معالجة مشكلة الازدواج الخطي باستخدام انحدار الحرف
دراسة تطبيقية على دالة الانفاق الاستهلاكي في الجزائر خلال الفترة 1970-2011

د. ساوس الشيخ*

جامعة أدرار

ملخص :

تم الاعتماد على أسلوب انحدار الحرف (Ridge) المناسب في حالة مشكلة الازدواج الخطي دون اللجوء لحدف المتغيرات المستقلة المتسببة في حدوث الازدواج الخطي، ذلك ان هذه المشكلة تظهر في كثير من النماذج القياسية خاصة دالة الانفاق الاستهلاكي التي فيها العلاقة بين الاستهلاك الحالي والاستهلاك السابق، ولقد تم الحصول على البيانات الاحصائية من مصادرها المختلفة خلال الفترة (1970-2011).

الكلمات المفتاح : الإزدواج الخطي؛ انحدار الحرف؛ الصادرات؛ دالة الاستهلاك؛ الجزائر.

Abstract:

This paper attempts to solve the multicollinearity problem which appears in the multiple linear regression models by deploying Ridge's regression and without omitting any of the explanatory variables which may cause this problem. We have applied the Ridge's regression to solve the multicolineraty problem in the consumption function in the case of Algeria, using real data from different sources that cover the period from 1970 to 2011.

***Keywords*:** Multicollinearity; Ridge Regression; the consumption function; Algeria.

* عضو مخبر التكامل الاقتصادي، كلية العلوم الاقتصادية والتجارية وعلوم التسيير، جامعة أدرار.

مقدمة:

يشكل الاستهلاك العائلي نسبة معتبرة من الناتج المحلي الاجمالي لأي بلد في العالم سواء كانت من الدول المتقدمة او النامية. وفي الجزائر فان نسبة الانفاق الاستهلاكي العائلي الى اجمالي الناتج المحلي معتبرة حيث بلغت سنة 2012، 2011، 2010، 2009، 2008، ما نسبته 32.89%، 31.32%، 34.32%، 37.56%، 30.18 على التوالي. بالإضافة الى مجموعة من العوامل التي تأكد على اهمية دراسة وتحليل سلوك الاستهلاك في الجزائر والتنبؤ به، منها الزيادة السكانية، والانفتاح على الثقافات العالمية المختلفة. حيث تناولت العديد من الدراسات في الجزائر موضوع الاستهلاك التي اكدت على تأثره ببعض المتغيرات الاقتصادية وغير الاقتصادية، باستخدام اساليب احصائية مختلفة.

اهمية موضوع الاستهلاك في الدراسات الاقتصادية، ترجع لارتباط الموضوع بالانسان، فضلا على ان تحليل سلوك الاستهلاك الكلي يساعد على رسم السياسات الاقتصادية على مستوى الدولة، كما ان الاستهلاك يخضع لمجموعة من المتغيرات الاقتصادية وغير الاقتصادية فهو يؤثر ويتأثر بها. ولا يقف تأثير دلك فقط على المتغيرات الداخلة وإنما كذلك الاسلوب المستخدم في التقدير فهناك فرضيات للاستهلاك تعاني من ازدواج خطي، وهو ما يتعارض مع بعض الاساليب المستخدمة في التقدير.

تتمثل الاشكالية الرئيسية التي تحاول الورقة البحثية معالجتها في السؤال التالي : **ما هي اهم العوامل المؤثرة في الانفاق الاستهلاكي في الجزائر في ظل مشكلة الازدواج الخطي؟** سوف يتم ذلك من خلال تحديد وتقدير نموذج قياسي باعتماد انحدار الحرف من خلاله يمكن قياس تأثير المتغيرات الاقتصادية المقترحة على سلوك الاستهلاك العائلي في الجزائر. وعلى ضوء الاشكالية السابقة يمكن صياغة الفرضية التالية : يتفق سلوك الاستهلاك العائلي في

الجزائر مع معظم النظريات الاقتصادية للاستهلاك. ولمعالجة الاشكالية السابقة تم تقسيم البحث الى ثلاث محاور رئيسية كما يلي :

أولاً. مشكلة الازدواج الخطي.

ثانياً. انحدار الحرف.

ثالثاً: الدراسة التطبيقية.

أولاً. مشكلة الازدواج الخطي Multicollinearity:

أن من الشروط اللازم توفرها لاستخدام طريقة المربعات الصغرى في تقدير معلمات نموذج الانحدار عدم وجود ارتباط أو علاقة خطية بين اثنين أو اكثر من المتغيرات المفسرة في نموذج الانحدار. ويشير الازدواج الخطي الى الحالة التي يكون فيها بين متغيرين اثنين أو اكثر من المتغيرات المفسرة في نموذج الانحدار ارتباط قوي، مما يجعل من الصعب أو المستحيل عزل تأثيراتها الفردية على المتغير التابع، في وجود هذا الإزدواج الخطي، فان تقدير معاملات نموذج الانحدار باستخدام طريقة المربعات الصغرى قد تكون غير معنوية احصائياً وقد تأخذ الاشارة الخاطئة بالرغم من أن معامل التحديد R^2 قد تكون عالية[1].

رياضياً يكون العلاقة الخطية بين متجهات المتغيرات المستقلة $X_1, X_2,..., X_n$ إذا تحقق الشرط التالي:

$$C_1X_1 + C_2X_2 + ... + C_nX_n = \underline{0}$$

حيث أن الثوابت $C_1, C_2,..., C_n$ لا تساوي جميعها الصفر، وفي اغلب حالات الانحدار الخطي فانه من غير المحتمل أن تكون العلاقة تامة، فإذا كان:

$$C_1X_1 + C_2X_2 + ... + C_nX_n \approx \underline{0}$$

[1]- دومينك سلفاتور، **ملخصات شوم نظريات ومسائل في الاحصاء والاقتصاد القياسي**، ترجمة سعدية حافظ منتصر، دار ماكجرو هيل للنشر، القاهرة، 1982، ص 210.

