استخدام انحدار الحرف (Ridge) لدراسة اثر بعض العوامل على المؤشر العام لسوق الأوراق المالية

إعداد المدرس المساعد

رواء صالح محمد

الجامعة المستنصرية- كلية الإدارة والاقتصاد

قسم الإحصاء

الملخص

في العديد من البحوث والدراسات الاقتصادية قد تعاني البيانات في هذه البحوث والدراسات من مشكلة التعدد الخطي وتنص هذه المشكلة على وجود علاقات (معادلات) ، خطية بين متغيرين توضيحيين أو أكثر من ذلك .

وعند وجود هذه المشكلة في البيانات فهذا يعني بان مقدّر طريقة المربعات الصغرى الاعتيادية سوف يفشل لعدم تحقق واحدة من الفرضيات الاساسية لطريقة (OLS) ، والتي تنص على عدم وجود ارتباط خطي بين المتغيرات التوضيحية وبالتالي سوف لا نحصل على مقدَّر يمتاز بخاصية (BLUE) Best linear unbiased estimater .

ولأجل معالجة هذه المشكلة والحصول على مقدّرات جيدة تمتاز بخاصية (BLUE) ، لابد من استخدام طريقة أسلوب انحدار الحرف (Ridge regression) .

والهدف من البحث هو دراسة طريقة (Ridge regression) ، بكافة تفاصيلها وتطبيقها على بيانات واقعية والتي تمثلت بالعلاقة بين المؤشر العام لسوق الاوراق المالية حيث اعتبر المتغير المعتمد (y)، الذي اعتمد على بقية المؤشرات الأخرى المتمثلة بالمتغيرات التوضيحية والمعرفة بعرض النقد الضيق (X1) ، وسعر الفائدة على الدوافع الثابتة (X2) ، وعرض النقد الواسع (X3) ، وطبقت هذه الدراسة على عينة حجمها (١٩) مفردة .

١ ـ المقدمة

عند اجراء البحوث والدراسات الاقتصادية التي تستلزم استخدام نموذج الانحدار المتعدد مراجعة الباحثون-وخصوصاً عند استخدام البيانات المقطعية لمشكلة وجود حالة تعدد العلاقات الخطية بين المتغيرات التوضيحية (Multicollinerity) سواء كان ذلك التعدد جزئياً او تاماً .

ان تطبيق بعض طرق القياس الاقتصادي وخصوصاً طريقة المربعات الصغرى (ols) للحصول على خصائص مرغوب فيها للتقديرات الناتجة يستلزم عدم وجود تلك العلاقات او التخفيف من حدة وجودها وذلك باستخدام الطرق والاساليب الاحصائية او القياسية ومن هذه الاساليب استخدام اسلوب انحدار الحرف .

اعتمدت دراسة البحث على سوق العراق للاوراق المالية والذي يمثل تنظيم ذاتي وقانوني ذو استغلال مالي واداري تأسس بموجب الامر الاداري المرقم (٧٤) وذلك بعد ان تم احلال سوق بغداد للاوراق المالية.

يجري التداول في هذا السوق على وفق الاسلوب اليدوي وبالاعتماد على نظام المزايدة العلنية المكتوبة على لوحات وقد تم تخصيص لوحة بلاستيكية لكل شركة مساهمة في السوق [٥].

ولاهمية هذا السوق في زيادة وانخفاض سعر صرف الدولار محلياً وعالمياً والذي من شأنه يعمل على تحديد الاقتصاد في البلد وعلاقته مع دول العالم، تم دراسة هذا البحث حول هذا السوق للنهوض بالواقع الحضاري الاقليمي للاقتصاد في العراق.

٢- هدف البحث:

ان الهدف الاساسي للبحث هو استخدام طريقة انحدار الحرف لدراسة تأثير بعض العوامل على المؤشر العام لسوق الاوراق المالية.

٣- مشكلة البحث

عند غياب العلاقة الخطية بين المتغيرات التوضيحية غياباً تاماً يقال عن هذه المتغيرات انها متعامدة ولكن اغلب تطبيقات الانحدار تكون المتغيرات التوضيحية غير متعامدة ومرتبطة ارتباطاً قوياً حيث يصعب تقدير تاثير كل متغير توضيحي تقديراً منفرداً في النموذج.

الجانب النظري

١-١ مشكلة الارتباط الخطي المتعدد

اصبحت مشكلة الارتباط الخطي المتعدد (Multicollinerity) معروفة منذ اكتشفها Fisher في عام ١٩٣٤ عند دراسته لسلسة زمنية تشمل عدّة متغيرات وقد تبين بعد ذلك مدى خطورتها على قيمة التقديرات ومقدار دقتها وخطورة استخدامها .ان مشكلة الحصول على مجموعة من المتغيرات التوضيحية التي تحقق كل الفرضيات التي يستلزمها تطبيق طرائق التقدير المعروفة وتعطي خاصية افضل تقدير خطي غير متحيز (Blue) التي تمتلك خاصية اقل تباين ممكن ،هي من المشاكل التي تواجه الباحث في مجال القياس الاقتصادي ،لان مشكلة (Multicollinerity) قائمة الوجود في العلاقات الاقتصادية ولكونها ذات تأثيرات خطيرة في دراسة الظاهرة المدروسة. كما انها من الصعوبة الشديدة تجنبها في معظم التطبيقات العملية حيث لا توجد علاقة خطية تامة او شبه تامة بين اي من المتغيرات التوضيحية،اضافة الى ذلك يجب ان يكون عدد المعلمات المطلوب تقديرها اقل من حجم العينة تحت البحث،اي ان [1].

.Rank (x)=k+1<n

ويعتقد klein ان مشكلة التعدد الخطي تحصل عندما تكون $r_{ij} \succ R_Y$ حيث تمثل

r_{ij} معامل الارتباط الجزئي بين المتغير التوضيحي $(x_i),(x_j)$ وان $i \neq j$.

