اختيار افضل نموذج لقياس أثر رأس المال البشري على انتاجية العمل في قطاع الصناعة التحويلية في العراق

د. حازم منصور كوركيس　　　　　　الباحثة سها طالب عبد الرحمن
قسم الرياضيات- كلية التربية- ابن الهيثم- جامعة بغداد

الخلاصة

تم في هذا البحث تطبيق اسلوب كل الانحدارات الممكنة واسلوب الانحدار المتدرج لاختيار افضل نموذج لقياس أثر رأس المال البشري ممثلاً بمستويات مختلفة من الكوادر البشرية على انتاجية العمل في قطاع الصناعة التحويلية في العراق من خلال توظيف بيانات سلسلة زمنية مؤلفة من 21 سنة وقد استخدم البرنامج الاحصائي SPSS لاجراء الحسابات المطلوبة.

Abstract

In this paper all possible regressions procedure as well as stepwise regression procedure were applied to select the best regression equation that explain the effect of human capital represented by different levels of human cadres on the productivity of the processing industries sector in Iraq by employing the data of a time series consisting of 21 years period. The statistical program SPSS was used to perform the required calculations.

1. المقدمة

لعل من ابرز اهداف التنمية الاقتصادية باعتبارها ثورة علمية وتكنولوجية هو احداث تغيرات عميقة وشاملة في صميم الهيكل الاقتصادي والاجتماعي للبلد ويجمع المهتمون في الشأن الاقتصادي على ان التنمية لا يمكن ان تتحقق بمجرد نقل التكنولوجيا الحديثة او تحقيق مستويات عالية من تراكم رؤوس الاموال المادية او من خلال سياسات الادخار والاستثمار المتبعة بل يجب ان يرافق كل ذلك العمل على تنمية المهارات البشرية الفنية والادارية بغية تطوير النظم القائمة والاتجاهات الاجتماعية والفكرية السائدة مما يسهل استخدام وتطوير الاساليب العلمية والوسائل التكنولوجية الحديثة على نحو يتلائم وظروف الاقتصاد الوطني، [1] (الحبيب، ص 18).

ان بحثنا هذا يصب في ذلك الاتجاه. فقد حاولنا قياس أثر مستويات مختلفة من الكوادر البشرية على انتاجية العمل في قطاع الصناعة التحويلية في العراق من خلال ايجاد افضل نموذج انحدار يمثل تلك العلاقة. ولبلوغ ذلك قمنا باستخدام اسلوبين احصائيين معروفين هما اسلوب كل الانحدارات الممكنة واسلوب الانحدار المتدرج. وقد تضمن الجانب النظري من البحث شرحاً مفصلاً لهذين الاسلوبين وفي الجانب التطبيقي من البحث قمنا بالاستعانة بالبرنامج الاحصائي SPSS لتطبيق هذين الاسلوبين على بيانات المسالة قيد البحث بالنظر لما يمتاز به هذا البرنامج من سرعة وكفاءة عالية.

2. الجانب النظري

2-1. نموذج الانحدار الخطي العام

يستند نموذج الانحدار الخطي العام على افتراض وجود علاقة خطية مابين متغير معتمد (y_i) وعدد من المتغيرات المستقلة (متغيرات توضيحية) $x_1, x_2, \ldots, x_n$ بالاضافة الى الخطأ العشوائي u_i. ويعبر عن هذه العلاقة لـ n من المشاهدات و k من المتغيرات المستقلة بالشكل:

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \ldots + \beta_k x_{ik} + u_i$$
$$= \sum_{j=0}^{k} \beta_j x_{ij} + u_i, \; i = 1, 2, \ldots, n \qquad \ldots(1)$$

ومن خلال الاستعانة بالمصفوفات يمكن اعادة كتابة المعادلة (1) كالآتي:

$$Y = X\beta + u \qquad \ldots(2)$$

حيث ان:

Y يمثل متجه مشاهدات المتغير المعتمد ($n \times 1$).

X تمثل مصفوفة مشاهدات المتغيرات التوضيحية ($n \times k+1$) علماً ان العمود الأول من هذه المصفوفة يمثل الحد الثابت وعناصره كلها مساوية للواحد.

β يمثل متجه معلمات النموذج ($k+1 \times 1$) المطلوب تقديرها ويكون العنصر الأول فيه هو الحد الثابت β_0.

u يمثل متجه الاخطاء العشوائية ($n \times 1$).

ان ايجاد مقدر المربعات الصغرى الاعتيادية لمتجه المعلمات β يتطلب تحقق فروض اساسية تتمثل في ان متجه الاخطاء العشوائية u يتبع التوزيع الطبيعي بوسط حسابي مساوٍ الى المتجه الصفري ومصفوفة تباين وتباين مشترك $\sigma^2 I_n$ أي ان $u \sim N(0, \sigma^2 I_n)$ حيث ان I_n تمثل مصفوفة الوحدة (identity matrix) ($n \times n$).

ان هذه الفرضية تعني عدم وجود تباين مشترك بين الاخطاء بالنظر لكون العناصر غير القطرية من المصفوفة $\sigma^2 I_n$ والتي تمثل التباين المشترك بين الاخطاء مساوية للصفر. كما تعني هذه الفرضية ايضاً تجانس تباين الخطأ وذلك لان:

$$\sigma^2_{u_1} = \sigma^2_{u_2} = \ldots = \sigma^2_{u_n} = \sigma^2$$

اضافة الى ما تقدم يشترط عدم وجود علاقة خطية تامة او شبه تامة بين المتغيرات التوضيحية وان يكون عدد المشاهدات اكبر من عدد المعلمات المطلوب تقديرها وهذا يعني ان عدد اعمدة المصفوفة (X) في النموذج (2) والبالغ ($k + 1$) يجب ان يقل عن عدد صفوفها البالغ (n).