أي مساوياً للصفر تقريباً، فاننا نقول أن متجهات المتغيرات $X_1, X_2, ..., X_n$ مرتبطة خطياً.[1]

1. أسباب الازدواج الخطي :

هناك العديد من مصادر تعدد العلاقات الخطية ومن أهم تلك المصادر ما يلي :-

- ✓ اتجاه بعض المتغيرات للتغير سوياً، فمثلاً نجد أن متغيرات دخل الموظف وسنوات خبرته وعمره ومرتبته الوظيفية غالباً ما تتغير سوياً ويكون بينها ارتباط قوي؛
- ✓ ادراج متغيرات مستقلة ذات تباينات صغيرة أو تشتت أيقل؛
- ✓ اساليب المعاينة يظهر هذا المصدر من خلال معرفة أو عدم معرفة صاحب التجربة بانه ياخذ بعض المتغيرات المستقلة وهذه المتغيرات تعرف واحد أو أكثر من معادلات الارتباطات الخطية، ويظهر هذا المصدر في المجالات الصناعية، فعلى سبيل المثال عندما يرغب الباحث بالتنبؤ بالأرباح من خلال دراسة (الدخل – تكاليف العمل) وبتحليل هذا النوع من البيانات تظهر علاقة خطية قوية بين الدخل وتكاليف العمل حيث ان زيادة تكاليف العمل سوف تؤذي الى زيادة الدخل.
- ✓ القيود الفيزياوية في النموذج أو المجتمع، يظهر هذا المصدر كثيرا في التحليلات الكيماوية حيث ان تكوين مركب ما يعتمد على اضافة مجموعة من العناصر بأوزان ثابتة، في بعض الاحيان يكون من الصعب معرفة القيود الفيزياوية في النموذج التي تكون السبب في الازدواج الخطي بين المتغيرات المستقلة.

[1] Gunst, R. F . and Mason, R. L. **Biased estimation in regression: an evaluation using mean squared error**. J. Amer . Statist. Assoc. 72, 1977. p p 616-628.

- ✓ يؤدي صغر حجم العينة بحيث يصبح عدد المشاهدات قريبا من عدد المتغيرات التفسيرية الى ظهور مشكلة الازدواج الخطي.[1]

2. الكشف عن مشكلة الازدواج الخطي:

هناك العديد من الاساليب المستخدمة للكشف عن مشكلة الازدواج الخطي بين المتغيرات المستقلة من اهمها :-

2-1. مصفوفة معاملات الارتباط الجزئي :

يعتبر هذا الاسلوب سهل وبسيط للكشف عن الازدواج الخطي بين المتغيرات المستقلة حيث يتم فحص المعاملات غير القطرية في مصفوفة معاملات الارتباط الجزئي $(X'X)$ التي تمثل قيم الارتباط البسيط بين المتغيرات المستقلة، فإذا كان $|r(X_i, X_j)| \approx 1$ معنى ذلك وجود علاقة خطية بين X_i, X_j ، وهناك من يقترح الكشف عن مشكلة الازدواج الخطي اذا كان $|r(X_i, X_j)| > 0.7$ وذلك من خلال رسم الشكل الانتشاري X_i, X_j فإذا امكن التعبير عن الشكل الانتشاري بخط مستقيم معنى ذلك انه هناك ازدواج خطي بين المتغيرين X_i, X_j .

2-2. عامل تضخم التباين:

يمكن كتابة عناصر قطر المصفوفة $C = (X'X)^{-1}$ كما يلي[2]:

$$C_{jj} = \frac{1}{(1 - R_j^2)} \quad j = 1,2,...,k$$

حيث R_j^2 يمثل معامل تحديد المتعدد لانحدار X_j على باقي المتغيرات المستقلة $k-1$ ، فإذا كان المتغير X_j متعامد تقريباً مع المتغيرات المتبقية فان R_j^2 سيكون صغيراً وبالتالي فان C_{jj} سيكون قريبا من 1، وإذا كان X_j مرتبطا

---

[1]- عبد القادر محمد عبد القادر عطية، **الحديث في الاقتصاد القياسي بين النظرية والتطبيق**، الدار الجامعية، الاسكندرية، 2005، ص 472.

[2] -Douglas C. Montgomery, George C. Runger, **Applied Statistics and Probability for Engineers**, John Wiley & Sons, Inc. USA, 2002, p 460.

خطيا مع بعض المتغيرات الاخرى فان R_j^2 سيكون قريبا من 1 وبالتالي C_{jj} سيكون كبير جداً، وحيث ان تباين معامل الانحدار للمتغير X_j يمكن كتابته على الشكل التالي :

$$V(\hat{B}_j) = \sigma^2 C_{jj}$$

وعليه فان C_{jj} عامل مهم في زيادة تباين $\hat{B}_j$ اذا كان هناك ازدواج خطي بين المتغيرات المستقلة، وبالتالي فان معامل تضخم التباين $VIF_j = C_{JJ}$ لكل متغير في نموذج الانحدار يقيس تاثير العلاقة الخطية بين المتغيرات المستقلة على تباينات معلمات الانحدار.[1]

اقترح المقياس VIF من قبل Glauber & Farrar عام 1967 وسماه Marguardt في عام 1970 عامل تضخم التباين variance inflation factor. وحدد[2] Gunst & Mason, 1980 انه اذا كان $C_{JJ} \geq 4$ فان ذلك دليلا كافياً على أن X_j مرتبطا مع بقية المتغيرات.

2-3. القيم الذاتية:

ايجاد القيم الذاتية للمصفوفة $X'X$ حيث أن :

$$L = V'(X'X)V$$

حيث L مصفوفة قطرية ذات بعد $P \times P$ عناصر قطرها الرئيسي عبارة عن القيم الذاتية $\lambda_1, \lambda_2, ..., \lambda_p$ لمصفوفة الارتباط $X'X$ ، V مصفوفة لها خاصية التعامد المعياري. فاذا كان هناك قيمة ذاتية قريبة من الصفر $\lambda_i \approx 0$ دل ذلك على وجود ارتباط خطي شبه تام. واذا كانت القيمة الذاتية تساوي الصفر $\lambda_i = 0$ دل ذلك على وجود ارتباط خطي تام. وتساعد القيم الذاتية في

[1] -Edward R. Mansfield, Billy P. Helms, Detecting Multicollinearity, **The American Statistician**, V 36, 3, 1982, p p pages 158-160.
[2] -Gunst, R. F . and Mason, R. L. **Op cit**. p p 616-628.