R_Y معامل الارتباط المتعدد بين المتغير المعتمد Y والمتغيرات المستقلة (X_i) في النموذج[٢].

$$x'x = \begin{bmatrix} 1 & r_{12} \\ r_{21} & 1 \end{bmatrix}$$

$$\operatorname{var}(b_1) = \operatorname{var}(b_2) = \frac{\sigma_e^2}{1-e^2} = \frac{\sigma_e^2}{1-r_{12}^2}$$

٢-١ الكشف عن وجود ظاهرة التعدد الخطي:

تحدث هذه الظاهرة عندما تكون هناك علاقة خطية ما بين اثنين او اكثر من المتغيرات التوضيحية فمن الصعوبة ايجاد معكوسة مصفوفة المعلومات $(x'x)$ لكون محدد هذه المصفوفة مساوياً الى الصفر يعبر عنها رياضياً [٣] .

$$y_i = b_o + b_1 x_1 + b_2 x_2 + u ....(1)$$

وبفرض ان $x_1 x_2$ مرتبطان ببعضها بعلاقة تامة اي ان $x_2=kx_1$ ،حيث ان k اي ثابت اعتباطي فان تقدير معلمات هذه العلاقة $(\hat{b}_2,\hat{b}_1)$ هما[١].

$$\hat{b}_1 = \frac{\sum(X_1-\bar{X}_1)(Y-\bar{Y})\sum(X_2-\bar{X}_2)^2 - \sum(X_2-\bar{X}_2)(Y-\bar{Y})\sum(X_1-\bar{X}_1)(X_2-\bar{X}_2)}{\sum(X_1-\bar{X}_1)^2\sum(X_2-\bar{X}_2)^2 - \sum\left[(X_1-\bar{X}_1)(X_2-\bar{X}_2)\right]^2}$$

$$b_2 = \frac{\sum(X_2-\bar{X}_2)(Y-\bar{Y})\sum(X_1-\bar{X}_1)^2 - \sum(X_1-\bar{X}_1)(Y-\bar{Y})\sum(X_1-\bar{X}_1)(X_2-\bar{X}_2)}{\sum(X_1-\bar{X}_1)^2\sum(X_2-\bar{X}_2)^2 - \sum\left[(X_1-\bar{X}_1)(X_2-\bar{X}_2)\right]^2}$$

نعوض عن X_2 بقيمة KX_1 نحصل على

$$\hat{b}_1 = \frac{k^2\sum(X_1-\bar{X}_1)(Y-\bar{Y})\sum(X_1-\bar{X}_1)^2 - k^2\sum(X_1-\bar{X}_1)(Y-\bar{Y})\sum(X_1-\bar{X}_1)^2}{k^2(\sum(X_1-\bar{X}_1)^2)^2 - k^2(X_1-\bar{X}_1)^2)^2} = \frac{0}{0}$$

$$\hat{b}_2 = \frac{k\sum(X_1-\bar{X}_1)(Y-\bar{Y})\sum(X_1-\bar{X}_1)^2 - k\sum(X_1-\bar{X}_1)(Y-\bar{Y})\sum(X_1-\bar{X}_1)^2}{k^2(\sum(X_1-\bar{X}_1)^2)^2 - k^2(X_1-\bar{X}_1)^2)^2} = \frac{0}{0}$$

لذلك فان معلمات العلاقة $(\hat{b}_2,\hat{b}_1)$ هي معلمات غير نهائية وليس هناك اي طريقة لايجاد قيم منفصلة لكل معلمة من المعلمات.

ويمكن اثبات ان كل من $(\hat{b}_2,\hat{b}_1)$ يكون مساوياً الى مالانهاية في ظل وجود العلاقة الخطية التامة بين المتغيرات التوضيحية وينسحب هذا كذالك للحد الثابت فيكون تباينه كبيراً اما اذا كانت العلاقة الخطية غير تامة بين المتغيرات التوضيحية[١].

فأن

$$\text{var}-\text{cov}(b)_{ls} = S_E^2(x'x)^{-1}$$

$$\text{var}(b_1) = \frac{S_e^2\sum x_2^2}{(\sum x_1^2)(\sum x_2^2)(\sum x_1x_2)^2}$$

$$\text{var}(b_2) = \frac{S_e^2\sum x_1^2}{(\sum x_1^2)(\sum x_2^2)(\sum x_1x_2)^2}$$

وبالتعويض عن $X_2 = kX_1$ نحصل على

$$\text{var}(b_1) = \frac{S_2^2\sum x_j}{0.0} = \text{var}(b_2) = \infty$$

اما اذا كانت العلاقة الخطية غير تامه بين المتغيرات التوضيحية كمافي النموذج رقم(١) يتأثران بعامل ثالث في مثل هذه الحاله معلمات التقدير للنموذج سوف تكون غير دقيقة وغير واضحة لمشكلة الدراسة وضآلة قيمة محدد مصفوفة المعلومات $(x'x)$ التي يكون تباين المعلمات المقدرة كبيرة جداً . وذلك بسبب[١].

$$\text{var}-\text{cov}(b)_{ls} = \frac{s_e^2 Adj(x'x)}{|x'x|}$$

وبالتالي قد يستنتج خطأ بان بعض المتغيرات التوضيحية ليست لها اهمية في النموذج.اذ يظهر اختبار (t) عدم معنوية معلمات تلك المتغيرات، في حين انها في الواقع معنوية ولكن بناء النموذج يعجز عن اظهار اثر كل منها بشكل منفصل ، نظراً لارتباط هذه المتغيرات بعضها مع البعض.

اختبار وجود مشكلة التعدد الخطي:

يبدو ان كلاين يقبل بان الارتباط الخطي المتعدد ليس بالضرورة مشكلة مالم يكن ذلك لارتباط الخطي المتعدد اكبر نسبياً من درجة الارتباط المتعدد الكلي بين كل المتغيرات آنياً،يعتقد كلاين ان الارتباط الخطي يكون مؤذياً اذا كان:

$$r^2_{x_i x_j} \geq R^2_{Y,X_1,X_2 ....X_N}$$

ان اسلوب كلاين تم تحجيمه من قبل كل من فرار-وكلارير(Farrar-Glauber) في بحثهما المعنون Multicollinerity in regression analysis مشكلة الارتباط الخطي المتعدد في تحليل الانحدار، والمنشور في مجلة Review of econamics and statistic، عام ١٩٦٧ ويستند اختبار (Farrar-Glauber) الى إحصائه (χ^2) حيث يتم اختبار الفرضية التالية[٣] .

