**Case 1: Validation of Keynes’ Consumption Hypothesis.**

Keynes Stated:

The fundamental psychological law… is that men [women] are disposed, as a rule and on average, to increase their consumption as their income increases, but not as much as the increase in their income.

Validate this hypothesis of Keynes through the following steps of econometrics:

1. Statement of the theory or hypothesis.

2. Specification of the mathematical model of the theory.

3. Specification of the statistical or econometric model.

4. Obtaining the data.

5. Estimation of the parameters of the econometric model.

6. Hypothesis testing.

7. Forecasting or prediction.

8. Using the model for policy purposes.

While discussing the steps ensure to elaborate on autonomous consumption, marginal propensity to consume (MPC), source of the data in step 4, goodness of the model (𝑅2, p-values etc.) some example predictions, income multiplier and policy examples.

**Case Study – Econometric Modelling of Keynesian Consumption Function using Simple Linear Regression**

**1.1. Statement of Theory -**Keynes stated (The General Theory of Employment, Interest and Money): The fundamental psychological law … is that men [women] are disposed, as a rule and on average, to increase their consumption as their income increases, but not as much as the increase in their income.

In short, Keynes postulated that the marginal propensity to consume (MPC), the rate of change of consumption for a unit (say, a rupee) change in income, is greater than zero but less than one.

**1.2. Specification of the Mathematical Model of Consumption**

Although Keynes postulated a positive relationship between consumption and income, he did not specify the precise form of the functional relationship between the two. For simplicity, Keynesian Consumption Function is taken as follows.

𝐶 = 𝛼 + 𝛽Y; 0 ≤ 𝛽 ≤ 1… (2.1)

Where 𝐶 = consumption expenditure, 𝑌 = income. The intercept 𝛼 measures the autonomous consumption and the slope 𝛽 measures the marginal propensity to consume (MPC).

**1.3. Specification of the Statistical or Econometric Model of Consumption**

Equation (2.1) specifies a deterministic or non-stochastic relationship between income and consumption. In practice the relationship is not exact instead it is stochastic or random. To allow for the inexact relationship one can define a stochastic model by adding a random error component.

𝐶 = 𝛼 + 𝛽𝑌 + 𝜖; 0 ≤ 𝛽 ≤ 1… (2.2)

**1.4. Data Collection**

The equation (2.2) represents a simple linear regression model. To analyse this model we need to estimate the unknown parameters using a random sample. Hence we need data on consumption and income. One can take the private final consumption expenditure (PFCE) as the consumption variable and gross domestic product (GDP) as income variable. Consider the data file Module 1.1 – SLRM.xlsx for analysis which has PFCE and GDP data for 1950-51 to 2006-2007, both in 1999-2000 prices measured in crore rupees.

Source: National Accounts Statistics (2000, 2007, 2009), Central Statistical Organization, Ministry of Statistics and Programme Implementation, Government of India (http://www.mospi.gov.in/)

**1.5. Estimation of the Econometric Model**

We analyse the model in equation (2.2) as a classical simple linear regression model and obtain a sample regression function (SRF) of the Keynesian Consumption Function as follows. Also check goodness of the model using 𝑅2 and 𝑅𝐴𝑑𝑗 2.

| Model Summaryb | | | | |
| --- | --- | --- | --- | --- |
| Model | R | R Square | Adjusted R Square | Std. Error of the Estimate |
| 1 | .998a | .995 | .995 | 30,184.533 |
| a. Predictors: (Constant), Y b. Dependent Variable: C | | | | |
|  | | | | |

* The Model Summary Table reports the strength of the relationship between the model and the dependent variable. The multiple Correlation Coefficient (R) is the linear correlation between the observed and model predicted values of the dependent variable. A large value (R=0.998) indicates a strong relationship.
* R2 is 0.995 meaning that 99.5% variance in consumption is explained by the model. The standard error of estimate is a measure of accuracy of the prediction. In the case of a linear regression with more than one variable we prefer adjusted R-squared. **For a single independent variable model, both statistics are interchangeable.**

The ANOVA table tests the acceptability of the model of the model from the statistical perspective. We set the following hypothesis :( F test)

**Ho**: The dependent variable is not significantly explained by the independent variables.

**H1:** The dependent variable is significantly explained by the independent variables.

| ANOVAb | | | | | | |
| --- | --- | --- | --- | --- | --- | --- |
| Model | | Sum of Squares | df | Mean Square | F | Sig. |
| 1 | Regression | 1.005E13 | 1 | 1.005E13 | 11029.640 | .000a |
| Residual | 5.011E10 | 55 | 9.111E8 |  |  |
| Total | 1.010E13 | 56 |  |  |  |
| a. Predictors: (Constant), Y b. Dependent Variable: C | | | | | | |

* ANOVA Table tests the significance of the regression model. We can see from the table that Sig (p value) =0.00.As p<0.05 our predictors are significantly better than would be expected by chance. The regression line predicted by the independent variables explains a significant amount of variance in the dependent variable.

**1.6. Statistical Significance of the Econometric Model**

Test for significance of MPC using an appropriate t-test. Set up the null and alternative hypotheses as follows.