حساب نوعين من الاحصاءات التي تساهم في الكشف عن وجود ازدواج خطي هما :

- مؤشر الحالة :

$$CI = \sqrt{\frac{\lambda_{\max}}{\lambda_j}} \quad j = 1,2,..., p$$

حيث:

$\lambda_{\max}$: اكبر قيمة ذاتية. λ_j : القيمة الذاتي في الحالة j

- رقم الحالة :

$$CN = \sqrt{\frac{\lambda_{\max}}{\lambda_{\min}}}$$

إذا كانت $10 < CN > 30$ يدل ذلك على وجود ازدواج خطي معتدل، اذا كانت $CN > 30$ يدل على وجود ازدواج خطي حاد.[1]

3. نتائج مشكلة الازدواج الخطي :

ان وجود ارتباط خطي قوي بين المتغيرات يؤدي الى ان تكون مقدرات المربعات الصغرى كبيرة الحجم وبعيدة عن المعلمات الحقيقية، كما هو موضح كما يلي:

ان مربع الخطأ أو مربع المسافة بين مقدر المربعات الصغرى $\widehat{B}_{ls}$ ومتجه المعلمات الحقيقي B يساوي:

$$L^2 = \left(\widehat{B}_{ls} - B\right)' \left(\widehat{B}_{ls} - B\right)$$

وعليه فان متوسط مربع الخطأ يساوي:

[1]- آيات عبد المهدي محمود السباح، **مقارنة بين انحدار المركبات الرئيسية والجذور الصماء بأستخدام أسلوب المحاكاة**، رسالة ماجستير، ك قسم الرياضيات، كلية العلوم، جامعة ال البيت، الاردن، 2008، ص 07.

$$E(L^2) = E\left(\left(\hat{B}_{ls} - B\right)'\left(\hat{B}_{ls} - B\right)\right)$$

$$= \sum_{j=1}^{p} \mathrm{E}\left(\hat{B}_{ls} - B\right)^2$$

$$= \sum_{j=1}^{p} Var\left(\hat{B}_{jls}\right)$$

$$= \sigma^2 tr(X'X)^{-1}$$

$$= \sigma^2 \sum_{j=1}^{p}\left(\frac{1}{\lambda_j}\right)$$

وبالتالي فان العلاقة بين مربع المسافة $E(L^2)$ و λ_j عكسية والذي يكون صغير جدا عند وجود ارتباط خطي مما ينتج عنه كبر $E(L^2)$ وبالتالي فان مقدرات المربعات الصغرى ستكون كبيرة. وهناك تأثير اخر للازدواج الخطي على معنوية مقدرات المربعات الصغرى نتيجة كبر معامل التحديد وبالتالي تضخم تباين مقدرات المربعات الصغرى.

4. معالجة مشكلة الازدواج الخطي :

اهتم العديد من الباحثين في ايجاد الحلول المناسبة لمعالجة مشكلة الازدواج الخطي سواء كان باختزاله او جمع بيانات أكثر (رفع حجم العينة)، وباستخدام معلومات مسبقة، أو تحويل العلاقة الدالية، أو بالتخلص من واحد من المتغيرات المفسرة ذات الارتباط العالي[1]، أو استخدام طرق تقدير اخرى ولكنها متحيزة، بخلاف طريقة المربعات الصغرى غير المتحيزة، والتي يمكن ان تكون أكثر دقة، ومن أهمها[2]:-

[1] - دومينك سلفاتور، **المرجع السابق الذكر**، ص 210.

[2] -عبد المهدي رضا الجراح، **مقارنة طريقتي المربعات الصغرى والمكونات الرئيسية في تحليل الإنحدار بأستخدام أسلوب المحاكاة**، رسالة ماجستير، قسم الرياضيات، كلية الادب والعلوم، جامعة ال البيت، الاردن، 2004، ص 21.

4-1. انحدار المكونات الرئيسية : اقترحها وطورها كل من kendal & hotelling, 1967 وتهدف هذه الطريقة الى تحويل المتغيرات المرتبطة خطياً الى متغيرات جديدة غير مرتبطة خطياً.

4-2. انحدار الجذور الصماء : اقترحها وطورها كل من Webster, 1974 Hawkins, 1973 وتعتمد هذه الطريقة على استخدام القيم الذاتيه والأشعة الذاتيه في تقدير معلمات نموذج الانحدار.

4-3. انحدار الحرف : اقترحها لأول مرة Horel في عام 1962 بعد ذلك طورت من طرف كل من Horel & Kennard في عام 1970 والتي تتضمن اضافة الثابت k إلى عناصر قطر المصفوفة $(X'X)$ قبل اخذ معكوسها. وسيتم تناوله بشكل مفصل لاحقاً.[1]

ثانياً. انحدار الحرف :

اقترحها لأول مرة Horel في عام 1962 بعد ذلك طورت من طرف كل من Horel & Kennard في عام 1970 والتي تتضمن اضافة الثابت k إلى عناصر قطر المصفوفة $(X'X)$ قبل اخذ معكوسها. وسيتم تناوله بشكل مفصل لاحقاً.

1. تقدير معلمات انحدار الحرف:

يمكن توضيح طريقة تقدير معلمات انحدار الحرف من شكلين اثنين كما يلي:

1-1. تقدير معلمات انحدار الحرف في الشكل المعياري :

يتم تقدير معلمات الانحدار المعيارية وفقاً لأسلوب انحدار الحرف كالأتي[2]:

[1] E. Hoerl and Robert W. Kennard, Ridge Regression: Biased Estimation for Nonorthogonal Problems**, Technometrics**, Vol. 42, No. 1, Special 40th Anniversary Issue (Feb., 2000), pp. 80-86.
[2] Ibid.

$$\hat{\beta}^{*}_{(k)} = (X'X + kI)^{-1} X'Y; \quad 0 \le k \le 1 \qquad .....(1)$$

حيث :

$\beta^{*}_{(k)}$: شعاع معاملات انحدار الحرف المعيارية المقدرة (لا يتضمن المقدار الثابت أو مقدر المقطع)

k : كمية ثابتة تمثل مقدار التحيز وقيمتها محصورة بين الصفر والواحد الصحيح

$X'X$: مصفوفة معاملات الارتباط الجزئي البسيط ما بين المتغيرات المستقلة وبعضها البعض

I : مصفوفة الوحدة من الرتبة $p \times p$

$X'Y$: مصفوفة معاملات الارتباط الجزئي البسيط ما بين المتغيرات المستقلة والمتغير التابع

2-1. تقدير معلمات انحدار الحرف في الشكل الاصلي :