$$H_0 : X_J \quad orthogonal$$
$$H_1 : X_J \quad not\ orthogonal$$

ويمكن التعبير عن احصاءة الاختبار رياضياً كالاتي:

$$\chi^2 = -\left[n - 1 - \frac{1}{6}(2k+5)\right] \ln|D|$$

حيث ان n تمثل حجم العينة، k تمثل عدد المتغيرات التوضيحية $\ln|D|$ تمثل اللوغاتيم الطبيعي لمحدد مصفوفة معاملات الارتباط التالية [٣] .

$$D = \begin{bmatrix} 1 & r_{12} & r_{13}.........r_{1k} \\ r_{21} & 1 & r_{23}........r_{2k} \\ . & & \\ . & & \\ . & & \\ r_{k1} & r_{k2} & r_{k1}..........1 \end{bmatrix}$$

نقارن قيمة (χ_o^2) المحسوبة مع قيمة (χ^2) النظرية بدرجة حرية $k(k-1)l2$ ومستوى معنوية معين عندما تكون $\chi^2 \prec \chi_o^2$ ترفض H_0 ولانرفض H_1 اي هناك مشكلة التعدد الخطي بين المتغيرات التوضيحية والعكس صحيح.

وبعد ثبوت مشكلة التعدد الخطي بموجب الاختبار اعلاه،سيتوجب تحديد اي متغير من المتغيرات المستقلة مرتبطة خطياً والتي ادى الى حدوث مشكلة التعدد الخطي. ويتم التشخيص من خلال اختبار F وحسب الصيغة التالية[٣].

$$F = \frac{R^2{}_{j.23.....k} /(k-1)}{(1-R^2_{j.23...k})/(n-k)}$$

حسب اختبار فرضية العدم:

$$H_0 : R^2_{j.23...k} = 0$$
$$H_1 : R^2_{j.23...k} \neq 0$$

نقارن قيمة F_j المحسوبة مع قيمة F النظرية بدرجتي حرية (n-k,k-1) ومستوى معنوية α فاذا كانت $F_j>F$ ترفض H_0 ان المتغيرات التوضيحية مرتبطة مع بعضها وبعكسه ترفض الفرضية البديلة H_1 اي ان المتغيرات التوضيحية لاترتبط مع بعضها ولا يشكل مصدر قلق لمشكلة التعددالخطي، وبذلك يتم تشخيص كافة المتغيرات المرتبطة مع بقية المتغيرات التوضيحية ولغرض تحديد العوامل المسببة لحصول مثل هذه المشكلة للمتغيرات التوضيحية لذلك يجب اجراء اختبار ثالث وهو اختبار t الذي يعتمد على قيم معاملات الارتباط الجزئية ما بين كل اثنين من المتغيرات التوضيحية. وبموجب الصيغة التالية[3].

$$t_{ij} = \frac{r_{ij.12.....k}\sqrt{n-k}}{\sqrt{1-r^2_{ij.12.....k}}}$$

حيث ان $r^2_{ij.12.....k}$ يمثل مربع معامل الارتباط الجزئي ما بين المتغيرين التوضيحيين (x_i, x_j)باعتبار بقية المتغيرات التوضيحية ثابتة .حيث ان [3].

$H_0 : r_{ij.12...k} = 0$
$H_1 : r_{ij.12...k} \neq 0$

نقارن مع القيمة المحسوبة ومقابل الجدولية بدرجة حرية (n-k) ومستوى معنوية معين فاذا كانت (الجدولية t) > (t المحسوبة) نرفض H_0 ولانرفض H_1 ، اي ان الارتباط الجزئي بين x_j, x_i معنوي وبذلك يتم تشخيص بشكل نهائي المتغيرات التوضيحية التي تكون سبباً في حصول مشكلة التعدد الخطي.

اسلوب مقدّر انحدار الحرف Ridge regression estimation [4]

يعتبر اسلوب انحدار الحرف Ridge regression احد بدائل التقدير عندما يكون هناك مشكلة تعدد خطي بين المتغيرات التوضيحية للنموذج الخطي العام $(y = x\beta + \mu)$ وذلك يهدف في معالجة هذه المشكلة حيث ان :

$\mu_j \rightarrow N(0, \sigma^2 I_n)$
$E(\mu_i, \mu_j) \neq 0 \qquad \forall i \neq j$
$E(\mu_i, \chi_i) = 0$

من الملاحظ في اتباع هذا الاسلوب تستخدم الصيغة القياسية للمتغيرات المعتمدة والتوضيحية من اثناء طرح الوسط الحسابي والقسمة على الانحراف المعياري لكل متغير اذ تجري جميع الحسابات الخاصة بانحدار الحرف على اساس ذلك . وبوضع المتغيرات بالصيغة القياسية تتحول المصفوفة $(x'x)$ الى مصفوفة ارتباط المتغيرات التوضيحية.

وكما هو معلوم في حالة التعدد الخطي شبه التام ،يمكن الحصول على مقدرات اوليه لمعلمات النموذج الخطي العام،من خلال تطبيق اسلوب (OLS) وكالاتي[4] :

$b_{Ls} = (x'x)^{-1} x'y$

وكذلك لمصفوفة التباين والتباين المشترك كالاتي

$v - \text{cov}(b_{ls}) = \sigma_m^2 (x'x)^{-1}$

وعند تعويض قيمة مقدر تباين العينة (s_e^2)نحصل على مصفوفة التباين والتباين المشترك المقدرة كالاتي:

$v - \text{cov}(b_{ls}) = s_m^2 (x'x)^{-1}$
$E(s_e^2) = \sigma_m^2$

وتعد طريقة انحدار الحرف تحسين لطريقة (OLS) عند وجود التعدد الخطي شبه التام وذلك باضافة كمية موجبة صغيرة (C) للعناصر القطرية للمصفوفة $(x'x)$قبل اخذ معكوسها ليصبح التقدير للمعلمة β لنموذج الانحدار بالشكل التالي.

$$b_{RR} = \begin{bmatrix} b_{1RR} \\ b_{2RR} \\ . \\ . \\ . \\ b_{kRR} \end{bmatrix} = (x'x + cI_n)^{-1} x'y$$

حيث ان $c \geq 0$. تمثل قيمة ثابت الحرف وهي معلمة غير عشوائية عندما $(c = 0)$ فان مقدرات الـ Ridge هي نفسها مقدرات OLS .