عند تحقق الفروض اعلاه يمكن استخدام طريقة المربعات الصغرى الاعتيادية (ordinary least squares) (OLS) لتقدير معلمات النموذج (2) والذي يمكن اعادة كتابته على الشكل:

$$u = Y - X\beta$$
$$u'u = (Y - X\beta)'(Y - X\beta)$$
$$= Y'Y - Y'X\beta - \beta'X'Y + \beta'X'X\beta$$
$$= Y'Y - 2\beta'X'Y + \beta'X'X\beta$$

وباشتقاق u'u بالنسبة الى β ومساواة المعادلة الناتجة بالصفر نحصل على:

$$\frac{\partial u'u}{\partial \beta} = -2X'Y + 2X'Xb = 0$$

ومنه نستنتج ان:

$$b_{OLS} = (X'X)^{-1}X'Y \qquad \ldots(3)$$

وبالتعويض عن Y بالمتجه $X\beta + u$ في المعادلة (3) وملاحظة ان $var(u) = \sigma^2 I_n$ ، $E(u) = O$ فانه بالامكان اثبات ان b_{OLS} هو تقدير غير متحيز الى β أي ان:

$$E(b_{OLS}) = \beta \qquad \ldots(4)$$

كما يمكن اثبات ان:

$$var - cov(b_{OLS}) = E(b - \beta)(b - \beta)' = \sigma^2 (X'X)^{-1} \qquad \ldots(5)$$

ان عناصر المتجه b_{OLS} تمثل دوال خطية للمشاهدات $y_1, y_2, \ldots y_n$ وهي تقديرات غير متحيزة ذات اقل تباين لعناصر متجه المعلمات β وعلى هذا الاساس يعتبر b_{OLS} افضل تقدير خطي غير متحيز (best linear unbiased estimator BLUE) الى β، [2] (أ.د.الحسناوي، ص 13).

2-2. مشاكل النماذج الخطية

1- عدم تجانس تباين الخطأ:

ان احدى الفرضيات الاساسية التي يتم الاعتماد عليها في تقدير معلمات نموذج الانحدار الخطي هي فرضية تجانس تباين الخطأ (homoscedasticity) وفي حالة عدم تحقق هذه الفرضية يكون تطبيق طريقة المربعات الصغرى الاعتيادية في تقدير معلمات نموذج الانحدار الخطي غير دقيق حيث ان المعلمات المقدرة بهذه الطريقة سوف لن تمتلك خاصية اقل تباين وبالتالي فان b_{OLS} سوف لن يكون افضل تقدير خطي غير متحيز الى β. بشكل عام تظهر مشكلة عدم تجانس التباين في حالة تقدير معلمات النماذج المعتمدة على بيانات مقطعية (cross section data) حيث يظهر تفاوت كبير في قيمها كما هو الحال في بيانات بحوث ميزانية الاسرة التي تضم اسراً متباينة وبشكل كبير في مستويات الدخل والانفاق، [2](أ.د. الحسناوي، ص 142).

لقد وضعت عدة اختبارات للكشف عن مشكلة عدم تجانس تباين الخطأ منها اختبار سبيرمان لارتباط الرتب واختبار Bartlett واختبار Golfeld-Qundil ولمعالجة هذه المشكلة يتم تطبيق طريقة المربعات الصغرى الموزونة (weighted least squares) في تقدير معلمات نوذج الانحدار الخطي، [3] (أ.د.كاظم، مسلم، ص 101).

2- التعدد الخطي

ان تفسير معادلة الانحدار المتعدد يعتمد ضمنا على فرضية عدم وجود علاقات خطية قوية بين المتغيرات التوضيحية. فمن المعروف ان معامل الانحدار هو مقياس للتغير الحاصل في المتغير المعتمد (متغير الاستجابة) عند زيادة المتغير التوضيحي المناظر له بمقدار وحدة واحدة مع ثبات بقية المتغيرات التوضيحية. ان هذا التفسير لن يكون صحيحاً في حالة وجود علاقات خطية تامة او شبه تامة بين المتغيرات التوضيحية. عند غياب العلاقة الخطية بين المتغيرات التوضيحية بشكل تام يقال عن هذه المتغيرات انها متعامدة (orthogonal) ولكن في اغلب تطبيقات الانحدار تكون المتغيرات التوضيحية غير متعامدة وهذه الظاهرة لاتكون من الخطورة بما يكفي للتأثير على نتائج التحليل الا في الحالات التي تكون فيها المتغيرات التوضيحية مرتبطة ارتباطاً تاماً أو شبه تام بحيث يصعب تقدير تأثير كل متغير توضيحي بشكل منفرد في معادلة الانحدار. ورياضياً يتعذر تقدير معلمات النموذج باستخدام طريقة المربعات الصغرى الاعتيادية في حالة وجود علاقة خطية تامة بين اثنين او اكثر من المتغيرات التوضيحية ويعود السبب في ذلك الى استحالة ايجاد معكوس المصفوفة (X'X) بالنظر الى ان محدد هذه المصفوفة سوف يكون مساوياً للصفر في هذه الحالة. اما اذا كانت العلاقة الخطية شبه تامة بين متغيرين توضيحيين او أكثر فانه بالامكان ايجاد مقدر المربعات الصغرى الاعتيادية غير ان هذا المقدر سوف لن يكون دقيقاً ولا يمثل واقع المشكلة المدروسة وذلك لان قيمة محدد المصفوفة (X'X) سوف يكون قريباً من الصفر وبالتالي يصبح تباين المعلمات المقدرة كبيراً جداً كما يتضح من المعادلة (5). ان ظاهرة وجود علاقة خطية بين المتغيرات التوضيحية تسمى بالتعدد الخطي (multi collinearity) ان اهم المؤشرات للكشف عن ظاهرة التعدد الخطي يتمثل بمعرفة الباحث المسبقة بطبيعة المتغيرات التوضيحية فمثلاً اذا كانت الدخول ترتفع معاً مع الاسعار فيمكن للباحث ان يتوقع وجود علاقة خطية بين هذين المتغيرين عند دراسة العوامل التي تحدد الطلب، [4](أ.د. محبوب، ص 210).