كما هو الحال بالنسبة لطريقة المربعات الصغرى عند تحويل المقدرات من الشكل المصفوفي الى الشكل الأصلي سيتم التعامل مع طريقة انحدار الحرف بنفس الشكل مع استثناء مقدر المقطع $\hat{\beta}_0(k)$ والذي يتم ايجاده بطريقة مستقلة من خلال الشكل الاصلي باستخدام العلاقة التالية[1]:

$$\hat{\beta}_i(k) = \frac{S_y}{S_i} \hat{\beta}^{*}_{i,(k)} \qquad .....(2)$$

ثم ايجاد مقدر المقطع (الجزء الثابت) كما يلي :

$$\hat{\beta}_0(k) = \bar{y} - \hat{\beta}^{*}_{1,(k)} \bar{x}_1 - \hat{\beta}^{*}_{2,(k)} \bar{x}_2 .... \hat{\beta}^{*}_{p,(k)} \bar{x}_p \qquad .....(3)$$

وعليه يتبين ان انحدار الحرف لا يستخدم مباشرة في الحصول على مقدر المقطع، ومن اشهر طرق تقدير المقطع طريقة Brown, 1977 كما يلي :

[1] -M. El-Dereny, N. I. Rashwan, Solving Multicollinearity Problem Using Ridge Regression Models, **Int. J. Contemp. Math. Sciences**, Vol. 6, 2011, no. 12, 585 – 600.

$$\hat{\beta}_0(k) = \hat{\beta}^{*r}_{,(k)} z_i; \qquad z_i = \frac{\overline{x}_i^r}{\sqrt{\sum \left(x_{ij}^r - \overline{x}_i^r\right)^2}} \qquad \ldots\ldots\ldots(4)$$

2. تحديد القيمة المناسبة لمعلمة انحدار الحرف في ظل مشكلة الازدواج الخطي:

يترتب غالباً في ظل مشكلة الازدواج الخطي كبر حجم الاخطاء المعيارية للمعلمات المقدرة نتيجة كبر قيم العناصر القطرية للمصفوفة $(x'x)^{-1}$، دلك لان[1]:

$$\text{var}(b) = s^2 (x'x)^{-1} \ldots\ldots\ldots\ldots\ldots(5)$$

يتسبب ذلك في عدم معنوية التقديرات اضافة الى عدم استقرار المعلمات المقدرة واختلاف اشارتها عند حدوث أي تغير في حجم العينة ومن الطرق الأساسية الواسعة الاستخدام في الكشف عن وجود مشكلة الازدواج الخطي حساب عوامل تضخم التباين Variance Inflation Factors (VIF's) حيث نقيس عوامل تضخم التباين مدى تضخم تباينات معاملات الانحدار المقدرة في ظل وجود الازدواج الخطي و يتم حساب عامل تضخم التباين VIF's لكل متغير مستقل على حدة كالآتي:

$$VIF = \frac{1}{1 - R_i^2}; i = 1,2,3....n \ldots\ldots\ldots\ldots\ldots(6)$$

حيث : R_i^2 معامل التحديد لنموذج انحدار المتغير المستقل i على باقي المتغيرات المستقلة حيث نوجد انحدار كل من (x_i) على بقية المتغيرات الأخرى وذلك على النحو التالي :

$$x_1 = f(x_2, x_3, \ldots, x_n) \ldots\ldots\ldots\ldots\ldots(7)$$

[1]- أحمد كامل السيد أحمد، **نماذج احصائية مقترحة لتحليل دوال الاستهلاك باستخدام بيانات السلاسل الزمنية وابحاث الدخل والانفاق والاستهلاك في مصر**، رسالة دكتوراه، قسم الإحصاء، كلية التجارة، جامعة عين شمس، 2009، ص 81.

و هكذا في كل مرة يتم حساب قيمة معامل التحديد R_i^2 يتم استخدامها في ايجاد قيمة VIF بحيث أنه اذا كانت قيمة VIF أكبر من 10 دل ذلك على وجود ازدواج خطي ويقابل هذا الحد معامل تحديد قدره 0.90 لنموذج انحدار المتغير المستقل i على بقية المتغيرات المستقلة وتعتبر قيمة ال VIF أكبر من الواحد الصحيح.

وتستخدم قيم عوامل تضخم التباين VIF's لقياس مدى بعد مقدرات المربعات الصغرى عن قيمها الحقيقية حيث تأخذ القيمة المتوقعة لمجموع مربعات انحرافات معلمات الانحدار المقدرة عن قيمها الحقيقية الصيغة التالية:

$$E\left[\sum_{i=1}^{p}\left(\hat{\beta}_i-\beta_i\right)^2\right]=\sigma^2\sum_{i=1}^{p}VIF_i=\sigma^2 p \quad \text{.....(8)}$$

ومن ثم يمكن حساب النسبة التالية :

$$\frac{\sigma^2\sum_{i=1}^{p}VIF_i}{\sigma^2 p}=\frac{\sum_{i=1}^{p}VIF_i}{p} \quad \text{.....(9)}$$

وتشير هذه النسبة الى متوسط قيم VIF's لمعاملات الانحدار المقدرة و اذا كانت المتغيرات المستقلة متعامدة أى لا يوجد بينها ارتباط فان هذه النسبة تساوي الواحد الصحيح ولذلك نجد أنه كلما زادت قيمة متوسط VIF's عن الواحد الصحيح دل ذلك على وجود ارتباط خطي بين المتغيرات المستقلة وتستخدم بعض البرامج الجاهزة معكوس VIF للكشف عن وجود الازدواج الخطي وتحديد دخول أي متغير للنموذج من عدمه و يعرف هذا المقياس بـ Tolerance حيث :

$$\text{Tolerance}=\frac{1}{VIF_i}=1-R_i^2 \quad \text{.....(10)}$$

و تعكس قيمة الثابت k مقدار التحيز في المقدرات ويلاحظ أنه عندما تكون قيمة الثابت k مساوية للصفر فإنه يمكن الحصول على مقدرات المربعات الصغرى :

$$\beta_i = (x'x)^{-1} x'y \ldots\ldots\ldots\ldots\ldots\ldots\ldots (11)$$

وعندما تكون قيمة الثابت k أكبر من الصفر نحصل على مقدرات متحيزة الا أنها أكثر استقرارا من المربعات الصغرى ويعاب على طريقة انحدار الحرف صعوبة تحديد k التي تعطى أفضل نموذج.[1]