ان صيغة التقدير المعرفة بالمعادلة اعلاه تعرف بمقدار انحدار الحرف الاعتيادي (Ordinary Ridge regression) ويرمز له (ORR) والتي تعتمد على اضافة كمية ثابتة قيمتها (C) لكل عنصر من عناصر قطر المصفوفة $(x'x)$ [4].

ويمكن تحديد قيمة C بعدّة طرق منها الطريقة التحليلية والطريقة البيانية والطريقة الاخيرة افترضها العالم [6] (Hoerl & kannard) شكلاً بيانياً اسمياه اثر الحرف Ridge Trace وهو عبارة عن التمثيل الآتي لمقدّرات انحدار الحرف المناظره لقيمة C ويحدد مدى قيم C اعتيادياً ضمن المدة (0,1)

عندما تكون ظاهرة التعدد الخطي واضحة وتمثل مشكلة حقيقه فأن مقدرات الحرف تتغير تغيراً متذبذباً عند اي ابقاء طفيف في قيمة (C) عن الصفر .

واخيراً تتجه هذه المقدرات نحو الاستقرار عند القيم الكبيرة الى (C) .

وان قيمة (C) تختار اصغر تستقر عندها مقدّرات الحرف وعندها يكون متوسط مربعات الخطأ لمقدر انحدار الحرف اصغر من متوسط مربعات الخطأ لمقدر المربعات الصغرى ، وعند هذه القيمة يبقى (MSE) اقرب الى قيمته الصغرى[4] .

الجانب التطبيقي

المقدمة:

تحصل مشكلة التعدد الخطي عندما يرتبط اثنان او اكثر من المتغيرات التوضيحية بعلاقة خطية قوية جداً بحيث يصبح من الصعب فصل اثر كل متغير توضيحي عن المتغير المعتمد، كما تحدث هذه المشكلة حينما تكون قيمة احد المتغيرات التوضيحية متساوية لكافة المشاهدات .

وباستخدام احد اساليب المعالجة يمكن التغلب على هذه المشكلة كأن يكون تكبير حجم العينة او حذف المتغير او المتغيرات التوضيحية التي ترتبط خطياً مع بقية المتغيرات التوضيحية اضافة الى اساليب علمية اخرى.

وفي هذا الجانب من البحث سنتطرق الى كيفية الكشف عن مشكلة التعدد الخطي والمتغيرات التوضيحية المسببة لهذه المشكلة ومن ثم معالجة هذه المشكلة وذلك لتقدير معلمات النموذج المستخدم في هذه الدراسة وبالاعتماد على الوسائل الاحصائية المعتمدة.

عينة البحث:

اجريت هذه الدراسة بالاعتماد على بيانات تم الحصول عليها من البنك المركزي الدولي للمدة من ١٩٩٠ ولغاية ٢٠٠٨ ، حيث تمثلت هذه البيانات بالمؤشر العام للسوق للاوراق المالية واعتبر هذا المؤشر بالمتغير المعتمد (y) . وذلك بالاعتماد على بقيه المؤشرات الاخرى، حيث تمثل المتغير التوضيحي الاول (x_1) بعرض النقد الضيق، والمتغير التوضيحي الثاني (x_2) بسعر الفائده على الدوافع الثابتة في حين كان المتغير التوضيحي الثالث (x_3) هو عرض النقد الواسع، حسبت هذه المؤشرات بالدولار الامريكي . وباستخدام البرنامج الاحصائي الجاهز NCSS تم الحصول على النتائج التي تخص هذه الدراسة ، حيث بلغ حجم عينة الدراسة (١٩) مفردة .

اختبار وجود مشكلة التعدد الخطي :

يمكن ملاحظة وجود مشكلة التعدد الخطي من استخراج مصفوفة الارتباطات للمتغيرات التوضيحية ، حيث يلاحظ قوة هذه الارتباطات مما يعني وجود هذه المشكلة وكالاتي:

$$
\begin{array}{ccc}
 & & \begin{array}{ccc} \qquad X_1 \qquad & \qquad X_2 \qquad & \qquad X_3 \qquad \end{array} \\
 & \begin{array}{c} X_1 \\ X_2 \\ \ \end{array} & \begin{pmatrix} 1.000000 & -0.959994 & 0.993512 \\ -0.939994 & 1.000000 & -0.958013 \\ 0.993512 & -0.958013 & 1.000000 \end{pmatrix} \\
r = & &
\end{array}
$$

حيث يلاحظ من خلال مصفوفة الارتباطات Correlations matrix للمتغيرات التوضيحية $(X'S)$ بان هناك علاقة قوية جداً بين المتغيرات (X_1,X_3) اذ بلغ معامل الارتباط بينهما (0.993512) ،وهي

علاقة طردية بين المتغيرين ،وفي المرتبة الثانية كانت العلاقة قوية جداً ايضاً بين (X_1,X_2) اذ بلغ معامل الارتباط بينهما (-0.939994) وهي علاقة عكسية بين المتغيرين ، وفي المرتبة الثالثة كانت العلاقة قوية جداً ايضاً بين المتغيرين (X_2,X_3) ، اذ بلغ معامل الاتباط بينهما (-0.958013) وهي علاقة عكسية ايضاً بين المتغيرين وهذا يعني وجود مشكلة التعدد الخطي (Multicollinearity problem) في النموذج بين المتغيرات التوضيحية ($X'S$) .

