Ibn Al-Haitham Journal for Pure and Applied Science

مجلة إبن الهيثم للعلوم الصرفة و التطبيقية

No. 2 Vol. 25 Year 2012

2012 السنة 25 المجلد 2 العدد

Using Restricted Least Squares Method to Estimate and Analyze the Cobb-Douglas Production Function with Application

H. M. Gorgess, A. I. Abdul Naby

Department of Mathematics , College of Education -Ibn-Al-Haitham -, University of Baghdad

Received in : 8 September 2011 Accepted in : 18 October 2011

Abstract

In this paper, the restricted least squares method is employed to estimate the parameters of the Cobb-Douglas production function and then analyze and interprete the results obtained.

A practical application is performed on the state company for leather industries in Iraq for the period (1990-2010).

The statistical program SPSS is used to perform the required calculations.

Key Words: General linear model, Cobb-Douglas production function, OLS estimators, RLS estimators, multicollinearity.

1. Introduction

A production function is a mathematical description of the various production activities faced by a firm. Algebraically, It can be written as: [1]

$$q = f(x_1,x_2,\ldots,x_n) \quad \ldots(1)$$

where q represents the flow of output produced and $x_1,x_2,\ldots,x_n$ are the flows of inputs, each measured in physical quantities. Often production functions appear in literature written with two inputs as $q = f(K,L)$ where K denotes the amount of capital and L denotes the amount of labor. Equation (1) is assumed to provide for any conceivable set of inputs, the solution to the problem of how to best (more efficiently) combine different quantities of those inputs to get the output. However, the key question from an economic point of view is how the levels of output and inputs are chosen by firms to maximize profits.

Thus, economists use production function in conjunction with marginal productivity theory to provide explanations of factor prices and the levels of factor utilization.

The marginal productivity of an input is the additional output that can be produced by employing one more unit of the input while holding all other inputs constant [2].

Algebraically, $\frac{\partial q}{\partial K}$ is the marginal productivity of capital and $\frac{\partial q}{\partial L}$ is the marginal productivity of labor. It is assumed that both marginal productivities are positive, that is $\frac{\partial q}{\partial K} > 0$, $\frac{\partial q}{\partial L} > 0$.

The negative marginal productivity means that using more of the input results in less output being produced.

It is also usually assumed that the production process exhibits diminishing marginal productivity. This means that successive additions of one factor while keeping the other one constant yields smaller and smaller increases of output, that is:

$\frac{\partial^2 q}{\partial K^2} < 0$ and $\frac{\partial^2 q}{\partial L^2} < 0$.

Factor elasticity is the percentage change in output in response to an infinitesimal percentage change in a factor given that all other factors are held fixed [2], that is:

$$e_L = \frac{\partial q}{\partial L}\frac{L}{q}, \quad e_K = \frac{\partial q}{\partial K}\frac{K}{q}$$

where e_L, e_K represent the factor elasticity of labor and capital respectively.

2. The Cobb-Douglas Production Function

In economics the Cobb-Douglas functional form of production function is widely used to represent the relationship of an output to inputs. It was proposed by Kunt Wicksell (1851-1962), and tested against statistical evidence by Charles Cobb and Paul Douglas in 1928 [3]. In 1928 Charles Cobb and Paul Douglas published a study in which they modeled the growth of the American economy during the period 1899-1922. They considered a simplified view of the economy in which production output is determined by the amount of labor involved and the amount of capital invested. While there are many other factors affecting economic performance, their model proved to be remarkably accurate.

The function they used to model production was of the form

$$p(L,K) = b\, L^{\alpha} K^{\beta} \quad \ldots(2)$$

where

p = total production (the monetary value of all goods produced in a year).
L = labor input (the total number of persons-hours worked in a year)
K = capital input (the monetary worth of all machinery, equipment, and buildings)
b = total factor productivity

α and β are the output elasticities of labor and capital, respectively. Each b, α, β are the parameters that must be estimated by using suitable method of estimation. The property of production that examines changes in output subsequent to a proportional change in all inputs (where all inputs increase by a constant factor) is refereed to as returns to scale. If output increases by the same proportional change in all inputs, then there are constant returns to scale (CRTS). If output increases by less than that proportional change, there are decreasing returns to scale (DRTS). If output increases by more than that proportion there are increasing returns to scale (IRTS).