3. التنبؤ باستخدام انحدار الحرف:

يعتبر انحدار الحرف مفيداً عند استخدامه للمحاكاة، لكن إذا كان الهدف هو التنبؤ فلا يمكن استخدام متوسط مربعات الخطأ الكلي (TMSE) كمعيار لدقة التنبؤ، وبدلاً من ذلك يمكن تفضيل استخدام $E(y-\hat{y}^*)'(y-\hat{y}^*)$ كمعيار. ويلاحظ أنه في حالة تعويض $\hat{y}^* = \hat{y}$ ، حيث $(\hat{y})$ مقدر المربعات الصغرى، ينتج أن:

$$E\sum(y-\hat{y})^2 = E(SSE) = (n-p-1)\sigma^2 \ldots\ldots\ldots\ldots\ldots\ldots (12)$$

وهذا يعني أن التنبؤ باستخدام المربعات الصغرى لا يتأثر في حالة وجود الازدواج الخطي طالما استمرت نفس الظروف في فترة التنبؤ ومن ثم فان نفس النتيجة تنطبق في حالة انحدار الحرف لكن Hoerl and Robert W. Kennard, 1970 اعتبر ان المعيار المذكور يعطي خطأ التقدير ولا يعطي خطأ التنبؤ، خاصة وان الشكل العام لمجموع متوسط مربعات الخطأ لا يتضمن التنبؤ بصورة مباشرة. وبالتالي فان هناك احتياج الى معيار اخر مختلف عند التنبؤ بخلاف المعيار المستخدم عند التقدير. وفي دراسة Park, 1981 تم استخدام

[1] أحمد كامل السيد أحمد، **المرجع السابق**، ص ص 81-84.

معيار متوسط مربعات خطأ التنبؤ MSEP عند التنبؤ باستخدام انحدار الحرف وفقا للصيغة التالية :

$$\text{MSEP}\ (\bar{y}/x) = E(x'\beta^* - x'\beta)^2 \quad \ldots\ldots\ldots\ldots\ldots\ldots (13)$$

حيث β^* تمثل An arbitrary estimator أما (x) فتمثل توليفة خطية من المتغيرات المستقلة، وقد توصلت الدراسة التي اجراها Theobald, 1974 الى ان انحدار الحرف يعطي نتائج افضل من المربعات الصغرى عندما يكون الهدف هو التنبؤ وذلك من حيث قيمة MSEP اذا كانت $0 < k < 2\sigma^2 / \beta'\beta$ ؛ وذلك لان $\text{Var}\ (\bar{y}/x) = x'x\sigma^2 / \lambda_i$ حيث λ_i تطابق القيمة المميزة. يضاف الى دلك ان التنبؤات لا يتم عملها عادة عند نفس توليفات قيم المتغيرات المستقل التي تم استخدامها في العينة للحصول على تقديرات المعالم بطريقة المربعات الصغرى؛ ودلك لان الازدواج الخطي عادة ما يحدث بين متغيرات عشوائية لا يمكن السيطرة عليها لاحقا، ولا يمكن للمربعات الصغرى حينئدا ان تعطي افضل تنبؤات. [1]

ثالثاً. الدراسة التطبيقية:

يعرض البحث فيما يلي نتائج تطبيق انحدار الحرف على بيانات السلاسل الزمنية للإنفاق الاستهلاكي الحقيقي للفرد RC_t ، والذي يمثل المتغير التابع الرئيسي في الدراسة واهم العوامل المقترحة المؤثرة فيه خلال الفترة (1970-2011) الموضحة في الجدول رقم 01، وكذلك نتائج تطبيق هذه الاختبارات على بيانات السلاسل الزمنية لمحددات الانفاق الاستهلاكي الحقيقي، والتي تشمل كل من :-

RYC_t : الدخل الحقيقي المتاح للفرد.

RW_t : الثروة الحقيقية للفرد.

[1] **نفس المرجع**، ص 86.

P_t : معدل التضخم.

RYC_t : الاستهلاك الحقيقي للفرد السابق .

1. الكشف عن الازدواج الخطي:

يمكن الكشف عن القيم الشادة بين متغيرات النموذج المستقلة، باستخدام Matrix plot الموجود في برنامج Statgraphics حتى يتم التحقق من ان مشكلة الازدواج الخطي التي يسعى نموذج انحدار الحرف لعلاجها لا تتسبب فيه القيم الشادة، ويوضح الشكل رقم 01 اشكال Matrix plot لمتغيرات الدراسة وتكشف تلك الاشكال عن عدم وجود قيم شادة بين المتغيرات المستقلة لنموذج البحث.

وبالكشف عن مصفوفة الارتباط الجزئي لمتغيرات النموذج المستقلة والموضحة في الجدول رقم 03 يتضح ان كل الارتباطات ما بين المتغيرات المستقلة قوية جداً ما عدا الارتباطات ما بين متغير التضخم والمتغيرات المستقلة الاخرى، فيما عدا ذلك فان أقلها هو الارتباط بين متغير لوغاريتم الثروة الحقيقية للفرد ومتغير لوغاريتم الانفاق الاستهلاكي الحقيقي للفرد حيث بلغ 0.804، وبالتالي فان هناك اخلال بأحد فروض تطبيق طريقة المربعات الصغرى والتي يمكن معالجتها بنموذج انحدار الحرف.

2. تقدير قيمة معلمة انحدار الحرف او ثابت التحيز (k):

للوصول الى قيمة جيدة لمعلمة انحدار الحرف تم استخدام برنامج Statgraphics في الحصول على شكل Ridge Trace ، الذي يوضح العلاقة بين قيم المعلمات المعيارية للمتغيرات المتوقع تأثيرها على لوغاريتم الانفاق الاستهلاكي الحقيقي للفرد، والقيم المختلفة لمعلمة انحدار الحرف، ويوضح الشكل رقم 02 Ridge Trace لدالة لوغاريتم الانفاق الاستهلاكي الحقيقي للفرد، ويتبين من الشكل ان افضل قيمة لمعلمة انحدار الحرف هي $k=0.01$ تقريباً؛ وذلك لانها أقل قيمة تبدأ بعدها المعلمات المعيارية في الاستقرار وقد تم اختيار هذه القيمة بعد اجراء عدة مقارنات بينها وبين القيم السابقة واللاحقة.