Null Hypothesis (𝐻0) ∶ 𝑀𝑃𝐶 = 0

Alternative Hypothesis (𝐻1): 𝑀𝑃𝐶 ≠ 0

Calculate the p-value and note for the observed sample data 𝑝 < (= 0.05). Validate this from your own results. Hence MPC is significant.

| Coefficientsa | | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Model | | Unstandardized Coefficients | | Standardized Coefficients | |  | |
| B | Std. Error | Beta | | t | Sig. |
| 1 | (Constant) | 103736.049 | 6587.369 |  | | 15.748 | .000 |
| Y | **.630** | .006 | .998 | | 105.022 | .000 |
| * The fitted regression model is: 𝐶 ̂ = 𝛼 ̂ + 𝛽 ̂𝑌 = 103736.049 + 0.630𝑌, where ̂ is the value of the autonomous consumption that is the intercept of the line indicates the average level of consumption expenditure when the income is 0. We have tested for significance of MPC by using an appropriate t-test. As p<0.05 we conclude that MPC is significant. * The unstandardized coefficients B column, gives us the coefficients of the independent variables in the regression equation. Each of these β values has an associated standard error indicating to what extent these values would vary across different samples and these standard errors are used to determine whether or not the β values significantly differ from 0.For example:β =0.006 indicates that as Income increases by 1 standard deviation consumption increases by 0.006 standard deviations. * The standardised beta coefficient column shows the contribution that an individual variable makes to the model. The beta weight is the average amount the dependent variable increases when the independent variable increases by one standard deviation (all other independent variables are held constant). | | | | | | | |
|  | | | | | |
|  | | | | | |

For the observed sample data it can be seen that 0 ≤ 𝛽 ̂ ≤ 1. Validate this from your own results.

**1.7. Forecasting or Prediction**

The estimated PRF or sample regression function can be used to forecast of predict the consumption expenditure at future income values. The GDP value for 2007-2008 was 31, 29,717 crore, we can forecast the consumption expenditure for 2007-2008 using the SRF as follows.

**𝐶 ̂2007−2008 = 103736.0493 + 0.6303 × (31, 29,717) = 2076396.6744**

**1.8. Use of Model for Control or Policy Purposes**

The sample regression function 𝐶 ̂ = 𝛼 ̂ + 𝛽 ̂𝑌 = 103736.0493 + 0.6303 𝑌 can be used for policy formulation. Consider the following examples.

1. Suppose the government believes that consumer expenditure of about 25, 00, 000 crore will help increase employment rate in the country. What level of income will guarantee that target consumption to attain specific employment objective? Assuming that the sample regression function is adequate enough we can use it to answer this question by solving the following equation for 𝑌.

25, 00,000 = 103736.0493 + 0.6303 𝑌

Which gives 𝑌 = 38, 01, 783 crore, approximately. That is, an income level of about 38, 01, 783 crore, given an MPC of 0.63, will produce a consumption expenditure of about 25, 00, 000 crore.

By appropriate fiscal and monetary policy mix, the government can manipulate the income level to achieve the desired level of target consumption to attain specific employment objective.

2. Suppose the government decides to propose a reduction in income tax. What will be the effect of such a policy on income and thereby on consumption expenditure and ultimately on employment?

Suppose that as a result of the proposed policy change, investment expenditure increases. What will be the effect on economy? As macroeconomic theory shows, the change in income following, say, a rupee’s worth of change in investment expenditure is given by income multiplier M, which is defined as

𝑀 =1 / (1 – 𝑀𝑃𝐶)

**If we use the estimated MPC as 0.63 obtained from the data, income multiplier is estimated at about 2.70. That is, an increase (decrease) of a rupee in investment will eventually lead to more than a twofold increase (decrease) in income.**

Ultimately sample regression function can be used to obtain the effect on consumption due to a change in income.

**Case 2: Modelling Production Function**.

Consider the data Cobb-Douglas Production Function.xlsx and model the Cobb-Douglas production function for Taiwan for the years 1958-1972. Comment on state of production process in terms of the following:

1. Total Factor Productivity

2. Labour Elasticity of the Output

3. Capital Elasticity of the Output

4. Returns to Scale

Discuss all the steps involved in an elaborated manner. Whenever a regression model is built goodness of the model (𝑅2, p-values etc.) should be provided.

Solution

**Case Study – Econometric Modelling of Cobb-Douglas Production Function using Multiple Linear Regression**

The Cobb–Douglas production function is a particular functional form of the production function, widely used to represent the technological relationship between the amounts of two or more inputs, particularly physical capital and labour, and the amount of output that can be produced by those inputs.

In its most standard form for production of a single good with two factors, the function is

**𝑌 = 𝛼𝐿𝛽𝐾𝛾**

**Where Y = total production (the real value of all goods produced in a year), L = labour input (the total number of person-hours worked in a year), K = capital input (the real value of all machinery, equipment, and buildings), 𝛼 = total factor productivity, 𝛽 and 𝛾 are the output elasticities of labour and capital, respectively.** These values are constants determined by available technology.

Output elasticity measures the responsiveness of output to a change in levels of either labour or capital used in production, ceteris paribus. For example if 𝛽 = 0.45, a 1% increase in capital usage would lead to approximately a 0.45% increase in output.

* If **𝛽 + 𝛾 = 1**, the production function has constant returns to scale, meaning that doubling the usage of capital K and labour L will also double output Y.
* If **𝛽 + 𝛾 > 1,** the production function has increasing returns to scale.
* If **𝛽 + 𝛾 < 1**, the production function has decreasing returns to scale.