However, if $\alpha + \beta = 1$ the production function has constant returns to scale. If $\alpha + \beta < 1$, returns to scale are decreasing, and if $\alpha + \beta > 1$ returns to scale are increasing [4].

In our study we focus our attention on the Cobb-Douglas production function with (CRTS).

The assumptions made by Cobb and Douglas can be stated as follows:

1- If either labor or capital vanishes, then so will production, that is $p(K,0) = p(0,L) = 0$.

2- The marginal productivity of labor is proportional to the amount of production per unit of labor, that is $\frac{\partial p}{\partial L} \propto \frac{p}{L}$.

3- The marginal productivity of capital is proportional to the amount of production per unit of capital, that is $\frac{\partial p}{\partial K} \propto \frac{p}{K}$.

4- $\beta_j > 0$, $j = 1, 2$.

3. Deriving the Cobb-Douglas Production Function

Since the production per unit of labor is $\frac{p}{L}$, then according to assumption 2 we have $\frac{\partial p}{\partial L} = \alpha \frac{p}{L}$ for some constant α.

If we keep K constant ($K = K_0$) then this partial differential equation will become an ordinary first order separable differential equation $\frac{dp}{dL} = \alpha \frac{p}{L}$.

Re-arranging the terms and integrating both sides we obtain:

$\int \frac{1}{p} dp = \alpha \int \frac{1}{L} dL$ and this yields:

$\ln p = \alpha \ln L + c_1(K_0)$, then

$$p(L,K_0) = c_1(K_0)L^{\alpha} \quad \ldots(3)$$

where $c_1(K_0)$ is the arbitrary constant of integration and we write it as a function of K_0 since it could depend on the value of K_0.

Similarly, assumption 3 says that:

$\frac{\partial p}{\partial K} = \beta \frac{p}{K}$ for some constant β, keeping L constant ($L = L_0$), this differential equation can be solved to get:

$$p(L_0,K) = c_2(L_0)K^{\beta} \quad \ldots(4)$$

And finally, combining equations (3) and (4) to obtain equation 2 which is, [2]:

$p(L,K) = b L^{\alpha} K^{\beta}$

where b is a constant that is independent of both L and K.

Notice from equation (2) that if labor and capital are both increased by a factor m, then:

$p(mL,mK) = b(mL)^{\alpha}(mK)^{\beta} = m^{\alpha+\beta} b L^{\alpha} K^{\beta} = m^{\alpha+\beta} p(L,K)$.

If $\alpha + \beta = 1$, then $p(mL,mK) = m\, p(L,K)$, which means that production is also increased by a factor m, as discussed earlier.

4. The Case of Multicollinearity

In the general linear regression model $y = X\beta + \varepsilon$ where y is $(n \times 1)$ vector of response variables, X is $(n \times p)$ matrix, $(n > p)$ of explanatory variables, β is $(p \times 1)$ vector of unknown parameters and ε is an $(n \times 1)$ vector of unobservable random errors, where $E(\varepsilon) = 0$, $var(\varepsilon) = \sigma^2 I_n$.

The problem of multicollinearity exists, when there exists a linear relationship or an approximate linear relationship among two or more explanatory variables.

Multicollinearity can be thought of as a situation where two or more explanatory variables in the data set move together, as a consequence it is impossible to use this data set to decide which of the explanatory variables is producing the observed change in the response variable.

Some multicollinearity is nearly always exist, but the important point is whether it is serious enough to cause appreciable damage to the regression analysis. The best way to deal with this problem may be to find a different data set, simplify the model by using variable selection techniques or using additional data to break the association between the related variables. Some indicators of multicollinearity include a low determinant of the information matrix (X'X), the smallest eigenvalue of the information matrix is very close to zero, a very high correlation among two or more explanatory variables, [5].