كما يظهر الشكل رقم 03 العلاقة بين القيم المختلفة لمعلمة انحدار الحرف وعوامل تضخم التباين VIF ويتضح من الشكل ايضا ان انسب قيمة لمعلمة انحدار الحرف هي $k=0.01$ ؛ وذلك لانها اقل قيمة تبدأ بعدها عوامل تضخم التباين في التغير ببطء، وتنخفض عندها عوامل تضخم التباين الى اقل من 10.

ويوضح الجدول 03 رقم قيم عوامل تضخم التباين عند القيم المختلفة لدالة لوغاريتم الانفاق الاستهلاكي الحقيقي للفرد ويظهر ايضا من فحص VIF عند القيم المختلفة لمعلمة انحدار الحرف الواردة بالجدول أن VIF قد اخذت في التناقص بشكل ملحوظ.

وتمثل القيم الواردة في الصف الاول من الجدول رقم 02 قيم عوامل تضخم التباين لتقديرات المربعات الصغرى، والمقابلة لقيمة معلمة انحدار الحرف منعدمة $k=0$ ويتضح من فحص هذه القيم انها تزيد عن 10، مما يدل على تأثير التقديرات بمشكلة الازدواج الخطي، وتعتبر قيمة $k=0.01$ هي اقل قيمة تبدأ بعدها عوامل تضخم التباين في التناقص ببطء ، كما جميع قيم فيف عند تلك القيمة اقل تماما من 10، مما يشير الى عدم تاثر التقديرات بوجود الازدواج الخطي.

3. تحليل نتائج تقدير انحدار الحرف:

يوضح الجدول رقم 04 نتائج تقدير دالة LRC_t بافتراض $k=0.01$ ، حيث يلاحظ على النتائج الواردة في الجدول رقم 04 ما يلي :

✓ نموذج انحدار الحرف سوف يأخذ الشكل التالي :

$$\hat{LRC_t} = 0.669 + 0.203LRY_t - 0.002LRW_t + 0.0003P_t + 0.673LRC_{t-1}$$

$S.E$	(0.062)	(0.044)	(0.001)	(0.076)
t	$(3,276)$	$(-0,047)$	$(0,2301)$	$(8,7995)$

$D.W = 1.856$ $F = 249.455$ $n = 41$ $\overline{R}^2 = 0.961$ $k = 0.01$

- ✓ يتضـح مـن النتـائج ثبـوت معنويـة المعـاملات المقـدرة الخاصـة بمتغير لوغاريتم الدخل الحقيقي المتاح للفرد ومتغير لوغاريتم الانفـاق الاسـتهلاكي الحقيقـي للفـرد فـي الفتـرة السـابقة؛ حيـث يتضـح من فحص قيمة ستودنت المحسوبة t للمعلمات المقدرة لهما أنها تزيد عن القيمة الحرجة للتوزيع الطبيعي 1.96± وذلك عند مستوى معنوية α=0.05 والقيمة الحرجة للتوزيع ستوذنت عند درجة حرية 37 ومستوى معنوية α=0.05 .
- ✓ أظهرت نتائج التقدير أن مرونة الاستهلاك لتغيرات الاستهلاك السابق تفوق المرونة الخاصة بالدخل (نمو دخل الفرد الحقيقي المتـاح للتصـرف بمعـدل 1% يـؤدي الـى زيـادة انفاقـه الحقيقـي بمعدل 0.203%، بينما يؤدي نمو الاستهلاك السابق بمعدل 1% يؤدي الى زيادة انفاقه الحقيقي بمعدل 0.673%). أما فيما يتعلق بالثروة الحقيقية للفرد LRW فقد اوضحت نتائج التقدير ان لها اثر سلبي علـى دالـة الاسـتهلاك، وتبريـر دلـك ان زيـادة الثـروة المتراكمـة لـدى بعـض الافـراد قـد يتسـبب الـى انخفـاض الميـل الحدي للاستهلاك من الثروة نتيجة اتجاه اصحاب تلك الثروات في الغالب الى مشاريع الاستثمار المختلفة.
- ✓ ثبوت معنوية النموذج ككل وفقا لاختبار F ، ودلك عند مستوى معنوية α=0.05 . كما بلغت قيمة معامل التحديد الى 96.51% تقريبا، ويشير دلك الى ارتفاع القدرة التفسيرية للنموذج ككل، اما معامـل التحديـد المعـدل المناسـب فـي حالـة زيـادة المتغيـرات التفسيرية عن اثنان واكثر فقد بلغ 96.13% تقريبا.
- ✓ بلغت قيمة احصاء دربن واطسون D.W (1.856)، ونظرا لوجود متغير مبطئ في الدالة وهو LRC_{t-1} فيمكن الكشف عن وجود

الارتباط الذاتي من خلال حساب المعامل (h) المسمى Durbin's h كما يلي[1]:

$$h=(1-\frac{d}{2})\sqrt{\frac{n}{1-nV(\delta)}}=(1-\frac{1.856}{2})\sqrt{\frac{41}{1-41\times(0.076)^2}}=0.53$$

حيث : $V(\delta)$ تباين مقدر المتغير المبطئ. يتبين ان القيمة المحسوبة اقل من القيمة الحرجة للتوزيع الطبيعي من جانب واحد والتي تساوي 1.64 عند مستوى معنوية 5%، مما يدل على خلو النموذج من مشكلة الارتباط الذاتي من الرتبة الاولى بين البواقي.

خاتمة :

يمكن تلخيص اهم النتائج المتوصل اليها في ما يلي :

- تبين أن اسلوب انحدار الحرف بديل مهم من اسلوب انحدار المربعات الصغرى؛
- مقدرات انحدار الحرف متحيزة عكس مقدرات المربعات الصغرى المتحيزة، غير ان مقدرات انحدار الحرف أفضل من ناحية المعنوية الاحصائية والاختبار الاقتصادي؛
- يتضح في النماذج القياسية ان هناك تبادل في التأثير بين الاختبارات الاحصائية من خلال معنوية المعالم المقدرة والاختبارات الاقتصادية والمتمثلة في اشارة المعالم المقدرة، وبالتالي فان نتائج الدراسات التطبيقية او العملية تعتمد بالأساس على الاسلوب الاحصائي المستخدم.