We can estimate the parameters involved in the model using the multiple linear regression model. Cobb-Douglas production function, in its stochastic form, may be expresses as

**𝑌 = 𝛼𝐿𝛽𝐾𝛾𝑒𝜖**

Where 𝜖 is the stochastic disturbance term.

In fact the Cobb-Douglas production function relates the production and, labour and capital through a nonlinear function of unknown parameters. But we can analyse a transformed linear model after taking a log transformation of the original model.

ln𝑌 = ln𝛼 + 𝛽 ln𝐿 + 𝛾ln𝐾 + 𝜖

𝑌 ̃ = 𝛼 ̃ + 𝛽𝐿 ̃ + 𝛾𝐾 ̃ + 𝜖 … (3.1)

Analysing the transformed model is not completely equivalent to analyse the original model. In fact the exact analysis is to be performed using non-linear regression but the log transformation gives fairly good approximation. Exactly the same steps explained using SLRM can be executed to study the log-transformed Cobb-Douglas production function.

| Model Summaryb | | | | |
| --- | --- | --- | --- | --- |
| Model | R | R Square | Adjusted R Square | Std. Error of the Estimate |
| 1 | .943a | .889 | .871 | .074800281 |
| a. Predictors: (Constant), lnK, lnL b. Dependent Variable: lnY | | | | |
| * The Model Summary Table reports the strength of the relationship between the model and the dependent variable. The multiple Correlation Coefficient (R) is the linear correlation between the observed and model predicted values of the dependent variable. A large value (R=0.943) indicates a strong relationship. * R-squared measures the proportion of the variation in your dependent variable (Y) explained by your independent variables (X) for a linear regression model. Adjusted R-squared adjusts the statistic based on the number of independent variables in the model. * R2 is 0.889 meaning that 88.9% variance in the log of output is explained by the logs of labour and capital. The standard error of estimate is a measure of accuracy of the prediction. In the case of a linear regression with more than one variable we prefer adjusted R-squared. Adjusted r2 compares the explanatory power of regression model that contains different number of predictors. | | | | |

The ANOVA table tests the acceptability of the model of the model from the statistical perspective. We set the following hypothesis :( F test)

**Ho**: The dependent variable is not significantly explained by the independent variables.

**H1:** The dependent variable is significantly explained by the independent variables

| ANOVAb | | | | | | |
| --- | --- | --- | --- | --- | --- | --- |
| Model | | Sum of Squares | df | Mean Square | F | Sig. |
| 1 | Regression | .538 | 2 | .269 | 48.083 | .000a |
| Residual | .067 | 12 | .006 |  |  |
| Total | .605 | 14 |  |  |  |
| a. Predictors: (Constant), lnK, lnL b. Dependent Variable: lnY | | | | | | |
| * ANOVA Table tests the significance of the regression model. We can see from the table that Sig (p value) =0.00.As p<0.05 our predictors are significantly better than would be expected by chance. The regression line predicted by the independent variables explains a significant amount of variance in the dependent variable. | | | | | | |

| Coefficientsa | | | | | | |
| --- | --- | --- | --- | --- | --- | --- |
| Model | | Unstandardized Coefficients | | Standardized Coefficients |  | |
| B | Std. Error | Beta | t | Sig. |
| 1 | (Constant) | -3.339 | 2.449 |  | -1.363 | .198 |
| lnL | 1.499 | .540 | .373 | 2.777 | .017 |
| lnK | .490 | .102 | .644 | 4.801 | .000 |
| a. Dependent Variable: lnY | | | | | | |

The fitted regression model is: ln𝑌 = ln𝛼 + 𝛽 ln𝐿 + 𝛾ln𝐾 + 𝜖 ------- (1)

* 𝑌 ̃ = 𝛼 ̃ + 𝛽𝐿 ̃ + 𝛾𝐾 ̃ + 𝜖 𝑌 ̃ = -3.339 +1.499𝐿 ̃+.490𝐾 ̃ From (1) we have calculated **α=0.035472,**

is **total factor productivity** that is the intercept of the line indicating the total production when the labour and capital are 0. We have tested for significance of MPC by using an appropriate t-test. As p<0.05 we conclude that MPC is significant.

* The unstandardized coefficients B column, gives us the coefficients of the independent variables in the regression equation. Each of these β values has an associated standard error indicating to what extent these values would vary across different samples and these standard errors are used to determine whether or not the β values significantly differ from 0.For example:β =0.102 indicates that as Capital Input increases by 1 standard deviation consumption increases by 0.102 standard deviations.
* β =0.540 indicates that as Labour Input increases by 1 standard deviation consumption increases by 0.540 standard deviations.
* The standardised beta coefficient column shows the contribution that an individual variable makes to the model. The beta weight is the average amount the dependent variable increases when the independent variable increases by one standard deviation (all other independent variables are held constant).
* **We see that in Taiwanese agricultural sector for the period 1958-1972 the output labour elasticities for labour and capital were 1.499 and 0.490 respectively. In other words over the period of study holding the capital input constant, a 1% increase in the labour input led on the average to about a 1.5% increase in the output.**
* **Similarly holding the labour input constant, a 1% increase in the capital input led on the average to about a 0.5% increase in the output.**
* **Adding the 2 output elasticities, we obtain 1.989 which gives the value of the returns to scale parameter. So the Taiwanese agricultural sector is characterised by increasing returns to scale.**
* **Our results show that a big part of Taiwan s economic growth was based on Labour although capital also has an important role to play. So our estimates indicate that given relative stable terms of trade its impact on explaining GDP growth of the Taiwanese economy is significant.**

**Case 3: Estimation of Price Elasticity of Demand.**

Consider the data and estimate the price elasticity of demand for some commodity. Assuming the demand function to be as follows:

𝑄𝑑 = exp (𝛽0 + 𝛽1 ln𝑃)

Where 𝑄𝑑 is the quantity demanded of the given commodity and 𝑃 is the price of the given commodity.