However, Farrar-Glauber test can be used to detect multicollinearity, where the null hypothesis to be tested is:
H_0: x_j's are orthogonal
against the alternative hypothesis
H_1: x_j's are not orthogonal, j = 1, 2,…,p.
The test statistic is:

$$c_0^2 = -[n-1-\frac{1}{6}(2p+5)]\ln|D| \quad \text{…(5)}$$

where, n is the sample size, p is the number of explanatory variables and D is the determinant of the correlation matrix of explanatory variables.

The calculated value of c_0^2 from equation (5) will be compared with the theoretical value obtained from the chi square table with p(p – 1)/2 degrees of freedom and specified level of significant. The null hypothesis H_0 will be rejected when the calculated value is more than the tabulated value, which means that the explanatory variables are not orthogonal and hence the multicollinearity problem is presented.

5. Restricted Least Squares Estimator:

The restricted least squares (RLS) method of estimation is used when one or more equality restrictions on the parameters of the model are available, [6].
Suppose the general linear model y = xb + e is subject to j equality restrictions represented by the matrix equation

$$Rb = r \quad \text{…(6)}$$

where R is (j´ p) matrix of restrictions, b is a p´ 1 vector of parameters and r is a j´ 1 vector of values of the restrictions.
In order to minimize the error sum of squares we have to minimize the Lagrangean function

$$e'e = (y - xb)'(y - xb) - 2l'(Rb - r) \quad \text{…(7)}$$

where l is a j´ 1 vector of Lagrange multipliers.
Equation (8) can be written as:

$$e'e = y'y - 2bx'y + b'x'xb - 2l'(Rb - r)$$

Differentiating partially with respect to b and l then set the derivatries equal to zero we obtain, [3]:

$$b_{RLS} = b_{OLS} + (x'x)^{-1}R'[R(x'x)^{-1}R']^{-1}(r - Rb_{OLS}) \quad \text{…(8)}$$

where $b_{OLS} = (x'x)^{-1}x'y$ is the ordinary least squares estimator.
The efficiency of the restricted least squares estimator is

$$eff\,(b_{RLS})_i = \frac{var(b_{RLS})_i}{var(b_{OLS})_i}, i = 0,1,2,...,p$$

In matrix terms we have, [3]:

$$eff(b_{RLS}) = I_p - R'[R(x'x)^{-1}R']^{-1}R(x'x)^{-1} \quad \text{…(9)}$$

The matrix in equation (9) is a p´ p square matrix; the elements on the main diagonal represent the relative efficiency of the parameters estimated by using the RLS method, these diagonal elements must be less than or equal to one, the off-diagonal elements are meaningless.

6. The Practical Application:

In this section we try to estimate, analyze and then interpret the Cobb-Douglas production function for the state company for leather industries in Iraq by employing the data obtained from the company for the period (1990-2010). In our study we propose an extension of the usual Cobb-Douglas production function to include the technological progress represented by the time T_t as well as the capital K_t and labor L_t as input variables. Thus our proposed function is:

$$P_t = a L_t^{b_1} K_t^{b_2} e^{CT_t} e^{U_t} \quad \ldots(10)$$

where C represent the rate of annual growth in production as a consequence to technological progress, moreover we suppose that the production function is CRTS that is $b_1 + b_2 = 1$. The linear from of this equation is:

$\ln P_t = \ln a + b_1 \ln L_t + b_2 \ln K_t + CT_t + U_t$

setting $y_t = \ln P_t$, $b_0 = \ln a$, $x_1 = \ln L_t$, $x_2 = \ln K_t$, $x_3 = T_t$, $b_3 = C$, $e = U_t$ we get:

$$y_t = b_0 + b_1 x_1 + b_2 x_2 + b_3 x_3 + e \quad \ldots(11)$$

The first four columns of table (1) below represent the values of production P_t, labor L_t, capital K_t and time T_t, the natural logarithms of these quantities are presented in the next three columns.