ملحق الجداول والأشكال البيانية :

الجدول رقم 01: تطور متغيرات الدراسة في الجزائر خلال الفترة 1970- 2011

[1] Isabelle Cadoret, **Centre de Recherche en Économie et Management**, http://perso.univ-rennes1.fr/isabelle.cadoret-david/econometrie/lexique/teststatistique.htm

معدل التضخم	عدد السكان الاجمالي	النقود واشباه النقود M2	الارقام القياسية للأسعار	الدخل الوطني المتاح	الاستهلاك الفردي للاسر	السنة
Pt	NPOP	Wt	PC	yt	Ct	t
6,599999999	14690855	13 076	3,346830503	22 206,70	13 443,80	1970
2,626641651	15098212	13 925	3,434739747	24 350,00	14 283,50	1971
3,65630713	15511892	18 139	3,560324381	29 075,60	16 772,10	1972
6,172839506	15935594	22 930	3,780097491	32 663,20	18 861,80	1973
4,699612403	16374671	25 772	3,957747422	50 529,70	23 924,50	1974
8,230316653	16833516	33 749	4,283482567	56 899,40	29 522,40	1975
9,430735402	17311457	43 605	4,687446474	67 532,50	34 010,50	1976
11,98928332	17808536	51 950	5,249437712	78 937,80	42 646,00	1977
17,52392345	18330535	67 458	6,169345158	95 176,40	48 793,20	1978
11,34860051	18884689	79 688	6,869479494	114 318,80	56 216,60	1979
9,517824498	19475204	93 538	7,523304496	145 017,70	67 816,50	1980
14,65484264	20103931	109 154	8,625832931	171 962,50	84 570,30	1981
6,54250963	20766568	137 889	9,190178881	184 528,00	92 648,50	1982
5,96716393	21453278	165 926	9,73857192	207 298,70	103 048,00	1983
8,116397955	22150483	194 718	10,52899317	234 886,50	122 372,50	1984
10,48228704	22847437	223 860	11,63267246	258 403,40	136 423,10	1985
12,37160917	23539383	227 017	13,07182123	261 707,10	152 195,10	1986
7,441260913	24225748	257 896	14,04452956	275 195,40	149 865,70	1987
5,911544964	24904931	292 963	14,87477824	334 347,60	208 875,70	1988
9,304361258	25576596	308 146	16,25878134	424 954,30	257 269,80	1989
16,65253439	26239708	343 005	18,96628049	543 473,60	305 042,30	1990
25,88638693	26893663	415 270	23,87596525	812 210,60	410 049,90	1991
31,66966191	27535151	515 902	31,43740272	1 023 831,50	538 844,50	1992
20,54032612	28157560	627 427	37,89474776	1 107 132,30	639 067,60	1993
29,04765612	28752749	723 514	48,90228378	1 407 803,70	826 754,50	1994
29,77962649	29315463	799 562	63,46520123	1 877 457,50	1 103 081,90	1995
18,67907586	29845208	915 058	75,31991432	2 346 682,80	1 319 393,20	1996
5,733522754	30345466	1 081 518	79,63839874	2 570 235,20	1 411 669,40	1997
4,950161638	30820435	1 592 461	83,5806282	2 590 658,20	1 531 502,90	1998
2,645511134	31276295	1 789 350	85,79176303	2 919 214,00	1 642 338,50	1999
0,339163189	31719449	2022500	86,08273711	3 755 594,70	1 684 862,80	2000
4,225988349	32150198	2473500	89,72058355	3 925 430,30	1 817 277,40	2001
1,418301923	32572977	2901500	90,99309231	4 184 666,90	1 955 242,00	2002
4,268953958	33003442	3299500	94,87754553	4 906 801,90	2 090 638,00	2003
3,961800303	33461345	3644300	98,63640441	5 730 157,10	2 333 218,50	2004
1,382446567	33960903	4070400	100	6 987 340,90	2 510 479,40	2005
2,314524087	34507214	4933700	102,3145241	7 859 227,40	2 647 004,70	2006
3,673827269	35097043	5994600	106,073383	8 867 483,80	2 908 907,00	2007

المصدر :

- الاستهلاك الفردي للأسر، الدخل الوطني المتاح، النقود واشباه النقود M2، عدد السكان الاجمالي ؛ الديوان الوطني للاحصاء، http://www.ons.dz

– الرقم القياسي للاسعار (سنة الاساس 2005)، معدل التضخم؛ البنك الدولي، http://www.albankaldawli.org

الجدول رقم 02 : مصفوفة الارتباط الجزئي للمتغيرات المستقلة في دالة LRC_t

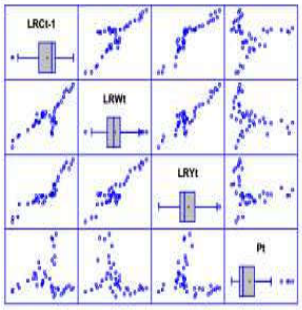
Correlations

	LRYt	LRWt	LRYt	LRCt-1
LRYt		0,9114	1,0000	0,9112
		(41)	(41)	(41)
		0,0000	0,0000	0,0000
LRWt	0,9114		0,9114	0,8040
	(41)		(41)	(41)
	0,0000		0,0000	0,0000
LRYt	1,0000	0,9114		0,9112
	(41)	(41)		(41)
	0,0000	0,0000		0,0000
LRCt-1	0,9112	0,8040	0,9112	
	(41)	(41)	(41)	
	0,0000	0,0000	0,0000	

Correlation
(Sample Size)
P-Value

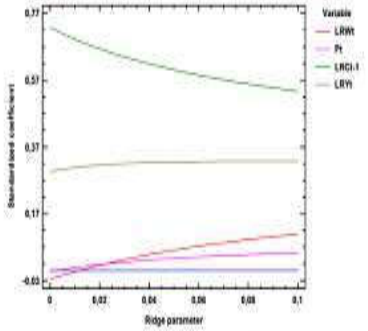
المصدر : من اعداد الباحث بالاعتماد على برنامج Statgraphics

الشكل رقم 01 : اشكال Matrix plot للمتغيرات المستقلة في دالة LRC_t



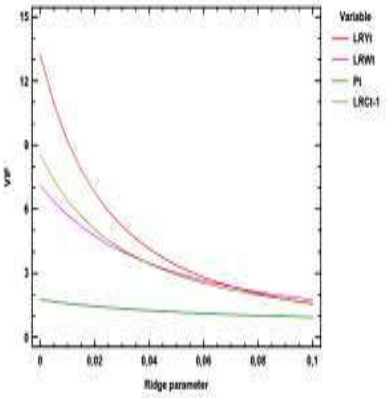
المصدر : من اعداد الباحث بالاعتماد على برنامج Statgraphics