ويمكن التعرف على ظاهرة التعدد الخطي من خلال إيجاد القيم المميزة لمصفوفة ارتباط المتغيرات التوضيحية فاذا كانت بعض هذه القيم مساوية للصفر أو قريبة جداً من الصفر كان ذلك مؤشراً على وجود التعدد الخطي. ويمكن استخدام اختبار Farrar-Glober للكشف عن التعدد الخطي [2] ص 261. وهنالك اساليب عديدة لمعالجة مشكلة التعدد الخطي نذكر منها اسلوب المركبات الرئيسية (principal components) لمعالجة التعدد الخطي التام واسلوب انحدار الحرف (ridge regression) لمعالجة التعدد الخطي شبه التام، [3]، (أ. د.كاظم، ص 190).

3- الارتباط الذاتي

من الفروض الاساسية التي يتم الاعتماد عليها في تقدير معالم النموذج الخطي باستخدام طريقة المربعات الصغرى الاعتيادية هي فرضية انعدام وجود ارتباط ذاتي بين اخطاء المشاهدات المختلفة في عينة البحث. اي ان

$$cov(e_i,e_j)=0,\ i,j=1,2,\ldots,n$$

حيث تشير كل من e_i، e_j الى الاخطاء العشوائية للمشاهدات i و j على التوالي وتشير n الى عدد المشاهدات. وتظهر مشكلة الارتباط الذاتي في اغلب الدراسات التي تأخذ شكل السلاسل الزمنية وقد تنشأ هذه الظاهرة نتيجة لحذف بعض المتغيرات التوضيحية من العلاقة المدروسة اي نتيجة التشخيص غير الدقيق للعلاقة بين متغير الاستجابة والمتغيرات التوضيحية وفي حالة ظهور هذه المشكلة فان طريقة المربعات الصغرى العامة GLS (Generalized Least Squares) هي البديل المناسب لتقدير معلمات نموذج الانحدار الخطي ، [2]، (أ.د.الحسناوي، ص 210).

ويمكن صياغة فرضية عدم وجود ارتباط ذاتي بين الاخطاء العشوائية للسلسلة الزمنية المدروسة كالاتي

$$H_0:\rho=0$$

ضد الفرضية البديلة

$$H_1:\rho\neq0$$

ويستخدم لهذا الفرض اختبار ديربن واتسون .D.W وفق الصيغة:

$$D.W.=2-\frac{2\,\hat{cov}(e_i,e_{i-1})}{\hat{var}(e_i)},\ =1,2,\ldots,n \qquad \ldots(6)$$

حيث تشير e_i الى الخطأ العشوائي للمشاهدة i والتي تمثل الفرق بين القيمة الحقيقية والقيمة التقديرية لـ y_i اي ان:

$$e_i=y_i-\hat{y}_i \qquad (7)$$

وتشير $\hat{cov}(e_i,e_{i-1})$ الى القيمة التقديرية للتباين المشترك بين الاخطاء e_i و e_{i-1} وتشير $\hat{var}(e_i)$ الى القيمة التقديرية لتباين الخطأ e_i.

ويمكن كتابة صيغة ديربن واتسون كالآتي:

$$D.W.=2-2\hat{\rho}=2(1-\hat{\rho}) \qquad \ldots(8)$$

حيث تشير $\hat{\rho}$ الى القيمة التقديرية لمعامل ارتباط e_i و e_{i-1} وما دام $-1\le\hat{\rho}\le1$ فانه يمكن ملاحظة ان $0\le D.W.\le4$ ولغرض اجراء الاختبار لابد من ايجاد القيمة العليا D_u والقيمة الدنيا D_L لمعامل ديربن واتسون والموجودة في جداول خاصة محسوبة على اساس درجات الحرية n وعدد معلمات النموذج ولمستوى دلالة معين ويتم قبول فرضية العدم اذا كان: [3]، (أ. د. كاظم، ص 165).

$$d_u < D.W. < 4-d_u \qquad \ldots(9)$$

2-3.النماذج اللاخطية

قد لا تكون الظاهرة الاقتصادية قيد الدراسة ممثلة بصيغ خطية فقد تكون بصيغة متعدد حدود (polynomial) او بصيغة أسية أو جذرية أو نسبية. ومن الامثلة على ذلك دالة الانتاج لـ (Gobb-Douglas) التي تظهر بالصيغة:

$$Y_i = \beta_0 L_i^{\beta_1} K_i^{\beta_2} u_i$$

حيث ان Y_i، L_i، K_i تشير على التوالي الى حجم الناتج والاستخدام (العمالة) ورأس المال الثابت ويشير u_i الى الخطا العشوائي. ومن الواضح انه يمكن تحويل هذه الدالة الى الحالة الخطية بأخذ اللوغاريتم الطبيعي لطرفيها. في الواقع فانه على الرغم من ان الكثير من العلاقات الاقتصادية قد لا تكون خطية فان الباحث يفترض العلاقة الخطية لسببين الاول هو لتبسيط النموذج والثاني هو امكانية تحويل اغلب العلاقات غير الخطية الى خطية، [4]، (أ. د. محبوب، ص 170).

2-4. اختيار أفضل نموذج خطي

عند توفيق معادلة انحدار خطي متعدد فان بعضا من المتغيرات التوضيحية قد تكون ذات قدرة بسيطة في تفسير التغير الحاصل في متغير الاستجابة وبالتالي فهي اقل اهمية من باقي المتغيرات وتجرى محاولة حذفها من المعادلة النهائية والهدف هو الحصول على افضل معادلة اي افضل نموذج خطي يمثل العلاقة بين المتغيرات التوضيحية ومتغير الاستجابة. عندما يقرر الباحث بناء نموذج انحدار خطي متعدد لتفسير العلاقة بين متغير الاستجابة y ومجموعة مؤلفة من k من المتغيرات التوضيحية $x_1, x_2, \ldots, x_k$ فانه يجب ان ياخذ في الحسبان اعتبارين مهمين هما:

1. ان تكون معادلة الانحدار الخطي قادرة على تفسير معظم التغيرات الحاصلة في متغير الاستجابة y. وان تكون مفيدة لاغراض التنبؤ وهذا يستلزم ادخال اكبر عدد ممكن من المتغيرات التوضيحية.
2. بالنظر للكلفة العالية والوقت والجهد المبذول للحصول على معلومات عن عدد كبير من المتغيرات التوضيحية يسعى الباحث الى ان تشتمل المعادلة على اقل عدد من المتغيرات التوضيحية كلما كان ذلك ممكناً، [6] ، (Drapper, N.R. and Smith, H., p.age 294).