6.1 The OLS Estimators:

According to the principles of ordinary least squares estimation method [6]. The following results were obtained:

$$x'x = \begin{bmatrix} 21.00 & 169.23 & 409.81 & 231.00 \\ 169.23 & 1364.99 & 3306.52 & 1880.32 \\ 409.81 & 3306.52 & 8058.14 & 4695.67 \\ 231.00 & 1880.32 & 4695.67 & 3311.00 \end{bmatrix}$$

$$(x'x)^{-1} = \begin{bmatrix} 119.844 & -11.984 & -1.561 & 0.658 \\ -11.984 & 1.423 & 0.053 & -0.048 \\ -1.561 & 0.053 & 0.068 & -0.018 \\ 0.658 & -0.048 & -0.018 & 0.007 \end{bmatrix}$$

$$x'y = \begin{bmatrix} 461.93 \\ 3724.35 \\ 9055.24 \\ 5263.94 \end{bmatrix}$$

$$b_{OLS} = (x'x)^{-1} x'y = \begin{bmatrix} 57.8153 \\ -4.0659 \\ -0.4012 \\ 0.4341 \end{bmatrix}$$

The regression equation is:

$$\hat{y}_t = 57.8 - 4.07x_1 - 0.401x_2 + 0.434x_3$$

The analysis of variance calculations are summarized in table (2).

The hypothesis $H_0: b_j = 0,\ j = 0, 1, 2, 3$ is tested by comparing the calculated F from the ANOVA table with an appropriate percentage point of the $F(v_1,v_2)$ distribution where v_1, v_2 are degrees of freedom due to regression and error respectively.

If the calculated value is more than the theoretical value obtained from the F table then we reject H_0. At the basis of this test we have:

$F(3.17,0.05) = 3.59$. Since $19.10 > 3.59$, then we reject H_0. This means that on the basis of this test at least, we have no reason to doubt the adequacy of our model.

The variance-covariance matrix of ordinary least square estimators is given as, [3]:

$$\text{var-cov}(b_{OLS}) = s^2 (x'x)^{-1} = \begin{bmatrix} 116.859 & -11.686 & -1.522 & 0.642 \\ -11.686 & 1.387 & 0.052 & -0.047 \\ -1.522 & 0.052 & 0.066 & -0.017 \\ 0.642 & -0.047 & -0.017 & 0.007 \end{bmatrix}$$

The coefficient of determination R^2 is a convenient measure of the success of the regression equation in explaining the variation of the data.

$$R^2 = \frac{SS(\text{Regression})}{SS(\text{tatal})} = 0.77123$$

which means that 77.123 % of variations in the data can be explained by the regression equation.

6.2 The Farrar-Glauber Test:

In order to perform the farrar-Glauber test of multicollinearty we have to compute the determinant of the correlation matrix of explanatory variables which has the form:

$$|D| = \begin{vmatrix} 1 & r_{x_1x_2} & r_{x_1x_3} \\ r_{x_2x_1} & 1 & r_{x_2x_3} \\ r_{x_3x_1} & r_{x_3x_2} & 1 \end{vmatrix}$$

In our case of consideration.

$$D = \begin{bmatrix} 1 & 0.427 & 0.622 \\ 0.427 & 1 & 0.867 \\ 0.867 & 0.622 & 1 \end{bmatrix} \quad \text{and} \quad |D| = 0.147718.$$

Applying the formula in equation (5) we obtain:

$$c_0^2 = -(20 - \frac{1}{6}(11))\,(-1.91245) = 34.74289.$$

The theoretical value obtained from the Chi-square table with 3 degrees of freedom and 5% level of significant is equal to 7.816.
Since 34.74289 > 7.816 we conclude that the explanatory variables are not orthogonal and hence we face the multicollinearty problem.