الشكل رقم 02: Ridge Trace لدالة لوغاريتم الانفاق الاستهلاكي الحقيقي للفرد



المصدر : من اعداد الباحث بالاعتماد على برنامج Statgraphics

الشكل رقم 03: العلاقة بين VIF والقيم المختلفة لمعلمة انحدار الحرف في دالة LRC_t



المصدر : من اعداد الباحث بالاعتماد على برنامج Statgraphics

الجدول رقم 03: نتائج تقدير دالة LRC_t في حالة $k = 0.01$

R-Squared = 96,5178 percent
R-Squared (adjusted for d.f.) = 96,1309 percent
Standard Error of Est. = 0,0508679
Mean absolute error = 0,0316922
Durbin-Watson statistic = 1,85564

Ridge Regression Coefficient Section for k = 0,010000

Independent Variable	Regression Coefficient	Standard Error	Stand'zed Regression Coefficient	VIF
Intercept	0,6692156			
LRYt	0,2030668	0,06199036	0,3088	9,1867
LRWt	-0,002074801	0,04398296	-0,0035	5,7221
Pt	0,000313961	0,001364348	0,0090	1,5838
LRCt_1	0,6728307	0,07646211	0,6931	6,4147

Analysis of Variance Section for k = 0,010000

Source	DF	Sum of Squares	Mean Square	F-Ratio	Prob Level
Intercept	1	1636,201	1636,201		
Model	4	3,093518	0,7733794	249,4549	0,000000
Error	36	0,11161	0,003100277		
Total(Adjusted)	40	3,205128	0,0801282		

المصدر : من اعداد الباحث بالاعتماد على برنامج NCSS 9

الجدول رقم 03: قيم VIF المقابلة لقيم معلمة انحدار الحرف عند تقدير دالة LRC_t

Variance Inflation Factors					
k	LRYt	LRWt	Pt	LRCt-1	R-Squared
0,0	13,2156	7,1111	1,76292	8,54078	97,11
0,005	10,9241	6,3506	1,66655	7,34876	96,81
0,01	9,18669	5,72212	1,58378	6,41466	96,52
0,015	7,83784	5,19372	1,5117	5,66582	96,23
0,02	6,76967	4,74323	1,44823	5,05406	95,96
0,025	5,90926	4,35474	1,39182	4,54627	95,69
0,03	5,20596	4,01648	1,34129	4,11906	95,43
0,035	4,62367	3,71954	1,29569	3,75543	95,18
0,04	4,13608	3,45701	1,2543	3,44278	94,93
0,045	3,72368	3,22346	1,21652	3,17156	94,69
0,05	3,37175	3,01456	1,18187	2,93443	94,45
0,055	3,06898	2,82679	1,14993	2,72566	94,21
0,06	2,80662	2,65725	1,12039	2,54071	93,98
0,065	2,57775	2,50358	1,09295	2,37593	93,76
0,07	2,3769	2,36377	1,06737	2,22837	93,53
0,075	2,19966	2,23616	1,04346	2,09563	93,31
0,08	2,04245	2,11931	1,02103	1,97569	93,10
0,085	1,90236	2,01202	0,999933	1,86691	92,88
0,09	1,77699	1,91323	0,980044	1,76789	92,67
0,095	1,66433	1,82206	0,961244	1,67745	92,46
0,1	1,56271	1,73771	0,943432	1,5946	92,26

المصدر : من اعداد الباحث بالاعتماد على برنامج Statgraphics

المراجع والاحالات:

1) دومينك سلفاتور، **ملخصات شوم نظريات ومسائل في الاحصاء والاقتصاد القياسي**، ترجمة سعدية حافظ منتصر، دار ماكجرو هيل للنشر، القاهرة، 1982.

2) عبد القادر محمد عبد القادر عطية، **الحديث في الاقتصاد القياسي بين النظرية والتطبيق**، الدار الجامعية، الاسكندرية، 2005.

3) آيات عبد المهدي محمود السباح، **مقارنة بين انحدار المركبات الرئيسية والجذور الصماء بأستخدام أسلوب المحاكاة**، رسالة ماجستير، ك قسم الرياضيات، كلية العلوم، جامعة ال البيت، الاردن، 2008.

4) عبد المهدي رضا الجراح، **مقارنة طريقتي المربعات الصغرى والمكونات الرئيسية في تحليل الإنحدار بأستخدام أسلوب المحاكاة**، رسالة ماجستير، قسم الرياضيات، كلية الادب والعلوم، جامعة ال البيت، الاردن، 2004.

5) أحمد كامل السيد أحمد، **نماذج احصائية مقترحة لتحليل دوال الاستهلاك باستخدام بيانات السلاسل الزمنية وابحاث الدخل والانفاق والاستهلاك في مصر**، رسالة دكتوراه، قسم الإحصاء، كلية التجارة، جامعة عين شمس، 2009.

7) الديوان الوطني للإحصاء، **حوصلة احصائية 1962-2011**،

-.htmlحوصلة-إحصائيةwww.ons.dz/-1962-2011

8) البنك الدولي، **بيانات**،

www.albankaldawli.org/

8) Douglas C. Montgomery, George C. Runger, **Applied Statistics and Probability for Engineers**, John Wiley & Sons, Inc. USA, 2002.

9) Gunst, R. F . and Mason, R. L. **Biased estimation in regression: an evaluation using mean squared error**. J. Amer . Statist. Assoc. 72, 1977. p p 616-628.

10) Edward R. Mansfield, Billy P. Helms, Detecting Multicollinearity, **The American Statistician**, V 36, 3, 1982, p p pages 158-160.

11) E. Hoerl and Robert W. Kennard, Ridge Regression: Biased Estimation for Nonorthogonal Problems**, Technometrics**, Vol. 42, No. 1, Special 40th Anniversary Issue (Feb., 2000), pp. 80-86.

12) M. El-Dereny, N. I. Rashwan, Solving Multicollinearity Problem Using Ridge Regression Models, **Int. J. Contemp. Math. Sciences**, Vol. 6, 2011, no. 12, 585 – 600.

13) Isabelle Cadoret, **Centre de Recherche en Économie et Management**, http://perso.univ-rennes1.fr/isabelle.cadoret-david/econometrie/lexique/teststatistique.htm