ان التوفيق بين هذين الاعتبارين المتناقضين هو مايسمى "اختيار افضل معادلة". وفي حقيقة الامر انه لا يوجد اسلوب احصائي وحيد لبلوغ ذلك بل ان هنالك مجموعة اساليب سوف نقتصر على ذكر اثنين منها.

1. اسلوب كل الانحدارات الممكنة All Posible Regressions Procedure

يتطلب هذا الاسلوب اولا توفيق كل معادلة انحدار ممكنة. ففي حالة وجود k من المتغيرات التوضيحية $x_1, x_2, \ldots, x_k$ فان عدد معادلات الانحدار الممكنة يكون مساو الى:

$$C_0^k + C_1^k + C_2^k + \ldots + C_k^k = (1+1)^k = 2^k \qquad \ldots(10)$$

ومن هنا نلاحظ انه كلما ازداد عدد المتغيرات التوضيحية ازداد وبشكل سريع عدد المعادلات. وتحدد اهمية كل معادلة باعتماد معايير معينة مثل معامل التحديد (Coefficient of Determination R^2) ومتوسط مربعات الخطأ (Mean Square Error S^2) وقيمة الاحصاءه (Mallow's C_p) ويتم ايجاد معامل التحديد R^2 وفق الصيغة:

$$R^2 = \frac{SSR}{SST} \qquad \ldots(11)$$

حيث ان $SST=\sum_{i=1}^{n}(y_i-\bar{y})^2$ هو مجموع المربعات الكلي لانحرافات المشاهدات عن وسطها الحسابي

وتشير **n** الى عدد المشاهدات كما ان $SSR=\sum_{i=1}^{n}(\hat{y}_i-\bar{y})^2$ هو ما يسمى مجموع المربعات العائدة للانحدار،[7](Prof.Fomby, page 2).

عند اعتماد R^2 كمعيار للمفاضلة يتم تقسيم المجموعة التي تضم كل المعادلات الى مجموعات جزئية تحتوي المجموعة الجزئية الاولى على معادلة واحدة فقط وتكون خالية من اي متغير توضيحي وتحتوي على الحد الثابت للنموذج فقط.

اما المعادلات في المجموعة الثانية فانها تحتوي على متغير توضيحي واحد ومعادلات المجموعة الثالثة تحتوي على متغيرين توضيحيين وهكذا. ثم يتم ترتيب المعادلات داخل كل مجموعة بالاعتماد على قيمة معامل التحديد R^2 وبعد ذلك يتم اختيار المعادلة التي تتحقق من خلالها اعلى قيمة لـ R^2 ضمن كل مجموعة. ثم يترك للباحث حرية اختيار المعادلة الافضل من بين المعادلات المنتخبة من كل المجموعات وذلك بالاعتماد على طبيعة المسالة قيد التحليل وطبيعة العلاقة بين المتغيرات التوضيحية وقوة الارتباط فيما بينها. ويمكن ان تكون مصفوفة معاملات الارتباط مفيدة لهذا الغرض [5]، (د.الراوي، ص 262).

عند اعتماد متوسط مربعات الخطأ S^2 كمعيار للمفاضلة فان المعادلة التي يتم اختيارها من كل مجموعة هي تلك التي يتحقق من خلالها اقل قيمة لـ S^2 ثم تتم المفاضلة بين المعادلات المنتخبة. وجدير بالذكر هنا انه عند اضافة المزيد من المتغيرات التوضيحية الى المعادلة فان متوسط مربعات الخطا يميل الى الاستقرار ويقترب من القيمة الحقيقية لتباين الخطا σ^2 [8]،(Carside,M.J., page 112). ان المعيار البديل للمفاضلة يتمثل بقيمة الاحصاءه **Mallow's C_p** التي اقترحها **C.L.Mallow's [6]**، (Drapper, N.R. and Smith, H., page 299) وتظهر بالصيغة

$$C_p=\frac{RSS_p}{S^2}-(n-2p)\qquad ...(12)$$

حيث RSS_p تمثل مجموع مربعات الخطأ للنموذج الذي يحتوي على **p** من المعلمات (بضمنها الحد الثابت β_0) وان S^2 يمثل متوسط مربعات الخطأ من المعادلة التي تحتوي على جميع المتغيرات التوضيحية والذي يفترض ان يكون تقديرا غير متحيز لتباين الخطأ σ^2 وحيث ان:

$$E(RSS_p)=(n-p)\,\sigma^2\qquad ...(13)$$

فانه وبشكل تقريبي يكون

$$E(RSS_p/S^2)=n-p$$

وبالتالي يكون وبشكل تقريبي ايضاً [6, page 300].

$$E(C_p)=p\qquad ...(14)$$

لذلك وعند اعتماد C_p كمعيار للمفاضلة فان معادلة الانحدار الافضل هي تلك المعادلة ذات القيمة الاقل لـ C_p والاكثر قربا من **p**.