وباستخراج التباين وقيمة R^2 للمتغيرات التوضيحية يلاحظ التضخم بالتباين (التباين المضخم) مما يدل على وجود هذه المشكلة ايضاً وحسب الجدول رقم (١):

جدل رقم (١)

التباين المضخم

Variable	Variance inflation	R-Squard vs other xs
X_1	٨٢.٥٣٣٣	٠.٩٨٧٩
X_3	١٢.٩٨٤١	٠.٩٢٣٠
X_3	٧٨.٧١٩٣	٠.٩٨٧٣

<u>استخدام اختبار فارار كلوبر (Farrar-Glavbery):</u>

اعتمد هذا الباحث في الكشف على مشكلة التعدد الخطي باستخدام احصاءة الاختبار $\chi^{2`}$ حيث اختبر الفرضية الاتية .

$$H_0 : X_j \quad orthogonal \quad , j = 1,2,3$$
$$H_1 : X_j \quad not\, orthogonal$$

ومن مصفوفة الارتباطات للمتغيرات التوضيحية يتم استخراج المحدد لهذه المصفوفة وكانت النتيجة كالاتي:

$$|D| = -6.91162$$

ويطبق احصاءة χ_0^2 نحصل على الاتية.

$$\chi_O^2 = -\left[(19-1) - \frac{1}{6}(2(3)+5)\right] * Ln(-6.91162)$$
$$= 31.2533$$

وبالمقارنة مع χ^2 الجدولية المبينه ادناه.

$$\chi_T^2((3(\frac{3-1}{2}),0.05) = 7.816$$

ومن ملاحظة النتائج تبين لنا χ^2 المحسوبة هي اكبر من χ^2 الجدولية وهذا يعني رفض الفرضية وهذا واضحاً لوجود المشكلة .

وبعد ثبوت مشكلة التعدد الخطي بموجب الاختبار اعلاه يجب تحديد اي متغير من المتغيرات التوضيحية مرتبطاً خطياً بحيث ادى الى حدوث مشكلة التعدد الخطي .

حيث يتم الاعتماد على اختبار (F) من استخراج القيمة المحسوبة لهذا الاختبار بعد تقدير معامل الارتباط المتعدد بين المتغيرات التوضيحية (X_j) وبقية المتغيرات التوضيحية الاخرى وكالاتي:

حيث يعتمد هذا الاختبار على الفرضية الاتية.

$$H_0 : R^2 j.23....k = 0$$
$$H_1 : R^2 j.23.......k \neq 0$$

ولاختبار R^2 للمتغير التوضيحي الاول باستبعاد البقية. حيث تمثلت نتيجة هذا الاختبار بان

$$F_1 = 15.6214$$

اما R^2 للمتغير الثاني باستبعاد البقية كانت

$$F_2 = 13.7147$$

اما R^2 للمتغير التوضيحي الثالث باستبعاد البقية تمثل

$F_3 = 15.6020$

ومن مقارنة هذا الاختبار للمتغيرات التوضيحية مع قيمة F الجدولية والمبينه ادناه نحصل على :

$F_{j.23}(19.3)(3-1),0.05) = 3.63$

نلاحظ ان F المحسوبة اكبر من F الجدولية ترفض الفرضية يعني ان جميع المتغيرات التوضيحية مرتبطة خطياً مع بعضها وهي مصدر لوجود مشكلة التعدد الخطي .

ومن خلال هذا الاختبار يتم تشخيص كافة المتغيرات التوضيحية التي ترتبط خطياً مع بقية المتغيرات التوضيحية الاخرى .

ولغرض تحديد المتغيرات التوضيحية المسؤولة عن حصول مشكلة التعدد الخطي بحيث ان يجري اختبار ثالث وهو اختبار (t) والذي يعتمد على قيم معاملات الارتباطات الجزئية ما بين كل اثنين من المتغيرات المستقلة وكالاتي :

حيث يتم الاعتماد على الفرضية

$H_0 : rij.12.......k = 0$
$H_1 : rij.12........k \neq 0$

وتطبق هذه الصيغة للمتغيرات التوضيحية الثلاثة باخذ كل متغيرين سوياً باستبعاد الاخر وكالاتي :

باختبار الارتباط الجزئي للمتغيرين التوضيحيين الاول والثاني باستبعاد الثالث وبتطبيق اختبار (t) نحصل :

$t_{23.1} = 0.5432$

اما بالنسبة الاول والثالث باستبعاد الثاني وباختبار t نحصل على.

$t_{12.3} = 1.07344$

اما حساب الارتباط الجزئي ما بين المتغيرين الثاني والثالث باستبعاد الاول وبتطبيق صيغة t نحصل على :

$t_{13.2} = 23.8027$

ومن مقارنة النتائج اعلاه مع قيمة t الجدولية وكما مبينة ادناه نحصل على :

$t(19-3,0.05)=1.746$

نلاحظ ان قيمة ($t_{12.3}$ ، $t_{23.1}$) المحسوبة هي اقل من قيمة t الجدولية هذا يعني لا ترفض الفرضية اي ان الارتباطات الجزئية غير معنوية بين هذه المتغيرات التوضيحية ،اما قيمة $t_{13.2}$ المحسوبة تكون اكبر من الجدولية لذلك لا تقبل الفرضية اي الارتباطات الجزئية معنوية والتي تكون مسؤولة عن وجود مشكلة التعدد الخطي بين المتغيرين التوضيحين . وهذا الاختبار يشخص لنا بشكل نهائي المتغيرات التوضيحية التي تسببت في حصول هذه المشكلة .

ومن خلال هذا الاختبار يتبين لنا ان المتغيرين التوضيحيين x_3,x_1 هما اللذان يسببان مشكلة التعدد الخطي لارتباطهما خطياً مع المتغير التوضيحي الثاني .

اسلوب انحراف الحرف :

تعتمد هذه الطريقة على تقدير معلمات النموذج عند وجود مشكة تعدد خطي بين متغيرين توضيحيين .

حيث يتم استخراج معاملات انحدار الحرف القياسي وباعطاء قيم لـ k والذي يزول من هذه المشكلة وحسب جدول رقم (٢):

جدول رقم (٢):(معاملات التقاطع لانحدار الحرف القياسي)

K	X_1	X_2	X_3
٠.٠٠٠٩	١.٠٩٢٨	٠.٤٤٨٢	٠.٣١٢٣
٠.٠٠١٠	١.٠٣٠٧	٠.٤٣٢٦	٠.٣٥٨٦
٠.٠٠٢٠	٠.٩٨٢٠	٠.٤١٧٨	٠.٣٩٢٥
٠.٠٠٣٠	٠.٩٤٢٤	٠.٤٠٣٨	٠.٤١٧٩
٠.٠٠٤٠	٠.٩٠٩٤	٠.٣٩٠٥	٠.٤٣٧٥
٠.٠٠٥٠	٠.٨٨١٣	٠.٣٧٧٨	٠.٤٥٢٦
٠.٠٠٥٠	٠.٨٨١٣	٠.٣٧٧٨	٠.٤٥٢٦
٠.٠٠٦٠	٠.٨٥٧٠	٠.٣٦٥٥	٠.٤٦٤٦
٠.٠٠٧٠	٠.٨٣٥٦	٠.٣٥٣٨	٠.٤٧٤٠