**Price Elasticity**:

Law of Demand is described as the inverse relationship between the quantity demanded of a good and its price, provided all other factors are constant.

The quantification of the degree of inverse relationship is studied by the concept of elasticity which is a measure of responsiveness of the quantity demanded with respect to price change. Mathematically, it is defined as the ratio of proportionate change in quantity demanded and proportionate change in price. Given two prices and corresponding quantities demanded we can calculate elasticity as follows:

**𝑒𝑑 = (Δ𝑄𝑑/ 𝑄) / (Δ𝑃/ 𝑃) = (Δ𝑄𝑑/ Δ𝑃) × (𝑃/ 𝑄**)

For a given price and quantity demanded elasticity can be viewed as follows:

𝑒𝑑 = (d𝑄𝑑 /d𝑃) × (𝑃/ 𝑄) so depending upon the quantity demanded function under consideration 𝑒𝑑 will behave accordingly. Moreover, elasticity changes with the change in value of 𝑃 even for the same demand function

| Model Summaryb | | | | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Model | | R | R Square | | Adjusted R Square | | | Std. Error of the Estimate | |
| 1 | | 1.000a | 1.000 | | 1.000 | | | .00000 | |
| a. Predictors: (Constant), lnP b. Dependent Variable: lnQ | | | | | | | | | |
|  | | | | | | | | | |
| * The Model Summary Table reports the strength of the relationship between the model and the dependent variable. The multiple Correlation Coefficient (R) is the linear correlation between the observed and model predicted values of the dependent variable. A large value (R=1) indicates a very strong relationship. It is a perfect model. * R-squared measures the proportion of the variation in your dependent variable (Y) explained by your independent variables (X) for a linear regression model. * R2 is 1.0 meaning that 100% variance in the log of output is explained by the log of quantity. The standard error of estimate is a measure of accuracy of the prediction. | | | | | | | | | |
| Coefficientsa | | | | | | | | | | | | | |
| Model | | | | | Unstandardized Coefficients | | | Standardized Coefficients | | t | | Sig. |  | |
| B | | Std. Error | Beta | |
| 1 | | (Constant) | | | 5.000 | | .000 |  | | . | | . |
| lnP | | | -.530 | | .000 | -1.000 | | . | | . |
| a. Dependent Variable: lnQ | | | | | | | | | | | | | |

* As we have a perfect model where standard error is 0, so t values don’t exist.-0.530 is price elasticity of demand. The quantification of the degree of inverse relationship is studied by the concept of elasticity which is a measure of responsiveness of the quantity demanded with respect to price change.
* When PED is less than one**, demand is inelastic**. This can be interpreted as consumers being insensitive to changes in price: a 1% increase in price will lead to a drop in quantity demanded of less than 1%.

**Note**: When PED is greater than one, demand is elastic. This can be interpreted as consumers being very sensitive to changes in price: a 1% increase in price will lead to a drop in quantity demanded of more than 1%.

Source: Boundless. “Interpretations of Price Elasticity of Demand.” *Boundless Economics*. Boundless, 21 Jul. 2015. Retrieved 22 Apr. 2016 from https://www.boundless.com/economics/textbooks/boundless-economics-textbook/elasticity-and-its-implications-6/price-elasticity-of-demand-54/interpretations-of-price-elasticity-of-demand-210-12301/

**Case 4: Indirect Least Squares Estimation:**

Consider the following price mechanism model for some commodity:

Demand Function: 𝑄𝑡𝑑 = 𝛼0 + 𝛼1𝑃𝑡 + 𝛼2𝑌 𝑡 + 𝑢1𝑡; 𝛼1 < 0 and 𝛼2 > 0

Supply Function: 𝑄𝑡 𝑠 = 𝛽0 + 𝛽1𝑃𝑡 + 𝑢2𝑡; 𝛽1 > 0

Equilibrium Function: 𝑄𝑡 𝑠 = 𝑄𝑡 𝑑 = 𝑄𝑡

Where 𝑄𝑡𝑑 and 𝑄𝑡 𝑠 are the quantity demanded and quantity supplied respectively, of the given commodity; 𝑃𝑡 is price of the given commodity and 𝑌 𝑡 is the income.