2. طريقة الانحدار المتدرج Stepwise Regression Method

بموجب هذه الطريقة يتم ادخال المتغيرات التوضيحية الى معادلة الانحدار على التعاقب وحسب اهميتها ومؤشر ذلك هو قيمة **F** الجزئية (او قيمة معامل الارتباط الجزئي) ففي الخطوة الاولى يتم ايجاد معادلة الانحدار الخطي البسيط عند كل متغير من المتغيرات التوضيحية وتحتسب قيمة **F** من جدول تحليل التباين والمتغير التوضيحي الذي يدخل اولا الى معادلة الانحدار هو ذلك المتغير الذي يحقق اعلى قيمة لـ **F** المحتسبة والتي يجب ان تكون معنوية من خلال مقارنتها مع قيمة **F** الجدولية بمستوى دلالة يحدده الباحث في الخطوة التالية يتم ايجاد معادلة الانحدار عند كل متغير من المتغيرات التوضيحية المتبقية بوجود المتغير الذي يتم اختياره في الخطوة الاولى والمتغير الثاني الذي يدخل معادلة الانحدار هو ذلك الذي يحقق اعلى قيمة لـ **F** الجزئية والمحتسبة من جدول تحليل التباين والتي يجب ان تكون معنوية وهكذا نستمر في ادخال المتغيرات التوضيحية الى ان يتم الحصول على افضل معادلة انحدار وجدير بالذكر هنا الى ان جميع المتغيرات التوضيحية التي ادخلت الى المعادلة يحسب لها **F** جزئية في كل خطوة ويتم التقييم على اساسها مرة اخرى وذلك لانه عند الادخال المبكر لاحد المتغيرات التوضيحية الى معادلة الانحدار قد يحقق احياناً **F** جزئية غير معنوية في المراحل المتاخرة مما يستوجب حذفه من المعادلة.

3- الجانب التطبيقي

في هذا الجانب من البحث تم تطبيق اسلوب كل الانحدارات الممكنة اضافة الى اسلوب الانحدار المتدرج للحصول على افضل نموذج خطي يمثل العلاقة بين مستويات مختلفة من الكوادر البشرية وانتاجية العمل في قطاع الصناعة التحويلية في العراق.

وفي هذا الجانب واجهتنا مشكلة تتمثل في عدم قدرتنا على الحصول على بيانات دقيقة وشاملة عن متغيرات النموذج للسنوات التي تلت عام 1990 ويعود السبب في ذلك الى ظروف الحصار الاقتصادي وما تلا ذلك من احداث في عام 2003 والتي الحقت ضررا بالغاً في كافة القطاعات الاقتصادية ولا سيما قطاع الصناعة التحويلية. وفي هذا الشأن يشير الدكتور عادل عبد الغني محبوب الى ان احدى مشاكل البيانات في الاقتصاد القياسي تتمثل في مشكلة "التغير الهيكلي" إذ قد يكون هناك تغير طارىء في العالم الحقيقي بحيث لاتمثل البيانات المجتمع المقصود مثل فترة الحرب التي طالما يتم استبعادها من السلسلة الزمنية على أساس انها لا تصلح للتمثيل، [4] (أ.د.محبوب، ص 188). ومن أجل ذلك قمنا بتوظيف بيانات تم الحصول عليها من الجهاز المركزي للاحصاء وتمثل نتائج الاحصاء الصناعي للمنشأت الصناعية الكبيرة في القطاع العام والتي توفرت فقط للسنوات 1990-1970.

لقد افترضنا ان متغير الاستجابة y يمثل انتاجيه العمل مقاسة بطريقة القيمة المضافة وهو يمثل دالة خطية في خمسة متغيرات توضيحية ترمز الى المستويات المختلفة من الكوادر البشرية حيث تمثل x_1 اعداد الكوادر الادارية و x_2 اعداد الكوادر الفنية والهندسية و x_3 ترمز الى اعداد عمال الخدمات و x_4 اعداد العمال غير الماهرين وتمثل x_5 اعداد العمال الماهرين. وقد تم عرض البيانات في الجدول رقم (1).

وقبل البدء بالتحليل ينبغي التطرق الى المشاكل التي قد يعاني منها النموذج المفترض والتي أشرنا اليها في البند 2-2 من هذا البحث. ان المشكلة الرئيسية التي تظهر في الدراسات القياسية التي تأخذ شكل السلاسل الزمنية كما هو الحال في المسألة قيد البحث هي مشكلة الارتباط الذاتي. ولأجل ذلك قمنا باجراء اختبار ديربن واتسون لاختبار فرضية عدم وجود ارتباط ذاتي بين الاخطاء العشوائية للسلسلة الزمنية المدروسة. حيث قمنا بايجاد قيمة معامل ديربن واتسون وفق الصيغة (6) وكانت مساوية الى 1.891. ومن جداول خاصة تم ايجاد قيم d_L، d_u بمستوى دلالة 5% ودرجات حرية 21 و 5 وكانت قيمها مساوية الى 0.829 و 1.864 على التوالي. وبذلك تكون قيمة D.W. واقعة ضمن منطقة قبول فرضية العدم (المتباينة 9) مما يعني ان النموذج لا يعاني من مشكلة الارتباط الذاتي. ان مشكلة عدم تجانس تباين الخطأ تظهر غالباً في الدراسات التي تعتمد على بيانات مقطعية. اما مشكلة التعدد الخطي فانها تكون ملازمة لجميع النماذج القياسية تقريباً غير انها لا تكون من الخطورة بما يكفي للتأثير على نتائج التحليل الا اذا كانت تأخذ شكل التعدد الخطي التام أو شبه التام كما بينا ذلك في الجانب النظري ولغرض الكشف عن التعدد الخطي تم وضع المتغيرات التوضيحية بالصيغة القياسية من خلال طرح الوسط الحسابي والقسمة على الانحراف المعياري لكل متغير. وبذلك تتحول مصفوفة المعلومات $X'X$ الى مصفوفة ارتباط المتغيرات التوضيحية (جدول رقم 2).

