيم وإشارة ومعنوية بعض أوكل المتغيرات التفسيرية فى نموذج الانحدار الخطى المتعدد.

٢- يؤدى – غالباً استخدام الطريقتين (RR),(PCR)إلى تحسين فى إشارة ومعنوية معاملات الانحدار المتعدد مقارنة بطريقة (O.L.S).

٣- ليس بالضرورة أن يؤدى استخدام(RR),(PCR) إلى زيادة R^2 أو تخفيض قيمة sigma مقارنة بطريقة (O.L.S)

٤-لم تقل قيمة متوسط مربع الخطأ فى ظل أسلوبى (RR) ، (PCR) حيث كانت

٣٢,٦ فى (L.S) ٣٥ فى (PCR) ٣٧,٣ فى (RR)

٥-قيم المعاملات وإشارتها فى النماذج (PCR ، RR) أصبحت أكثر منطقية عن نموذج (LS).

٦-يفضل النموذج (PCR)عن نظيره (RR) فى عدة معايير:

- الارتفاع النسبى لـ (R^2) حيث يبلغ حوالى ٧٥% بدلاً من ٧١% فى النموذج(RR)

-الانخفاض النسبى فى Sigma حيث تقدر بـ ٣٥ فى حين تقدر بـ ٣٧,٣ فى(RR).

٧- قيم وإشارات معاملات (PCR) معنوية ومتوافقة مع المنطق النظرى.

(١٠-٢): التوصياتRecommendations

تشمل هذة التوصيات الاتى:

١- ضرورة التحقق من توافر الفروض الخاصة بنظرية GAUSS-Markov ومنهاالفرض الخاص بعدم وجود ارتباط شبه تام بين المتغيرات التفسيرية.

٢- النظره المتعمقة للنتائج المقدرة بطريقة(O.L.S)وليس مجرد التعليق على النتائج " الحسابية" حتى ولو كانت غير متوافقة للمنطق العلمى والواقعى.

٣- أن إستخدام طرق التقدير الأخرى- فى حاله وجود الازدواج الخطى مثل (RR),(PCR) يفيد فى الوصول لنتائج أكثر ثقوية.

٤- يمكن التوسع فى استخدام طرق تقدير أخرى ومنها طريقة الانحدار الجزئى(PLRS).

٥- أهمية أن تكون العينات المستخدمه فى البحوث كبيرة الحجم.

٦- أن أستخدام طرق RR و PCR وغيرها يجنب الباحث الوقوع فى مشاكل التحيز (Bias Problem) التى تنتج من حذف متغير تفسيرى أوأكثر من النموذج.

References(قائمة المراجع)

أولاً: المراجع العربية:

١. الصغير، سامح (٢٠٠٨)، " **نموذج الانحدار اللوجستيى فى ظل وجود ازدواج خطى شبه تام** "، رسالة ماجستير غير منشورة – كلية التجارة – جامعة الإسكندرية.

٢. سالم كمال سلطان (١٩٧٦) – "**تحليل اقتصادي قياسي للجدارة الانتاجية للموارد الزراعية الوطنية والقومية العربية**"، رسالة دكتوراه – جامعة الأسكندرية.

٣. سالم كمال سلطان(٢٠١٤) – الاقتصاد القياسي – الطبعة الأولى، مكتبة الوفاء القانونية – الإسكندرية.

ثانياً: المراجع الأجنبية:

4. Ahmed, M. and G.M. Gilmi, (2010) "comparing Ridge regression (RR), Principal component regression (PCR) and partial least square regression (PLSR)"methods to tackle the multicollinearity problem, Department of statistics, Government College University ,Faisalabad ,Pakistan.

5. Ana, M.A.,E., and Mariano, J.V.(2006)"using principal components for Estimating Logistic Regression with high Dimensional Multicollinearty Data" , computational statistics Data Analsis,50,1905-1924.

6. Batah, F, sh. M., and Gore, S.D., (2009) "Ridge Regression Estimator. Combining unbiased and ordinary Ridge Regression Methods of estimation", online at: http://www.utgjiu.ro/math/sma

7. Belsley, D. A. (1991)" Conditioning Diagnostics, Collinearity and Weak Data inRegression". New York: John Wiley and Sons.

8. Adnan, Norliza (2006) "Comparing three methods of handling multicollinearity using simulation approach". Masters thesis, Universiti Teknologi Malaysia, Faculty of Science.

9. Bjorkstrom,(2007) "A Ridge regression and inverse problems",Stockholm university, Sweden.

10. Chatelaine, J-B, and Ralf, K., (2002) "spurious Regressions and Near – multico linearity, with an application to Aid, policies and growth, centre d' Economie de laSorbonne, Paris School of Economics", online at http://mpra.ub.uni-muenchen.de/42533/.