Perform the following tasks:

1. Is the system of equations identified?

2. Obtain the reduced form of the given system.

3. Estimate the parameters using Indirect Least Squares method.

**Solution:** The reduced form is

𝑃𝑡 = 𝜋0 + 𝜋1𝑌 𝑡 + 𝑣1𝑡

𝑄𝑡 = 𝜋2 + 𝜋3𝑌 𝑡 + 𝑣2𝑡

Where 𝜋0 = (𝛽0 − 𝛼0)/ (𝛼1 − 𝛽1), 𝜋1 = −𝛼2/ (𝛼1 − 𝛽1), 𝜋2 =𝛼1𝛽0 − 𝛼0𝛽1 / (𝛼1 − 𝛽1) and 𝜋3 = −𝛼2𝛽1/ (𝛼1 − 𝛽1)

**Step 1**. Obtain ordinary least squares estimators of 𝜋0, 𝜋1, 𝜋2 and 𝜋3 using the reduced form equations as follows. From first reduced form equation we have,

𝜋 ̂1 =∑ 𝑃𝑡𝑌 𝑡 / ∑ 𝑌 𝑡 2 𝑡=1… 𝑛 And 𝜋 ̂0 = 𝑃 ̅ − 𝜋 ̂1𝑌 ̅

Where 𝑃 ̅ = (1/ 𝑛) ∑ 𝑃𝑡 𝑡=1…. and ̅ = (1/ 𝑛)∑ 𝑌 𝑡 𝑡=1.…n

From second reduced form equation we have,

𝜋 ̂3 =∑ 𝑄𝑡𝑌 𝑡 / ∑ 𝑌 𝑡 2 𝑡=1….. 𝑛 And 𝜋 ̂2 = 𝑄 ̅ − 𝜋 ̂3𝑌 ̅

Where 𝑄 ̅ = (1/ 𝑛) ∑ 𝑄𝑡 𝑡=1…n and 𝑌 ̅ = (1 /𝑛) ∑ 𝑌 𝑡 𝑡=1,…, n

**Step 2.** As supply equation is exactly identified. Indirect least squares estimators of the structural coefficients in it viz., 𝛽1 and 𝛽0 are obtained as follows.

⇒ 𝛽 ̂1,𝑆 =𝜋 ̂3/ 𝜋 ̂1 and 𝛽 ̂0,𝐼𝐿𝑆 = 𝜋 ̂2 –(𝜋 ̂3 /𝜋 ̂1)𝜋 ̂0

| Model | Variables Entered | | Variables Removed | | Method | |
| --- | --- | --- | --- | --- | --- | --- |
| 1 | Y (Income)a | | . | | Enter | |
| b. Dependent Variable: P (Price) | | | | | | |
| First of all we make a model where the dependent variable is price and independent variable is income. | | | | | | |
| **Model Summaryb** | | | | | | | |
| Model | R | R Square | | Adjusted R Square | | Std. Error of the Estimate | |
| 1 | .494a | .244 | | .217 | | 9.107 | |

* The Model Summary Table reports the strength of the relationship between the model and the dependent variable. The multiple Correlation Coefficient (R) is the linear correlation between the observed and model predicted values of the dependent variable. A moderate value (R=0.494) indicates a weak relationship.
* R-squared measures the proportion of the variation in your dependent variable (Y) explained by your independent variables (X) for a linear regression model. Adjusted R-squared adjusts the statistic based on the number of independent variables in the model.

| * The fitted regression model is: P ̂ = 𝛼 ̂ + 𝛽 ̂𝑌 = 90.961 + 0.001𝑌, where ̂ is the intercept of the line which indicates the average price when the income is 0. As p<0.05 we conclude that both the coefficients are significant. * The unstandardized coefficients B column, gives us the coefficients of the independent variables in the regression equation. Each of these β values has an associated standard error indicating to what extent these values would vary across different samples and these standard errors are used to determine whether or not the β values significantly differ from 0. * The standardised beta coefficient column shows the contribution that an individual variable makes to the model. The beta weight is the average amount the dependent variable increases when the independent variable increases by one standard deviation (all other independent variables are held constant). * The confidence intervals are related to the p values such that the coefficient will not be statistically significant if the confidence interval includes 0.These confidence intervals can help us to put the estimate from the coefficient into perspective by seeing how much the value can vary.   ANOVAb | | | | | | |
| --- | --- | --- | --- | --- | --- | --- |
| Model | | Sum of Squares | df | Mean Square | F | Sig. |
| 1 | Regression | 749.770 | 1 | 749.770 | 9.041 | .006a |
| Residual | 2322.096 | 28 | 82.932 |  |  |
| Total | 3071.867 | 29 |  |  |  |
| a. Predictors: (Constant), Y (Income) | | | | | | |
| b. Dependent Variable: P (Price) | | | | | | |

R2 is 0.224 meaning that 22.4% variance in the output is explained by income. In the case of a linear regression with more than one variable we prefer adjusted R-squared. Adjusted r2 compares the explanatory power of regression model that contains different number of predictors

| Coefficientsa | | | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Model | | Unstandardized Coefficients | | Standardized Coefficients | t | Sig. | 95.0% Confidence Interval for B | |
| B | Std. Error | Beta | Lower Bound | Upper Bound |
| 1 | (Constant) | 90.961 | 4.051 |  | 22.456 | .000 | 82.664 | 99.258 |
| Y (Income) | .001 | .000 | .494 | 3.007 | .006 | .000 | .001 |
| a. Dependent Variable: P (Price) | | | | | | | | |

The ANOVA table tests the acceptability of the model of the model from the statistical perspective. We set the following hypothesis :( F test)

**Ho**: The dependent variable is not significantly explained by the independent variables.

**H1:** The dependent variable is significantly explained by the independent variables

* ANOVA Table tests the significance of the regression model. We can see from the table that Sig (p value) =0.006.As p<0.05 our predictors are significantly better than would be expected by chance. The regression line predicted by the independent variables explains a significant amount of variance in the dependent variable.