ان متوسط مربعات الخطأ لمقدر المربعات الصغرى b_{OLS} يمكن ايجاده وفق الصيغة:

$$\mathrm{MSE}(b_{OLS}) = \sigma^L \,\mathrm{trace}(X'X)^{-1} = \sigma^L \sum_{i=1}^{p} \frac{1}{\lambda_i} \quad \ldots(15)$$

وهو عبارة عن مجموع العناصر القطرية في مصفوفة التباين والتباين المشترك لمقدر المربعات الصغرى b_{OLS} وتشير $\lambda_i, i=1,2,\ldots,p$ الى القيم المميزة (eigen values) لمصفوفة ارتباط المتغيرات التوضيحية (Drapper, N.R. and Smith, H., page 313) [6].

وواضح ان الصيغة (15) لا يمكن تطبيقها اذا كانت واحدة أو أكثر من القيم المميزة مساوية للصفر مما يدل على وجود تعدد خطي تام بين المتغيرات التوضيحية. اما التعدد الخطي شبه التام فبالامكان تشخيصه اذا كانت بعض أو جميع القيم المميزة قريبة جداً من الصفر. لقد قمنا بحساب القيم المميزة لمصفوفة ارتباط المتغيرات التوضيحية وتم ترتيبها تنازلياً كالآتي:

3.623598, 1.051811, 0.211248, 0.104043, 0.00930

ويلاحظ عدم وجود قيمة مميزة مساوية للصفر مما يعني عدم وجود تعدد خطي تام بين المتغيرات التوضيحية بحيث يتعذر معه اجراء تحليل الانحدار.

وبغية تطبيق اسلوب كل الانحدارات الممكنة ومن خلال الاستعانة بالبرنامج الاحصائي الجاهز SPSS تم ايجاد جميع معادلات الانحدار لـ $32 = 2^5$ من المجموعات الجزئية للمتغيرات التوضيحية وقد تم تقدير المعلمات باستخدام طريقة المربعات الصغرى الاعتيادية OLS وتم اعتماد ثلاثة معايير للمفاضلة وهي معامل التحديد R^2 ومتوسط مربعات الخطأ S^2 وقيمة الاحصاءه C_p وتم عرض النتائج في الجدول رقم (3).

ان الخطوة التالية تتمثل بتحديد المعادلة التي تحقق اعلى قيمة لـ R^2 واقل قيمة لكل من S^2 و C_p من كل مجموعة وهذا ما تم عرضه في الجدول رقم (4).

ويتضح من هذا الجدول انه عند مقارنة R^2 للمعادلات المنتخبة من كل مجموعة نرجح اختيار المعادلة المنتخبة من المجموعة C ذات المتغيرين التوضيحيين x_2 و x_5 حيث انه بامكان هذين المتغيرين تفسير 71% من التغيرات الحاصلة في متغير الاستجابة y وان اضافة اي من المتغيرات التوضيحية المتبقية (x_1 و x_3 و x_4) او بعض منها او جميعها الى هذه المعادلة سوف لن يحقق الا زيادة طفيفة في قيمة R^2 الامر الذي لا يبرر اضافتها. كما يلاحظ ايضاً ان الفرق ضئيل بين قيمة متوسط مربعات الخطأ S^2 لهذه المعادلة وتلك المنتخبة من المجموعات D و E و F والتي تحتوي عددا اكبر من المتغيرات التوضيحية. كما ان هذه المعادلة تحقق اقل قيمة للاحصاءه C_p وهذا كله يبرر اختيارنا لهذه المعادلة لتكون هي الافضل وعليه فان افضل نموذج خطي بموجب هذه الطريقة يتمثل بمعادلة الانحدار

$$y=833.915 + 0.0566x_2 + 0.02568x_5 \quad \ldots(16)$$

عند تطبيق اسلوب الانحدار المتدرج فان الخطوة الاولى تتمثل بحساب قيمة F لانحدار y على كل من المتغيرات التوضيحية (x_i, $i=1,2,\ldots,5$) وقد لاحظنا ان اعلى قيمة لـ F هي 35.236 والتي تم احتسابها من جدول تحليل التباين لانحدار y على المتغير التوضيحي x_5 (جدول رقم 5) وهذه القيمة هي اكبر من قيمة $F(0.10,1,19)$ الجدولية والبالغة 2.99 وبذلك يكون x_5 هو المتغير التوضيحي الاول الذي يدخل المعادلة. ان الخطوة التالية هي احتساب قيمة F الجزئية لانحدار y على كل متغير من المتغيرات التوضيحية المتبقية مع وجود المتغير x_5 وقد لاحظنا ان اعلى قيمة لـ F الجزئية هي 3.7097 والتي تم احتسابها من جدول تحليل التباين لانحدار y على x_2 بوجود x_5 (جدول رقم 6) وهي اكبر من قيمة $F(0.10,1,18)$ الجدولية والبالغة 3.01 وبذلك يكون x_2 هو المتغير الثاني الذي يدخل المعادلة. وقبل ترشيح متغير ثالث لدخول المعادلة لابد من التأكد من ان تأثير المتغير الذي ادخل اولاً x_5 لايزال معنوياً بوجود x_2 وهذا يتم من خلال احتساب قيمة F الجزئية لانحدار y علىx_5 بوجود x_2 (جدول رقم 7) وقد لاحظنا ان هذه القيمة كانت مساوية الى 9.315 وهي اكبر من قيمة $F(0.10,1,18)$ والبالغة 3.01 لذا فان المتغير x_5 يبقى في المعادلة. وبغية ترشيح متغير ثالث لدخول المعادلة لابد من ايجاد قيم F الجزئية لانحدار y على كل من المتغيرات التوضيحية المتبقية (x_1,x_3,x_4) بوجود المتغيرين x_2 و x_5 (الجداول 8 و 9 و10).