11. David, B. O., & Hamer, M. P. (2012)." Remediation of a perched stream culvert with ropes improves fish passage". Marine and Freshwater Research, 63(5), 440.

12. Dekker, D, Krackhardt,D.and Snijders, T,(2003) "Multicollinearity Robust QAP for Multiple Regression", university of Nijmegen, The Netherlands.

13. Draper, N.R., and Nostr and, R.C.V., (1979)" Ridge Regression and James - stein Estimation: Reriew, Technometric", vol. 4, Nov. 1979 onlineat: http://www.Jstor.org/stable/1268284

14. El – Dereny, M, and Rashvan, N. I., (2011) "Solving Multicollinearity problem using Ridge Regression Models", Int. J. Contemp, Math. Sciences, Vol.6 no.12, 585 – 600.

15. GG. Judge.(1988)"Introduction to Theory and Practice of Econometrics", John willy and sons, New York,.

16. Gujarati, d,n.,(2003) "Basic Econometrics", Fourth edition, Mc Graw-Hill Higher Education, Singapore.

17. Haistovisky,Y.(1969)"Multicollinearity in Regression Analysis: Coment" The Review of Economic and Statistics 59,486-489.–

18. Hayashi,F,(2000)"Econometrics", Princeton university press, New Jersey.

19. Hocking, RR.(1984). Discussion of "K – clustering a s a detection tool for influential subsers in regression by J.B. Gray and R.F. ling Technimetrics", 26, 321 – 323.

20. Hoerl, A.E. and Kennand, R.W.(1970). "Ridge regression Biased estimation for nonorthogonal problems, technometr", 12,55- 67.

21. Janathan Annon (2012)" Principal component analysis of customer shums in Ghanaian telecommunication Industry",vol.2nor2,December, Deportment of mathemics kvame Nkrumoh university of science and techmdoy.

22. Lawrence, B., and cedric, B(2001)"Logistic Regression when Binary Predictor variables are highly correlated Statistics" Medicine20,1431-1442.

23. Le Cessie, S., and van Houwelingen. J.c. (1992)."Ridge Estimation Logistic Regression "Applied statistics 41, 7, 191 – 201.

24. Maitra, S. and J. Yan, (2008) "principal component analysis and partial least squares: Two dimension reducation techniques" for regression. Casualty a actuarial society Discussion paper program, pp:79 – 90.

25. Marquardt, D, w., and Snee, R.D., (1975) "Ridge Regression in practice, A mericanstatisticalAssociation",onlineat: http://www.Jstor.org/stable/2683673

26. Mishra,Sk,(2004),"EstimationunderMulticollinearity:Applicationof Restricted liu and Maximum Entropy Estimators to the Portland cement Dataset",North-EasternHilluniversity,shillong,India,onlineat http://mpra.ub.unimuenchen.de/1809/

27. Nagai, I., (2008)"selection of Model selection criteria For Multivariate Ridge Regression". Dep. of Math, Graduate school of science, Hiroshima university, Japan.

28. Ncss statistical Software,(2012)" Ride Regression", chapter 335,1-21.

29. Pasha, G. Rand Shah, M,A.A.,(2004)"Application of ridge Regression to Multicollineairty Data, Journal of Research (Science), Bahauddin Zakariya university, Multan, pakistan, VoL.15.No.1pp97-106.

30. Robert, C. and Kmbokawa, T.A. KMD. E. Saleh (1992), "Empirical Bayes Estimation of the covariane Matrix of a Normal Distribution with unknown mew under an Entropy Loss", Sankhya, A, 54, 402 – 410.

31. Roso,V.M,Schenkel, F.s., Mille, S.p and Schaeffer, L, R(2005)" Estimation of genetic effects in the presence of multicollinearity in Multibreed beef Cattle evaluation" , J. Anim. Sci, 2005, 83:1788-1800.

32. Schaefer, R.L. (1983)" Bias Correction in Maximum–likelihool logistic Regression statistics in medicine 2" 71 – 78.

33. Studenmund,A.,H.,(2011)"using Econometics: Apractical Guide"Sixth Edition, pearson,Inc.New York.

34. Tang, T. and Weiner J,(2004)" Multicollinearity In CSAT studies" white paper, Ipsos - in sight.

35. Webster, J.T. Gunst, R.F. and Mason, R.L (1974) "Latent Root Regession Analysis", Technometrices, 16,513-522.

36. Welling,M.,(2013)" kernel ridge Regression, Department of computer Science", university of Toronto, canda.

37. Wold, H.,(1966)."Estmation of principal components and relate models by it erative least squares. In multivariate analysis, Ed. Krishnaiah", P.R, New York: Academic press, PP:391- 420.