| **Model Summaryb** | | | | |
| --- | --- | --- | --- | --- |
| Model | R | R Square | Adjusted R Square | Std. Error of the Estimate |
| 1 | .969a | .940 | .938 | 3.509 |
| a. Predictors: (Constant), Y (Income) b. Dependent Variable: Q (Quantity) | | | | |
| * The Model Summary Table reports the strength of the relationship between the model and the dependent variable. The multiple Correlation Coefficient (R) is the linear correlation between the observed and model predicted values of the dependent variable. A moderate value (R=0.969) indicates a strong relationship. * R-squared measures the proportion of the variation in your dependent variable (Y) explained by your independent variables (X) for a linear regression model. Adjusted R-squared adjusts the statistic based on the number of independent variables in the model * R2 is 0.94 meaning that 94% variance in the output is explained by income. Adjusted r2 compares the explanatory power of regression model that contains different number of predictors. | | | | |

| ANOVAb | | | | | | |
| --- | --- | --- | --- | --- | --- | --- |
| Model | | Sum of Squares | df | Mean Square | F | Sig. |
| 1 | Regression | 5384.760 | 1 | 5384.760 | 437.395 | .000a |
| Residual | 344.707 | 28 | 12.311 |  |  |
| Total | 5729.467 | 29 |  |  |  |
| a. Predictors: (Constant), Y (Income) b. Dependent Variable: Q (Quantity) | | | | | | |

* ANOVA Table tests the significance of the regression model. We can see from the table that Sig (p value) =0.00.As p<0.05 our predictors are significantly better than would be expected by chance. The regression line predicted by the independent variables explains a significant amount of variance in the dependent variable.

| Coefficientsa | | | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Model | | Unstandardized Coefficients | | Standardized Coefficients | t | Sig. | 95.0% Confidence Interval for B | |
| B | Std. Error | Beta | Lower Bound | Upper Bound |
| 1 | (Constant) | 59.771 | 1.561 |  | 38.299 | .000 | 56.574 | 62.967 |
| Y (Income) | .002 | .000 | .969 | 20.914 | .000 | .002 | .002 |
| a. Dependent Variable: Q (Quantity) | | | | | | | | |

* The fitted regression model is: Q ̂ = 𝛼 ̂ + 𝛽 ̂𝑌 = 59.771 + 0.002𝑌, where α is the intercept of the line which indicates the average quantity demanded when the income is 0. As p<0.05 we conclude that both the coefficients are significant.
* The unstandardized coefficients B column, gives us the coefficients of the independent variables in the regression equation. Each of these β values has an associated standard error indicating to what extent these values would vary across different samples and these standard errors are used to determine whether or not the β values significantly differ from 0.
* The standardised beta coefficient column shows the contribution that an individual variable makes to the model. The beta weight is the average amount the dependent variable increases when the independent variable increases by one standard deviation (all other independent variables are held constant).
* The confidence intervals are related to the p values such that the coefficient will not be statistically significant if the confidence interval includes 0.These confidence intervals can help us to put the estimate from the coefficient into perspective by seeing how much the value can vary.

**The reduced form of the given system is:**

𝑃𝑡 = 𝜋0 + 𝜋1𝑌 𝑡 + 𝑣1𝑡 𝑃𝑡 =90.961 + 0.001𝑌

𝑄𝑡 = 𝜋2 + 𝜋3𝑌 𝑡 + 𝑣2𝑡 𝑄𝑡 =59.771 + 0.002𝑌

**As the true demand equation is under identified but the supply equation is exactly identified. So the system of equations is under-identified.**

**Parameters estimated using** Indirect Least Squares method is:

𝛽 ̂1,𝑆 =𝜋 ̂3/ 𝜋 ̂1 and 𝛽 ̂0,𝐼𝐿𝑆 = 𝜋 ̂2 –(𝜋 ̂3 /𝜋 ̂1)𝜋 ̂0 𝛽 ̂1,𝐼𝐿𝑆 =2 , 𝛽 ̂0,𝐼𝐿𝑆= -122.151

**Case 5: Two Stage Least Square Estimation (2-SLS):**

Consider a form of quantity-theory-Keynesian approaches to income determination states that income is determined by money supply, investment expenditure and government expenditure.

Income Function: 𝑌 𝑡 = 𝛽10 + 𝛽11𝑀𝑡 + 𝛾11𝐼𝑡 + 𝛾12𝐺𝑡 + 𝑢1𝑡

Money Supply Function: 𝑀𝑡 = 𝛽20 + 𝛽21𝑌 𝑡 + 𝑢2𝑡

Where 𝑌 𝑡 is the income at time 𝑡, 𝑀𝑡 is the money supply at time 𝑡, 𝐼𝑡 is the investment expenditure at time 𝑡 and 𝐺𝑡 is the government expenditure at time 𝑡.

Perform the following tasks:

1. Is the system of equations identified?

2. Obtain the reduced form of the given system.

3. Estimate the parameters using Two Stage Least Squares method.

**Solution:**

We can estimate the structural coefficients in the money supply function using the 2SLS as follows.