وقد لاحظنا ان جميع هذه القيم هي اصغر من قيمة $F(0.10,1,17)$ الجدولية والبالغة 3.03 وهذا يعني ان تاثير بقية المتغيرات غير معنوي عند مستوى دلالة 0.10 الامر الذي لا يستوجب ادخال اي منها الى المعادلة. وبذلك يكون افضل نموذج خطي بموجب هذه الطريقة هو ذات النموذج الذي توصلنا اليه من تطبيق طريقة كل الانحدارات الممكنة والمتمثل بالمعادلة (16). لاختبار وجود ارتباط ذاتي بين الاخطاء تم حساب قيمة D.W. للمعادلة (16) التي اختيرت كأفضل معادلة وكان مساويةً الى 1.583 ومن جداول خاصة تم ايجاد قيم d_L و d_u بمستوى دلالة 5% ودرجات حرية 2.21 وكانت قيمها على التوالي 1.125 و 1.539 وبذلك تكون D.W. واقعة ضمن منطقة قبول فرضية العدم كما يتضح ذلك من المتباينة في (9) ونستنتج من ذلك ان النموذج لا يعاني من مشكلة الارتباط الذاتي.

الاستنتاجات

لقد توصلنا في هذا البحث الى نتيجة منطقية وهي ان الكوادر الفنية والهندسية واعداد العمال الماهرين هي الاكثر قدرة على تفسير التغيرات الحاصلة في انتاجية العمل وبالتالي فانها تعتبر المتغيرات الاكثر اهمية مما يستوجب الاهتمام بها من حيث تطوير قدراتها ورفع مستوى مهاراتها من خلال عمليات التدريب والتطوير المستمر.

في الجانب الاحصائي تم استخدام اسلوبين مختلفين للحصول على افضل معادلة انحدار خطي هما اسلوب كل الانحدارات الممكنة واسلوب الانحدار المتدرج. ان هذه الاساليب ليست هي الوحيدة لايجاد افضل نموذج خطي بل ان هنالك اساليب اخرى نذكر منها اسلوب الاختيار الامامي (Forward Selection Procedure) واسلوب الحذف الخلفي (Backward Elimination Procedure) وعلى الرغم من أن أتباع أي من هذه الاساليب يؤدي في الغالب الى الحصول على نفس المعادلة كما هو الحال في المسألة قيد البحث غير أنه في العديد من الحالات قد يؤدي أتباع أساليب مختلفة الى الحصول على معادلات مختلفة.

ان طريقة كل الانحدارات الممكنة لم يكن بالامكان استخدامها لو لا التطور الهائل الذي تحقق في مجال الحاسبات الالكترونية من حيث الكفاءة العالية والسرعة الفائقة وعلى اية حال فان اسلوب كل الانحدارات الممكنة لا يوفر جوابا قاطعا حول المعادلة الافضل التي ينبغي اختيارها بل يبقى للباحث قدر من المرونة في اختيار المعادلة الافضل بالاعتماد على طبيعة المسالة قيد البحث وطبيعة العلاقة بين المتغيرات التوضيحية.

ان طريقة الانحدار المتدرج والتي تجمع بين طريقتي الاختيار الامامي والحذف الخلفي هي الاكثر اسخداما في السنوات الاخيرة. وبموجب هذه الطريقة فانه لا يتم التعامل مع جميع المتغيرات التوضيحية بل ان هذه المتغيرات يتم ادخالها الى المعادلة وفق قواعد معينة سبق ذكرها. وبذلك تكون هذه الطريقة هي اكثر اختصاراً واكثر قدرة على تشخيص المعادلة الافضل.

جدول رقم (1)*

إعداد الكوادر البشرية المختلفة وانتاجية العمل في قطاع الصناعة التحويلية / النشاط الاشتراكي في العراق للسنوات 1990-1970

انتاجية العمل y	الكوادر الادارية x_1	الكوادر الفنية والهندسية x_2	عمال الخدمات x_3	العمال غير الماهرين x_4	العمال الماهرين x_5
1271	1516	473	24524	50096	14216
1440	1514	1923	26012	57133	16467
1355	2037	2490	30360	66795	18697
1509	2019	2734	29385	66263	18716
1254	2210	3322	36619	61718	20696
1220	2234	3563	37393	61028	23474
1546	2525	4205	40034	69344	25283
1916	2381	4293	10310	66181	26172
2381	2944	5278	49163	62039	28108
2585	3528	8529	57299	69339	30707
2810	4308	9661	60965	69671	32561
1440	4444	9840	56600	69047	33609
2493	4588	10830	55789	65441	32789
3285	4075	12968	44125	44175	52340
3062	4563	12875	46492	52476	47633
3403	4031	12163	46599	48322	63398
2875	4479	12712	49070	49672	62205
2861	4343	12610	49035	47330	47530
2596	4299	12615	48013	46160	46260
1710	4345	12630	40860	45123	4510
1777	4865	12955	49095	44925	44915

* المصدر: الجهاز المركزي للاحصاء.

جدول رقم (2)
مصفوفة ارتباط المتغيرات التوضيحية

	x_1	x_2	x_3	x_4	x_5
x_1	1.000000	0.967806	0.780634	- 0.380475	0.821688
x_2	0.967806	1.000000	0.688120	- 0.541955	0.912351
x_3	0.780634	0.688120	1.000000	- 0.028945	0.513239
x_4	- 0.380475	- 0.541955	- 0.028945	1.000000	- 0.645230
x_5	0.821688	0.912351	0.513239	- 0.645230	1.000000