٠.٤٨١٥	٠.٣٤٢٤	٠.٨١٦٦	٠.٠٠٨٠
٠.٤٨٧٥	٠.٣٣١٥	٠.٧٩٩٥	٠.٠٠٩٠
٠.٤٩٢٢	٠.٣٢١٠	٠.٧٨٤١	٠.٠١٠٠
٠.٥٠٥٤	٠.٢٣٢٧	٠.٦٨٠٤	٠.٠٢٠٠
٠.٤٩٦٩	٠.١٦٦٦	٠.٦٢٠٤	٠.٠٣٠٠
٠.٤٥٤٣	٠.١١٥٢	٠.٥٧٩٠	٠.٠٤٠٠
٠.٤٧١٥	٠.٠٧٤١	٠.٥٤٨٠	٠.٠٥٠٠
٠.٤٥٩٦	٠.٠٤٠٥	٠.٥٢٣٦	٠.٠٦٠٠
٠.٤٤٨٨	٠.٠١٢٦	٠.٥٠٣٧	٠.٠٧٠٠
٠.٤٣٩١	-٠.٠١٠٩	٠.٤٨٧١	٠.٠٨٠٠
٠.٤٣٠٣	-٠.٠٣١٠	٠.٤٧٣٠	٠.٠٩٠٠
٠.٤٢٢٤	-٠.٠٤٨٣	٠.٤٦٠٧	٠.١٠٠٠
٠.٣٧١١	-٠.١٤١٨	٠.٣٩٠٠	٠.٢٠٠٠
٠.٣٤٢٩	-٠.١٧٧٢	٠.٣٥٥٤	٠.٣٠٠٠
٠.٣٢٣٤	-٠.١٩٣٥	٠.٣٣٢٧	٠.٤٠٠٠
٠.٣٠٨٢	-٠.٢٠١٢	٠.٣١٥٦	٠.٥٠٠٠
٠.٢٩٥٥	-٠.٢٠٤٥	٠.٣٠١٨	٠.٦٠٠٠
٠.٢٨٤٦	-٠.٢٠٥٣	٠.٢٩٠٠	٠.٧٠٠٠
٠.٢٧٤٩	-٠.٢٠٤٦	٠.٢٧٩٦	٠.٨٠٠٠
٠.٢٦٦١	-٠.٢٠٣٠	٠.٢٧٠٣	٠.٩٠٠٠
٠.٢٥٨١	-٠.٢٠٠٧	٠.٢٦١٨	١.٠٠٠٠

ويلاحظ انه من خلال معاملات انحدار الحرف القياسي Standardized ridge regression coefficients ، بان قيمة k التي من الممكن ان تزيل مشكلة الارتباط الخطي المتعدد هي عندما (k=0.005) حيث ان هذه القيمة تكررت مررتين في الجدول (٢) وقد بلغت معاملات الانحدار القياسية (0.8813,0.3778,0.4526) على التوالي .

ومن استخراج عامل التقاطع لتضخم التباين وكما في جدول (٣) يلاحظ وضوح المشكلة . ***جدول رقم (٣) عامل التقاطع لتضخم التباين***

Variance inflation factor

K	X_1	X_2	X_3
٠.٠٠٠٠	٨٢.٥٣٣٣	١٢.٩٨٤١	٧٨.٧١٩٣
٠.٠٠١٠	٦٢.٥٢٤٣	١٢.٤٩٢٧	٥٩.٧٨٠٦
٠.٠٠٢٠	٤٩.١٥٢٦	١٢.٠٣٢٧	٤٧.١١٧٠
٠.٠٠٣٠	٣٩.٧٦٨٣	١١.٦٠٠١	٣٨.٢٢٣٦
٠.٠٠٤٠	٣٢.٩٢٣٧	١١.١٩٢٠	٣١.٧٣٢٠
٠.٠٠٥٠	٢٧.٧٧٣٨	١٠.٨٠٦١	٢٦.٨٤٣٣
٠.٠٠٥٠	٢٧.٧٧٣٨	١٠.٨٠٦١	٢٦.٨٤٣٣
٠.٠٠٦٠	٢٣.٧٩٨١	١٠.٤٤٠٦	٢٣.٠٦٥٦
٠.٠٠٧٠	٢٠.٦٦١٨	١٠.٠٩٣٩	٢٠.٠٨٢٣
٠.٠٠٨٠	١٨.١٤١٩	٩.٧٦٤٧	١٧.٦٨٢٥
٠.٠٠٩٠	١٦.٠٨٤٨	٩.٤٥١٧	١٥.٧٢١١
٠.٠١٠٠	١٤.٣٨٢٢	٩.١٥٣٨	١٤.٠٩٥٥
٠.٠٢٠٠	٦.٣٥٢٠	٦.٨٢٥٠	٦.٣٧٧٩
٠.٠٣٠٠	٣.٧٦٥٨	٥.٢٩٣٠	٣.٨٤٩٥
٠.٠٤٠٠	٢.٥٦٩٠	٤.٢٣١٠	٢.٦٦٠٢

٠.٠٥٠٠	١.٩٠١٠	٣.٤٦٤٥	١.٩٨٦٧
٠.٠٦٠٠	١.٤٨٣١	٢.٨٩٣٢	١.٥٦٠٤
٠.٠٧٠٠	١.٢٠١٢	٢.٤٥٥٩	١.٢٧٠٠
٠.٠٨٠٠	١.٠٠٠٦	٢.١١٣٧	١.٠٦١٦
٠.٠٩٠٠	٠.٨٥١٨	١.٨٤١٠	٠.٩٠٦١
٠.١٠٠٠	٠.٧٣٨١	١.٦٢٠٠	٠.٧٨٦٥
٠.٢٠٠٠	٠.٣٠١٥	٠.٦٥٠٤	٠.٣٢٠٦
٠.٣٠٠٠	٠.١٩٢٨	٠.٣٧٥٤	٠.٢٠٢٨
٠.٤٠٠٠	٠.١٤٧١	٠.٢٥٨٥	٠.١٥٣٢
٠.٥٠٠٠	٠.١٢٢٢	٠.١٩٦٩	٠.١٢٦٣
٠.٦٠٠٠	٠.١٠٦٣	٠.١٥٩٧	٠.١٠٩٣
٠.٧٠٠٠	٠.٠٩٥١	٠.١٣٥٠	٠.٠٩٧٣
٠.٨٠٠٠	٠.٠٨٦٦	٠.١١٧٤	٠.٠٨٨٣
٠.٩٠٠٠	٠.٠٧٩٧	٠.١٠٤٢	٠.٠٨١١
١.٠٠٠٠	٠.٠٧٤٠	٠.٠٩٣٩	٠.٠٧٥١