Step 1. As income 𝑌 𝑡 is the only explanatory endogenous variable in the supply function, built a liner regression model of 𝑌 𝑡 on all exogenous variables present in the whole model viz. 𝐼𝑡, the investment expenditure and 𝐺𝑡, the government expenditure

𝑌 𝑡 = 𝛼0 + 𝛼1𝐼𝑡 + 𝛼2𝐺𝑡 + 𝜖𝑡

Using OLS we obtain 𝑌 ̂𝑡 = 𝛼 ̂0 + 𝛼 ̂1𝐼𝑡 + 𝛼 ̂2𝐺𝑡. 𝛼 ̂0 and 𝛼 ̂1 are consistent for 𝛼0 and 𝛼1 respectively, i.e. plim (𝛼 ̂0) = 𝛼0 and plim (𝛼 ̂1) = 𝛼1.

Step 2. Replace 𝑌 𝑡 by 𝑌 ̂𝑡 in the money supply equation and build the following linear regression model.

𝑀𝑡 = 𝛽20 + 𝛽21𝑌 ̂𝑡 + 𝑢2𝑡

Using OLS estimate 𝛽20 and 𝛽21.

Now let us consider a numerical example.

**Note: 2SLS can be applied to an individual equation in the system without directly taking into account any other equation (s) in the system, i.e. estimation of structural parameters of one structural equation will not involve direct information from the other structural equations. Hence 2SLS is a Limited Information Estimation Method. 2SLS offers an economical method for solving econometric models involving a large number of equations, and is hence used extensively in practice.**

| Variables Entered/Removed | | | |
| --- | --- | --- | --- |
| Model | Variables Entered | Variables Removed | Method |
| 1 | G (Government Exp), I (Investment)a | . | Enter |
| a. All requested variables entered. | | | |

We build a multiple linear regression model with government expenditure and Investment as the explanatory variables and Income as the dependent variable.

| Model Summaryb | | | | |
| --- | --- | --- | --- | --- |
| Model | R | R Square | Adjusted R Square | Std. Error of the Estimate |
| 1 | .996a | .992 | .991 | 124.148 |
| a. Predictors: (Constant), G (Government Exp), I (Investment) b. Dependent Variable: Y (Income) | | | | |

* The Model Summary Table reports the strength of the relationship between the model and the dependent variable. The multiple Correlation Coefficient (R) is the linear correlation between the observed and model predicted values of the dependent variable. A moderate value (R=0.996) indicates a strong relationship.
* R-squared measures the proportion of the variation in your dependent variable (Y) explained by your independent variables (X) for a linear regression model. Adjusted R-squared adjusts the statistic based on the number of independent variables in the model.

R2 is 0.992 meaning that 99.2% variance in the output is explained by income. In the case of a linear regression with more than one variable we prefer adjusted R-squared. Adjusted R2 compares the explanatory power of regression model that contains different number of predictors.

| ANOVAb | | | | | | |
| --- | --- | --- | --- | --- | --- | --- |
| Model | | Sum of Squares | df | Mean Square | F | Sig. |
| 1 | Regression | 2.613E7 | 2 | 1.306E7 | 847.610 | .000a |
| Residual | 200365.303 | 13 | 15412.716 |  |  |
| Total | 2.633E7 | 15 |  |  |  |
| a. Predictors: (Constant), G (Government Exp), I (Investment) | | | | | | |
| b. Dependent Variable: Y (Income) | | | | | | |

| Coefficientsa | | | | | | |
| --- | --- | --- | --- | --- | --- | --- |
| Model | | Unstandardized Coefficients | | Standardized Coefficients |  | |
| B | Std. Error | Beta | t | Sig. |
| 1 | (Constant) | 3091.934 | 157.835 |  | 19.590 | .000 |
| I (Investment) | 1.670 | .190 | .475 | 8.786 | .000 |
| G (Government Exp) | 2.069 | .204 | .548 | 10.133 | .000 |
| a. Dependent Variable: Y (Income) | | | | | | |

* ANOVA Table tests the significance of the regression model. We can see from the table that Sig (p value) =0.00.As p<0.05 our predictors are significantly better than would be expected by chance. The regression line predicted by the independent variables explains a significant amount of variance in the dependent variable.
* The regression sum of squares is considerably larger than the residual sum of squares, which indicates that about most of the variation in *the dependent* is explained by the model

* The fitted regression model is: **Y ̂ = 𝛼 ̂ + 𝛽 ̂I+γG = 3091.934 +1.670I+2.069G**, where α is the intercept of the line which indicates the average income when the investment and government expenditure is 0. As p<0.05 we may reject null hypothesis and conclude that both the coefficients are significantly different from 0.
* The unstandardized coefficients B column, gives us the coefficients of the independent variables in the regression equation. Each of these β values has an associated standard error indicating to what extent these values would vary across different samples and these standard errors are used to determine whether or not the β values significantly differ from 0.
* The standardised beta coefficient column shows the contribution that an individual variable makes to the model. The beta weight is the average amount the dependent variable increases when the independent variable increases by one standard deviation (all other independent variables are held constant).
* The confidence intervals are related to the p values such that the coefficient will not be statistically significant if the confidence interval includes 0.