6.3 The RLS Estimators:

In order to remove the effects of multicollinearty we propose using the restricted least squares method of estimation. In this case where the Cobb-Douglas production function assumed to be CRTS we have to write the restriction $b_1 + b_2 = 1$ in matrix equation as follows:

Let $R = [0 \;\; 1 \;\; 1 \;\; 0]$, $b = \begin{bmatrix} b_0 \\ b_1 \\ b_2 \\ b_3 \end{bmatrix}$, $r = [1]$.

Hence:

$b_1 + b_2 = 1$ implies that $\begin{bmatrix} 0 & 1 & 1 & 0 \end{bmatrix} \begin{bmatrix} b_0 \\ b_1 \\ b_2 \\ b_3 \end{bmatrix} = [1]$.

$$(x'x)^{-1}R' = \begin{bmatrix} -13.545 \\ 1.476 \\ 0.121 \\ -0.066 \end{bmatrix}.$$

$$R(x'x)^{-1}R' = 1.597$$

$$(R(x'x)^{-1}R')^{-1} = 0.626174$$

$$Rb_{OLS} = \begin{bmatrix} 0 & 1 & 1 & 0 \end{bmatrix} \begin{bmatrix} 57.8153 \\ -4.0659 \\ -0.4012 \\ 0.4341 \end{bmatrix} = [-4.4671]$$

$r - Rb_{OLS} = 1 + 4.4671 = 5.4671.$

$$(R(x'x)^{-1}R')^{-1}(r - Rb_{OLS}) = 3.4233$$

$$(x'x)^{-1}R'(R(x'x)^{-1}R')^{-1}(r - Rb_{OLS}) = \begin{bmatrix} -46.36935 \\ 5.05287 \\ 0.414226 \\ -0.22594 \end{bmatrix}.$$

$$B_{RLS} = b_{OLS} + (x'x)^{-1}R'(R(x'x)^{-1}R')^{-1}(r - Rb_{OLS})$$

$$b_{RLS} = \begin{bmatrix} 11.446595 \\ 0.9869731 \\ 0.013026 \\ 0.208158 \end{bmatrix}.$$

Accordingly, the regression equation is:

$$\hat{y}_t = 11.4465 + 0.9869x_1 + 0.0130x_2 + 0.2081x_3$$

and hence, the Cobb-Douglas production function is:

$$P_t = 93573.266\, L_t^{0.9869}\, K_t^{0.0130}\, e^{0.2081T_t}.$$

To determine the efficiency of the restricted least squares estimators we apply the formula in equation (9) and setting I_p to be a (4×4) identity matrix hence we obtain:

$$eff(b_{RLS}) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 8.4815 & 0.0752 & -0.0801 & 0.0413 \\ -8.4815 & -0.9248 & 0.9192 & 0.0413 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

7. Conclusions

Applying the RLS method to estimate the parameters of Cobb-Douglas production function leads to the following conclusions:

1. In contrast with the ordinary least squares estimators, the restricted least squares estimators of b_1 and b_2 are with positive signs. Moreover $b_1 + b_2 = 1$ the fact that agrees with the assumptions of Cobb-Douglass production function.
2. The efficiency matrix of b_{RLS} reveals that the restricted least squares estimators are more efficient than the ordinary least squares estimators that is because the second and third elements on the main diagonal which represents $eff(b_1)$ and $eff(b_2)$ respectively are less than one.
3. The values of estimated parameters means that a 100% increase in labor would lead to approximately a 98% increase in production, on the other hand, a 100% increase in capital would lead to approximately a 1.3% increase in production.
   Also the production satisfies an annual increment of 20.8% during the period of study as a result of technological progress.