جدول رقم (3)
قيم مقاييس المفاضلة للمعادلات الانحدارية المختلفة

المجموعة	المتغيرات التوضيحية في المعادلة	R_p^2	S_p^2	C_p
A	لايوجد	0	547671.65	48.7673
B	x_1	0.453	315072.039	20.0368
	x_2	0.559	254104.574	12.8701
	x_3	0.286	411750.267	31.4013
	x_4	0.136	497969.765	41.5365
	x_5	0.650	201960.085	6.7407
C	x_1,x_2	0.599	243884.048	12.1598
	x_1,x_3	0.454	332350.832	22.0117
	x_1,x_4	0.468	323515.309	21.0275
	x_1,x_5	0.687	190693.095	6.2351
	x_2,x_3	0.560	267567.583	14.7972
	x_2,x_4	0.561	267091.806	14.7442
	x_2,x_5	0.710	176752.195	4.6837
	x_3,x_4	0.412	357761.928	24.8416
	x_3,x_5	0.677	196772.628	6.9133
	x_4,x_5	0.650	212847.966	8.7035
D	x_1,x_2,x_3	0.640	231880.485	11.3884
	x_1,x_2,x_4	0.659	219629.565	10.0999
	x_1,x_2,x_5	0.724	177869.108	5.7077
	x_1,x_3,x_4	0.476	337866.048	22.5356
	x_1,x_3,x_5	0.688	200715.009	8.1105
	x_1,x_4,x_5	0.687	201888.306	8.2339
	x_2,x_3,x_4	0.561	282733.231	16.7369
	x_2,x_3,x_5	0.711	186333.130	6.5979
	x_2,x_4,x_5	0.715	183787.882	6.3302
	x_3,x_4,x_5	0.681	205763.152	8.6415
E	x_1,x_2,x_3,x_4	0.681	218233.700	10.6029
	x_1,x_2,x_3,x_5	0.745	174435.051	6.2673
	x_1,x_2,x_4,x_5	0.768	158530.775	4.6929
	x_1,x_3,x_4,x_5	0.689	213042.824	10.0890
	x_2,x_3,x_4,x_5	0.715	195249.793	8.3277
F	x_1,x_2,x_3,x_4,x_5	0.779	161633.036	6.0000

جدول رقم (4)

اعلى قيم لـ R^2 و S^2 و C_p من كل مجموعة

المجموعة	المتغيرات التوضيحية في المعادلة	R^2	S^2	C_p
A	لايوجد	0	547671.65	48.7673
B	x_5	0.650	201960.085	6.7405
C	x_2,x_5	0.710	176752.195	4.6837
D	x_1,x_2,x_5	0.724	177869.108	5.7077
E	x_1,x_2,x_4,x_5	0.768	158530.775	4.6929
F	x_1,x_2,x_3,x_4,x_5	0.779	161633.036	6.0000

جدول رقم (5)

تحليل التباين لانحدار y على x_5

Source of variation	df	Sum of squares	Mean square	F
$R(x_5)$	1	7116191.6	7116191.6	35.236
$error(x_5)$	19	3837241.6	201960.085	
Total	20	10953433		

جدول رقم (6)

تحليل التباين لانحدار y على x_2 بوجود x_5

Source of variation	df	Sum of squares	Mean square	F
$R(x_2,x_5)$	2	7771893.7	3885946.8	
$R(x_5)$	1	7116191.6	7116191.6	
$R(x_2\backslash x_5)$	1	655702.1	655702.1	3.7097
$error(x_2\backslash x_5)$	18	3181539.5	176752.19	
Total	20	10953433		

جدول رقم (7)

تحليل التباين لانحدار y على x_5 بوجود x_2

Source of variation	df	Sum of squares	Mean square	F
$R(x_2,x_5)$	2	7771893.7	3885946.8	
$R(x_2)$	1	6125446.3	6125446.3	
$R(x_5\backslash x_2)$	1	1646447.4	1646447.4	9.315
$error(x_2,x_5)$	18	3181539.5	176725.19	
Total	20	10953433		

جدول رقم (8)

تحليل التباين لانحدار y على x_1 بوجود x_2 و x_5

Source of variation	df	Sum of squares	Mean square	F
$R(x_1,x_2,x_5)$	3	7929658.4	2643219.46	
$R(x_2,x_5)$	2	7771893.7	3885946.86	
$R(x_1\backslash x_2,x_5)$	1	157764.7	157764.7	0.8869
$error(x_1,x_2,x_5)$	17	3023774.8	177869.108	
Total	20	10953433		

جدول رقم (9)

تحليل التباين لانحدار y على x_3 بوجود x_2 و x_5

Source of variation	df	Sum of squares	Mean square	F
$R(x_2,x_3,x_5)$	3	7785770.0	2595256.67	
$R(x_2,x_5)$	2	7771893.7	3885946.86	
$R(x_3\backslash x_2,x_5)$	1	13876.3	13876.3	0.07447
$error(x_2,x_3,x_5)$	17	3167663.2	186333.13	
Total	20	10953433		

جدول رقم (10)

تحليل التباين لانحدار y على x_4 بوجود x_2 و x_5

Source of variation	df	Sum of squares	Mean square	F
$R(x_2,x_3,x_5)$	3	7829039.2	2609679.74	
$R(x_2,x_5)$	2	7771893.7	3885946.86	
$R(x_4\backslash x_2,x_5)$	1	57145.5	57145.5	0.31093
$error(x_2,x_4,x_5)$	17	3124394.0	183787.88	
Total	20	10953433		

المصادر

1. الحبيب، مصدق جميل.(1981). "التعليم والتنمية الاقتصادية"، دار الرشيد للنشر.
2. أ. د. الحسناوي، اموري هادي كاظم. (2002). "طرق القياس الاقتصادي"، دار وائل للنشر، عمان– الاردن.
3. أ. د. كاظم، أموري هادي، مسلم، باسم شليبه. (2002). "القياس الاقتصادي المتقدم– النظرية والتطبيق"، مطبعة الطيف، بغداد – العراق.
4. أ.د. محبوب، عادل عبد الغني. (1998). "اصول الاقتصاد القياسي– النظرية والتطبيق"، شركة الاعتدال للطباعة الفنية المحدودة. بغداد – العراق.
5. د. الراوي، خاشع محمود.(1987)." المدخل الى تحليل الانحدار"، جامعة الموصل.
6. Drapper, N.R. and Smith, H. (1981). "Applied Regression Analysis", Second Edition, John Wiley and Sons, New York.
7. Prof. Fomby, T. (2008). "Multiple Linear Regression and Subset Selection", Southern Methodist University. Dallas. TX 75275.
8. Carside, M.J. (1971). "Best Subset Search", Applied Statistics 20, 112-115.