Predictor Variables will be regressed on the instrumental variables, and the model-estimated values will then be used in place of the actual values of these problematic predictors when computing the model for the dependent.

| Variables Entered/Removedb | | | |
| --- | --- | --- | --- |
| Model | Variables Entered | Variables Removed | Method |
| 1 | Unstandardized Predicted Valuea | . | Enter |
| a. All requested variables entered. b. Dependent Variable: M (Money Supply) | | | |

We build a multiple linear regression model with government expenditure and Investment as the explanatory variables and Money Supply as the dependent variable

| Model Summaryb | | | | |
| --- | --- | --- | --- | --- |
| Model | R | R Square | Adjusted R Square | Std. Error of the Estimate |
| 1 | .957a | .916 | .910 | 352.377 |
| a. Predictors: (Constant), Unstandardized Predicted Value | | | | |
| b. Dependent Variable: M (Money Supply) | | | | |

* The Model Summary Table reports the strength of the relationship between the model and the dependent variable. The multiple Correlation Coefficient (R) is the linear correlation between the observed and model predicted values of the dependent variable. A moderate value (R=0.957) indicates a strong relationship.
* R-square measures the proportion of the variation in your dependent variable (Y) explained by your independent variables (X) for a linear regression model. Adjusted R-squared adjusts the statistic based on the number of independent variables in the model.

R2 is 0.916 meaning that 91.6% variance in the output is explained by money supply. In the case of a linear regression with more than one variable we prefer adjusted R-squared. Adjusted r2 compares the explanatory power of regression model that contains different number of predictors

While the ANOVA table is a useful test of the model's ability to explain any variation in the dependent variable, it does not directly address the strength of that relationship.

| ANOVAb | | | | | | |
| --- | --- | --- | --- | --- | --- | --- |
| Model | | Sum of Squares | df | Mean Square | F | Sig. |
| 1 | Regression | 1.907E7 | 1 | 1.907E7 | 153.581 | .000a |
| Residual | 1738378.599 | 14 | 124169.900 |  |  |
| Total | 2.081E7 | 15 |  |  |  |
| a. Predictors: (Constant), Unstandardized Predicted Value | | | | | | |
| b. Dependent Variable: M (Money Supply) | | | | | | |

* ANOVA Table tests the significance of the regression model. We can see from the table that Sig (p value) =0.00.As p<0.05 our predictors are significantly better than would be expected by chance. The regression line predicted by the independent variables explains a significant amount of variance in the dependent variable.
* The regression sum of squares is considerably larger than the residual sum of squares, which indicates that about most of the variation in *the dependent* is explained by the model

| Coefficientsa | | | | | | |
| --- | --- | --- | --- | --- | --- | --- |
| Model | | Unstandardized Coefficients | | Standardized Coefficients |  | |
| B | Std. Error | Beta | t | Sig. |
| 1 | (Constant) | -3046.617 | 619.495 |  | -4.918 | .000 |
| Unstandardized Predicted Value | .854 | .069 | .957 | 12.393 | .000 |
| a. Dependent Variable: M (Money Supply) | | | | | | |

* The fitted regression model is: **M ̂ = 𝛼 ̂ + 𝛽 ̂Y҇ = -3046.617 +0.854** where α is the Intercept of the line which indicates the average money supply when the income is 0. As p<0.05 we may reject null hypothesis and conclude that both the coefficients are significantly different from 0.
* The unstandardized coefficients B column, gives us the coefficients of the independent variables in the regression equation. Each of these β values has an associated standard error indicating to what extent these values would vary across different samples and these standard errors are used to determine whether or not the β values significantly differ from 0.
* The standardised beta coefficient column shows the contribution that an individual variable makes to the model. The beta weight is the average amount the dependent variable increases when the independent variable increases by one standard deviation (all other independent variables are held constant).

The confidence intervals are related to the p values such that the coefficient will not be statistically significant if the confidence interval includes 0.

| Coefficientsa | | | |
| --- | --- | --- | --- |
| Model | | 95.0% Confidence Interval for B | |
| Lower Bound | Upper Bound |
| 1 | (Constant) | -4375.301 | -1717.932 |
| Unstandardized Predicted Value | .706 | 1.002 |
| a. Dependent Variable: M (Money Supply) | | | |

The Income equation Y ̂ = 𝛼 ̂ + 𝛽 ̂I+γG is under identified. Money Supply equation M ̂ = 𝛼 ̂ + 𝛽 ̂Y is just identified.

Two Stage Least Squares can be summarised as follows:

* Standard linear regression models assume that errors in the dependent variable are uncorrelated with the independent variable(s). When this is not the case (for example, when relationships between variables are bidirectional), linear regression using ordinary least squares (OLS) no longer provides optimal model estimates.
* Two-stage least-squares regression uses instrumental variables that are uncorrelated with the error terms to compute estimated values of the problematic predictor(s) (the first stage), and then uses those computed values to estimate a linear regression model of the dependent variable (the second stage). Since the computed values are based on variables that are uncorrelated with the errors, the results of the two-stage model are optimal.
* One of the basic assumptions of the ordinary least-squares (OLS) regression model is that the values of the error terms are independent of the values of the predictors. When this "recursivity assumption" is broken, the two-stage least-squares (2SLS) model can help solve these problematic predictors. The 2SLS model assumes that there exist **instruments**, or secondary predictors, which are correlated with the problematic predictors but not with the error term.
* Given the existence of instrument variables, the 2SLS model:

1. Computes OLS models using the instrument variables as predictors and the problematic predictors as responses.

2. The model-estimated values from stage 1 are then used in place of the actual values of the problematic predictors to compute an OLS model for the response of interest.