ويمكن ان نلاحظ من خلال جدول (٣) عامل تضخم التباين بان قيمة k قد بلغت (0.005) لتكرارها وهو تأكيد للجدول رقم (٢) جدول معاملات انحدار الحرف القياسية وكان مستوى عامل تضخم التباين بالنسبة للمتغير التوضيحي x_1 هو (27.7738) في حين ان مستوى عامل تضخم التباين بالنسبة للمتغير التوضيحي x_2 هو (10.8061) ،اما مستوى عامل التضخم بالنسبة للمتغير التوضيحي x_3 هو(26.8433) .

<u>مقارنة بين تقدير المربعات الصغرى وتقدير الحرف:</u>

جدول رقم (٤) بين قيم معامل التقاطع والميل الحدي والخطأ المعياري ومعامل التحديد بطريقة التقدير OLS و Ridge وقيمة k=0.005 وكالاتي:

Independent variable	Ridge coeffs	LS Coeffs	Ridge standard error	LS Standarl error
Intercept	-٣٠٨٥.٩٦٤	-٣٥٣٤.٣٥١		
X_1	٠.٠٢٦٠٥٦٦٩	٠.٠٣٢٣٠٩٤٨	٠.٠٠٨٤٦١٧٠٣	٠.٠١٣٤٩٥٨٤
X_2	١٨٨.٦٨٠٨	٢٢٣.٨٦٠٦٦	٨٩.١٦٦٦٥	٩٠.٤٣١٣٣
X_3	٠.٠٠٢٨٨٦٤٢	٠.٠٠١٩٩١٢٨٧	٠.٠٠١٧٩٤١٦٦	٠.٠٠٢٨١٢٦٩١٦
R-Squard	٠.٩٥٥٨	٠.٩٦٢١		
Sigma	٢٤٠.٧٢٦٤	٢٢٢.٧٢٥٠		

ويلاحظ من خلال جدول (٤) بان قيمة معامل التقاطع في تقدير Ridge قد بلغ $\hat{\beta}_o = -3085.96$ في حين ان قيمة معامل الميل الحدي للتقدير (x_1) قد بلغ $\hat{\beta}_1 = 0.026$ ، اما قيمة معامل الميل الحدي لتقدير لـ (x_2) قد بلغ $\hat{\beta}_2 = 188.68$ في حين قيمة معامل الميل الحدي لتقديرلـ (x_3) قد بلغ $\hat{\beta}_3 = 0.00288$ في حين كان تقديرات OLS هي $\hat{\beta}_o = -3534.351$ وميل الحدي للمتغير الاول $\hat{\beta}_1 = 0.03230$ وللمتغير الثاني $\hat{\beta}_2 = 223.860$ وللمتغير الثالث $\hat{\beta}_3 = 0.001991$ على التوالي .
ويمكن معرفة قيمة K من خلال الجدول رقم (٤) وكالاتي :

$$K = 0.005 = \frac{5}{1000} = \frac{5}{10^3} = 5*10^{-3}$$
$$= 0.005$$

ويلاحظ من الجدول ايضاً بان قيمة الخطأ المعياري لتقدير Ridge بالنسبة للمتغير الاول x_1 قد بلغ ٠.٠٠٨٤٦ في حين ان قيمة الخطأ المعياري لتقدير Ridge بالنسبة للمتغير الثاني x_2 قد بلغ ٨٩.١٦٦ اما قيمة الخطأ المعياري لتقدير Ridge بالنسبة للمتغير الثالث x_3 قد بلغ ٠.٠٠١٧٩ ،ومن مقارنتها مع قيم الخطأ المعياري لتقدير OLS بالنسبة لجميع المتغيرات التوضيحية على التوالي قد بلغت (0.01349,90.431,0.00284) يتضح بان قيم الخطأ المعياري عند استخدام طريقة تقدير Ridge هو افضل واقل من قيم الخطألمعياري عند استخدام طريقة تقدير OLS ،مما يعني ان طريقة Ridge تقلل وتزيل مشكلة التعدد الخطي الموجود بين المتغيرات التوضيحية .

كما يلاحظ بان قيمة معامل التحديد عند استخدام طريقة Ridge قد بلغ ٠.٩٥٨٥ في حين كان $\sqrt{(m.s.E)}\sigma = 240.7269$ اما قيمة معامل التحديد عند استخدام طريقة OLS قد بلغ ٠.٩٦٢١ وان

$\sqrt{(m.s.E)} = \sigma = 222.725$ ، وباتبـاع اسـلوب Ridge ومـن خـلال حسـاب معلمـات النمـوذج بطريقـة Ridge والتي تبين النتائج انها الافضل من OLS نحصل على النموذج الملائم لهذه الدراسة والذي يعتبر افضل نموذج .

$$Y = -3085.964 + 0.02605669X_1 + 188.6808X_2 + 0.00288642X_3$$

اما جدول تحليل التباين للنموذج.