References

1. Border, K.C., (2004), On the Cobb-Douglas Production Function, Internet Reference, California Institue of Tegnologe, Division of the Humanities and Social Sciences.
2. Lovell, C.A.K., (2007), Production Function, International Encyclopedia of the Social Sciences, 2nd Eddition.
3. كاظم، اموري هادي، مسلم، باسم شليبه، (2002)، القياس الاقتصادي المتقدم: النظرية والتطبيق، مطبعة الطيف، بغداد، العراق.
4. Hajkora, D. and Hurnik, J., (2007), Cobb-Douglas Production Function: The Case of Converging Economy, Journal of Economics and Finance, Vol. 57, No.9-10, pp.465-468.

5. Kubokawa, T. and Srivastava, M.S., (2003), Improved Empirical Bayes Ridge Regression Estimators Under Multicollinearity, Internet Reference, CIRJE-F-190.
6. Drappers, N.R. and Smith, H., (1981), Applied Regression Analysis, Second Edition, John Wiley and Sons, New York.
7. Graybill, F.A., (1976), Theory and Application of the Linear Model, Wadsworth Publishing Company, INC.USA.

Table (1)

P_t	L_t	K_t	T_t	$y_t = \ln P_t$	$x_1 = \ln L_t$	$x_2 = \ln K_t$
77187839	3597	50000000	1	18.16175	8.18786	17.72753
45565859	2601	50000000	2	17.63467	7.86365	17.72753
159555332	2636	50000000	3	18.88790	7.87702	17.72753
416864650	2897	50000000	4	19.84827	7.97143	17.72753
1307335567	2928	50000000	5	20.99126	7.98207	17.72753
3518767839	3182	50000000	6	21.98138	8.06527	17.72753
4973225270	2957	50000000	7	22.32733	7.99193	17.72753
6260702000	2689	50000000	8	22.55756	7.89692	17.72753
10104612000	2376	50000000	9	23.03626	7.77317	17.72753
6451468000	2460	50000000	10	22.58757	7.80792	17.72753
12147023000	2451	1450000000	11	23.22035	7.80425	21.09483
17559452000	2628	1543829187	12	23.58886	7.87398	21.15753
27728280000	2699	1520000000	13	24.04572	7.90064	21.14198
7731060000	2727	1520000000	14	22.76851	7.91096	21.14198
4449594000	2788	1520000000	15	22.21608	7.93308	21.14198
3451738000	3687	1520000000	16	21.96214	8.21257	21.14198
2939702000	4680	1520000000	17	21.80157	8.45105	21.14198
7093426000	4925	1520000000	18	22.68243	8.50208	21.14198
8423365000	4561	1520000000	19	22.85428	8.42530	21.14198
38318430000	4434	1520000000	20	24.36920	8.39706	21.14198
39984891000	4471	1520000000	21	24.41177	8.40537	21.14198

Table (2) ANOVA

Source	DF	SS	MS	F
Regression	3	55.884	18.628	19.10

Error	17	16.577	0.975	
Total	20	72.460		

إستعمال طريقة المربعات الصغرى المقيدة لتقدير وتحليل معلمات دالة الانتاج لـ -Cobb Douglas مع تطبيق عملي

حازم منصور كوركيس ، أحمد عيسى عبد النبي
قسم الرياضيات - كلية التربية ابن الهيثم - جامعة بغداد

استلم البحث في : 8 أيلول 2011 قبل البحث في : 18 تشرين الأول 2011

الخلاصة

في هذا البحث تم استعملت طريقة المربعات الصغرى المقيدة لتقدير معلمات دالة الانتاج لـ (Cobb-Douglas) ومن ثم تحليل وتفسير النتائج التي تم التوصل إليها مع تطبيق عملي لتقدير دالة الانتاج في الشركة العامة للصناعات الجلدية في العراق للمدة (1990-2010) وقد قمنا بالاستعانة بالبرنامج الاحصائي الجاهز (SPSS) لاجراء الحسابات المطلوبة

الكلمات المفتاحية: الأنموذج الخطي العام ، دالة الانتاج لـ Cobb-Douglas ، مقدرات المربعات الصغرى الاعتيادية ، مقدرات المربعات الصغرى المقيدة، التعدد الخطي .