38. Woold ridge, J.M.,(2003)" Introductory Econometrics: modern Approach", south western , Thomson Learning Ohio USA.

الملاحق:Appendix

الملحق الأول: بيانات البحث

المشاهدات	Y	X_1	X_2	X_3	X_4	X_5
١	٢١٥٧	٧٢٥٠	٣٨,٥	١٠,٥	٣٨٠	٢,٩٠٥
٢	٢١٧٤	٧٧٤٤	٣٩,٣	١٠,٥	٣٩٨	٢,٩٧٠
٣	٢٠٦٢	٣٠٦٨	٤٠,١	٨,٩	١٨٥	٢,٣٥٠
٤	٢١١١	١٦٣٢	٢٢,٤	١١,٥	١١٧	٢,٥١١
٥	٢١٣٤	١٢٧١٠	٥٧,٧	٨,٨	٧٣٠	٢,٧٩١
٦	٢١٨٥	٧٧٠٦	٣٨,٦	١٠,٧	٣٨٢	٣,٠٤٠
٧	٢٢١٠	٩٣٣٨	٣٩,٠	١١,٢	٤٧٤	٣,٢٢٢
٨	٢١٠٥	٤٧٣٠	٣٩,٩	٩,٣	٢٥٥	٢,٤٩٣
٩	٢٢٦٧	٨٣١٧	٣٨,٩	١١,١	٤٣١	٢,٨٣٨
١٠	٢٢٠٥	٦٧٨٩	٣٨,٨	٩,٥	٣٧٣	٢,٣٥٦
١١	٢١٢١	٥٩٠٧	٣٩,٨	١٠,٣	٣١٢	٢,٩٢٢
١٢	٢١٠٩	٥٠٦٩	٣٩,٧	٨,٩	٢٧١	٢,٤٩٩
١٣	٢١٠٨	٤٦١٤	٣٨,٢	٩,٢	٢٥٩	٢,٧٩٦
١٤	٢٠٤٧	١٩٨٧	٤٠,٣	٩,١	١٣٩	٢,٤٥٣
١٥	٢١٧٤	١٠٢٣٩	٤٠,٠	١١,٧	٤٩٨	٣,٥٨٢
١٦	٢٠٦٧	٤٤٣٩	٣٩,١	١٠,٥	٢٣٩	٢,٩٠٩
١٧	٢١٥٩	٥٦٢١	٣٩,٣	٩,٥	٣٠٨	٢,٥١١
١٨	٢٢٥٧	٧٢٩٣	٣٧,٩	١٠,١	٣٩٢	٢,٥١٦
١٩	١٩٨٥	١٨٦٦	٤٠,٦	٦,٦	١٤٦	١,٤٢٣
٢٠	٢١٨٤	١١٢٤٠	٣٩,١	١١,٦	٥٦٠	٣,٦٣٦

تابع: بيانات البحث

المشاهدات	Y	X_1	X_2	X_3	X_4	X_5
٢١	٢٠٨٤	٥٦٥٣	٣٩,٨	١٠,٢	٢٩٦	٢,٩٨٣
٢٢	٢٠٥١	٢٨٠٦	٤٠,٠	٩,١	١٧٢	٢,٥٧٣
٢٣	٢١٢٧	٨٠٤٢	٣٩,٥	١٠,٨	٤٠٨	٣,٢٦٢
٢٤	٢١٠٢	٧٥٥٧	٣٩,٨	١٠,٧	٣٥٢	٣,٢٣٤
٢٥	٢٠٩٨	٤٤٠٠	٤٠,٦	٨,٤	٢٧٢	٢,٢٨٠
٢٦	٢٠٤٢	١٧٣٩	٤١,٨	٨,٢	١٤٠	٢,٣٠٤
٢٧	٢١٨١	٧٣٤٠	٣٩,٠	١٠,٢	٣٨٣	٢,٩١٢
٢٨	٢١٨٦	٧٢٩٢	٣٧,٢	١٠,٩	٣٥٢	٣,٠١٥
٢٩	٢١٨٨	٧٣٢٥	٣٨,٤	١٠,٦	٣٧٤	٣,٠١٠
٣٠	٢٠٧٧	١٣٧٠	٣٧,٤	٨,٢	٩٥	١,٩٠١
٣١	٢١٩٦	٦٨٨٨	٣٧,٥	١٠,٦	٣٤٢	٣,٠٠٩
٣٢	٢٠٩٣	١٤٢٥	٣٧,٥	٨,١	١٢٠	١,٨٩٩
٣٣	٢١٧٣	٧٦٢٥	٣٩,٢	١٠,٥	٣٨٧	٢,٩٥٩
٣٤	٢١٧٩	٧٧٧٩	٣٩,٤	١٠,٥	٣٩٧	٢,٩٧١
٣٥	٢٢٠٠	٧٨٨٥	٣٩,٢	١٠,٦	٣٩٣	٢,٩٨٠

Source:Gujarati,d,n.,(2003)Basic Econometrics, Fourth edition, Mc Graw-HillHigher Education, Singapore.