جدول تحليل التباين

Source	D.F	Sum of squares	Mean square	Frantic	Prob level
Intercept	١	٠.٠٠٠٠٠٠١٦٠	٠.٠٠٠٠٠٠١٦٠		
Model	٣	٠.٠٠٠٠٠٠٠١٨٧	٦٢٦٠١٢٦	١٠٨.٠٢٧٨	٠.٠٠٠٠٠٠
Error	١٥	٨٦٩٢٣٨	٥٧٩٤٩.٢		
Total(Adjusted)	١٨	٠.٠٠٠٠٠٠١٩٦	١٠٩١٦٤٥		

وان المتوسط	٩٢٠.٢٦٨٤	
متوسط مربعات الخطأ	٢٤٠.٧٢٦٤	$\sqrt{MSE}$
معامل التحديد	٠.٩٥٥٨	R^2
معامل الاختلاف	٠.٢٦١٣٨٢٥	

ومن خلال جدول تحليل التباين يلاحظ عند استخدام طريقة تقدير Ridge فأن قيمة F قد بلغت (108.027) وهي معنوية وذلك لان قيمة (Probability level) مستوى المعنوية كانت مساوية لـ Zero وهذا يعني بان قيم المعلمات الموجودة في النموذج هي معنوية وذات اهمية وضرورية في النموذج، وان النموذج هو الافضل .

ومن خلال استخدام النموذج الافضل يمكن استخراج قيم الخطأ والقيم التنبؤية .والجدول(٥) يوضح ذلك .

جدول رقم(٥) (القيم التنبؤية)

Row	Predicted	Residucl
١	-١٥٢.٩٣٣٩	٢٣٧.٩٣٣٩
٢	-٦٧.٥٨٩٩٤	١٥.٩٥٨٩٩
٣	٢٨.٦٦٠٣	٧١.٣٣٩٦٩
٤	١٦٧.٩٧٩١	-١٨.٩٧٩٠٩
٥	٢٧٢.٨٢٥	-٣٤.٨٢٥٠١
٦	٢٣٦.٢٠٧	-٢٢.٢٠٧٠٤
٧	٢٤٧.٤٤١٥	٤٩.٥٥٨٤٨
٨	٢٦٧.٠٩٩	٩١.٩٠١٠٣
٩	٣٨٣.١٢٩٢	-١.١٢٩٢٢٢
١٠	٣٣٠.١٦٦١	٩٣.٨٣٣٨٣
١١	٦٩٤.٣٢٠٦	-٦٨.٣٢٠٦٣
١٢	٨٣٦.٠١٥٣	-٢٢٣.٠١٥٢
١٣	١٠٤٩.٣٥١	-٤١٤.٣٣٠٨
١٤	١١٦١.٣٨٨	-٣٥٨.٣٨٧٧
١٥	١٤٩٨.١٧٥	٢٧٧.٠٧٥
١٦	١٨٤٥.٧٦٩	٣٩٢.٢٣٠٨
١٧	٢١٨٢.٧٦	١٩٧.٢٣٩٩
١٨	٢٩٤٢.٥٧٢	-٣٢.٥٧١٤٥
١٩	٣٣٦٢.٧٦٦	١٥٧.٢٣٣٦

ومن خلال استخراج القيم التنبؤية ($\hat{y}$) للنموذج الافضل يلاحظ وجود قيم تنبؤية سالبة في الفترتين الاولى والثانية مما يعني انها لسبب ذات اهمية كبيرة للاستعانة بها من قبل الباحث وانما يفضل الاعتماد على القيم التنبؤية لبقية الفترات والتي يمكن ان تفيد الباحث في الاستعانة لهذه القيم المستخرجة لبناء الخطط المستقبلية ولاجراء البحوث التي يمكن من شأنها ان تساعد الباحثين في اجراء الدراسات او اتخاذ ما يلزم في بعض الامور التي تستوجب التنبؤ بها مستقبلاً من اجل اتخاذ القرارات المناسبة وذلك لعدم حصول اي مشكلة في موضوع الدراسة.

الاستنتاجات

من خلال اجراء هذه الدراسة تم التوصل الى جملة من الاستنتاجات.

١- تحصل مشكلة التعدد الخطي عندما تكون قيمة التباين للمتغيرات التوضيحية كبيرة.

٢- تعتبر قيمة ($k=0.005$) هي القيمة المثلى التي من شأنها ان تزيل مشكلة التعدد الخطي والتي قد بلغت معاملات لانحدار القياسية عندها (0.8813,0.3778,0.4526) على التوالي للمتغيرات التوضيحية x_3,x_2,x_1

٣- يعتبر متوسط مربعات الخطأ افضل معيار للمقارنة فمن مقارنة طريقة تقدير Ridge مع طريقة OLSنلاحظ بان قيم تقدير Ridge هي الافضل وذلك لان قيم الخطأ المعياري اقل في حين كانت قيم الخطأ المعياري لـ OLS هي الاكثر.

٤- من معنوية قيمة F التي ظهرت في جدول تحليل التباين يتبين ان النموذج المقدر بطريقة Ridge هو النموذج الافضل.

التوصيات

توصل الباحث الى توصيتين هما

١- استخدام النموذج المقدّر بطريقة Ridge والذي تمثل بافضل نموذج لتقدير القيم التنبؤية التي من شأنها قد تفيد الباحث في اجراء البحوث المستقبلية او قد تستخدم في اتخاذ القرارات اللازمة .

٢- يفضل اجراء الدراسة على عينات كبيرة الحجم وذلك للتقليل من حدوث هذه المشكلة .

المصادر العربية

١- هادي كاظم/ د.اموري / 2002 / طرق القياس الاقتصادي المتقدم النظرية والتطبيق / بغداد / مطبعة الطيف.

2- الحيالي/ د. طالب /1991/ القياس الاقتصادي/ بغداد/ مطبعة التعليم العالي .

3- القريشي/ د. محمد صالح / 2004/ مقدمة في الاقتصاد القياسي /عمان –الاردن.

4- شاكر / حمدية 2007/ اتجاهات التغيرات الهيكلية في اقتصادات التحول من نظام التخطيط المركزي الى نظام اقتصاد السوق للمدة (1990-2004)/ اطروحة دكتوراه / الجامعة المستنصرية / كلية الادارة والاقتصاد.

5- عدنان / هدى / 2008/ تمثيل فضاء الحالة لنماذج السلاسل الزمنية التركيبية ونماذج بوكس – جينكز مع تطبيق في سوق العراق للاوراق المالية / رسالة ماجستير الجامعة المستنصرية / كلية الادارة والاقتصاد.

المصادر الاجنبية

1- Hoerl,A.E&Kannard,R.W/1975/Ridge regression .biased estimation for non orthogonal problems,Technometric,v. 12,N.2 /p100-102.
٢-Koutsoyiannis,A,(1977),Theory of econometrics,Elibs with Macmillan press., London.