الملحق الثالث: معاملات الانحدار باستخدام طريقة المكونات الرئيسية

Regression Coefficient Section with 2 Components Omitted

Independent Variable	Regression Coefficient	Standard Error	Stand'zed Regression Coefficient	VIF
Intercept	2352.29			
x1	0.0124122	1.878317E-03	0.5553	0.8056
x2	-8.077955	1.7115	-0.5542	1.5728
x3	9.976211	2.59222	0.1817	0.2543
x4	0.2952509	4.879176E-02	0.6292	1.2332
x5	-61.27533	20.0486	-0.4398	2.3622

Analysis of Variance Section with 1 Component Omitted prob

Source	DF	Sum of Squares	Mean Square	F-Ratio	Level
Intercept	1	1.598497E+08	1.598497E+08		
Model	5	106614.6	21322.93	18.6523	0.000000
Error	29	33152.12	1143.177		
Total(Adjusted)	34	139766.8	4110.787		

Mean of Dependent	2137.086
Root Mean Square Error	33.81089
R-Squared	0.7628
Coefficient of Variation	0.01582103

الملحق الرابع: معاملات الانحدار باستخدام طريقة ريدج

(k=0.02)

Ridge Regression Coefficient Section for k = 0.020000

Independent Variable	Regression Coefficient	Standard Error	Stand'zed Regression Coefficient	VIF
Intercept	2154.083			
x1	0.0163125	4.862074E-03	0.7299	5.1705
x2	-4.857755	2.481995	-0.3333	3.1683
x3	26.93358	13.77737	0.4906	6.8805
x4	0.1374891	0.0964466	0.2930	4.6156
x5	-86.40157	29.95308	-0.6202	5.0506

Analysis of Variance Section for k = 0.020000

Source	DF	Sum of Squares	Mean Square	F-Ratio	Prob Level
Intercept	1	1.598497E+08	1.598497E+08		
Model	5	102669.1	20533.82	16.0517	0.0000
Error	29	37097.62	1279.228		
Total(Adjusted)	34	139766.8	4110.787		

Mean of Dependent	2137.086
Root Mean Square Error	35.7663
R-Squared	0.7346
Coefficient of Variation	1.673602E-02

الملحق الخامس: ملخص نموذج الانحدار- معاملات النموذج- تحليل التباين

(1)Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics					Durbin-Watson
					R Square Change	F Change	df1	df2	Sig. F Change	
1	.884[a]	.781	.743	32.51877	781	20.634	5	29	.000	1.686

a. Predictors: (Constant), x5, x2, x4, x3, x1

b. Dependent Variable: y

(3)ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	109100.108	5	21820.022	20.634	.000[b]
	Residual	30666.635	29	1057.470		

(2)Coefficients

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	2077.621	219.787		9.453	.000
	x2	-2.455-	3.423	-.168-	-.717-	.479
	x1	.044	021	1.959	2.132	.042
	x3	41.747	20.630	.760	2.024	.052
	x4	-.401-	.407	-.854-	-.984-	.333
	x5	-143.713-	45.004	-1.031-	-3.193-	.003
Total		139766.743	34			

a. Dependent Variable: y
b. Predictors: (Constant), x5, x2, x4, x3, x1

الملحق السادس: مصفوفة معاملات الارتباط

Correlations

		y	x2	x1	x3	x4	x5
Pearson Correlation	y	1.000	-.077-	.719	.681	.680	.562
	x2	-.077-	1.000	.417	-.331-	.502	.043
	x1	.719	.417	1.000	.631	.988	.781
	x3	.681	-.331-	.631	1.000	.539	.883
	x4	.680	.502	.988	.539	1.000	.704
	x5	.562	.043	.781	.883	.704	1.000
Sig. (1-tailed)	y	.	.330	.000	.000	.000	.000
	x2	.330	.	.006	.026	.001	.402
	x1	.000	.006	.	.000	.000	.000
	x3	.000	.026	.000	.	.000	.000
	x4	.000	.001	.000	.000	.	.000
	x5	.000	.402	.000	.000	.000	.
N	y	35	35	35	35	35	35
	x2	35	35	35	35	35	35
	x1	35	35	35	35	35	35
	x3	35	35	35	35	35	35
	x4	35	35	35	35	35	35
	x5	35	35	35	35	35	35