ECONOMETRICS

(A case study demonstrating the applications of simple and multiple linear regression, indirect least squares and two-stage least squares to model consumption, production, elasticity of demand, price mechanism, national income and money supply.)

Submitted by: Vishal Kumar

Due Date: 24th April, 2016

Date of Submission: 24th April, 2016

Instructor's Remarks:

1. ECONOMETRICS (5 CASES)

1.1 SIMPLE AND MULTIPLE LINEAR REGRESSION

1.1.1 VALIDATING KEYNESIAN CONSUMPTION HYPOTHESIS

1.1.2 MODELING COBB DOUGLAS PRODUCTION FUNCTION

1.1.3 ESTIMATING PRICE ELASTICITY OF DEMAND

1.2 SIMULTANEOUS EQUATIONS MODEL

1.2.1 INDIRECT LEAST SQUARES – MODELING PRICE MECHANISM

1.2.2 TWO-STAGE LEAST SQUARES – MODELING QUANTITY-THEORY-KEYNESIAN APPROACH TO INCOME DETERMINATION

Econometrics

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.

Note: The data has been provided to you in the file Keynes' Consumption Function.xlsx.

Solution:

1. **Statement of Theory or Hypothesis:**

The Keynesian fundamental psychological law underlying the consumption function states that marginal propensity to consume (MPC) and marginal propensity to save (MPS) are greater than zero but less than one i.e. whenever national income rises by $1 part of this will be consumed and part of this will be saved.

In short, Keynes postulated that the marginal propensity to consume (MPC), the rate of change of consumption for a unit (say, a dollar) change in income, is greater than zero but less than 1.

2. **Specification of the Mathematical Model of Consumption:**

Although Keynes postulated a positive relationship between consumption and income. For simplicity, we might suggest the following form of the Keynesian consumption function:

$$Y = \beta_1 + \beta_2 X; \qquad 0 < \beta_2 < 1 \qquad \ldots\ldots(1)$$

where $Y = consumption\ expenditure,$

$X = income, and$

$\beta_1, \beta_2\ are\ known\ as\ the\ \boldsymbol{parameters}\ of\ the\ model.$

The slope coefficient β_2 measures the MPC. This equation, which states that consumption is linearly related to income, is an example of a mathematical model of the relationship between consumption and income that is called the consumption function in economics. A model is simply a set of mathematical equations. In the equation the variable appearing on the left side of the equality sign is called the dependent variable and the variable(s) on the right side are called the independent, or explanatory, variable(s). Thus, in the Keynesian consumption function, consumption (expenditure) is the dependent variable and income is the explanatory variable.

3. Specification of the Econometric Model of Consumption:

The purely mathematical model of the consumption function given in (1) is of limited interest to the econometrician, for it assumes that there is an exact or deterministic relationship between consumption and income. But relationships between economic variables are generally inexact. Thus, if we were to obtain data on consumption expenditure and disposable income we would not expect all observations to lie exactly on the straight line of the graph of (1) because, in addition to income, other variables affect consumption expenditure.

To allow for the inexact relationships between economic variables, the econometrician would modify the deterministic consumption function in (1) as follows:

$$Y = \beta_1 + \beta_2 X + u \qquad \dots \quad (2)$$

where u, is known as the disturbance, or error, term, is a random (stochastic) variable that has well-defined probabilistic properties. The disturbance term u may well represent all those factors that affect consumption but are not taken into account explicitly.

4. Obtaining the data:

To estimate the econometric model given in (2), that is, to obtain the numerical values of β_1 and β_2 we need data. The data is given in the table below:

Year	PFCE(Y)	GDP(X)	Year	PFCE(Y)	GDP(X)
1950	201,090.0	224,786.0	1978	509,819.0	631,839.0
1951	213,872.0	230,034.0	1979	498,384.0	598,974.0
1952	222,503.0	236,562.0	1980	543,243.0	641,921.0
1953	235,879.0	250,960.0	1981	566,866.0	678,033.0
1954	243,617.0	261,615.0	1982	572,536.0	697,861.0
1955	245,946.0	268,316.0	1983	616,974.0	752,669.0
1956	256,826.0	283,589.0	1984	634,757.0	782,484.0
1957	251,753.0	280,160.0	1985	661,249.0	815,049.0
1958	274,864.0	301,422.0	1986	682,116.0	850,217.0
1959	277,991.0	308,018.0	1987	705,495.0	880,267.0
1960	293,804.0	329,825.0	1988	749,530.0	969,702.0
1961	298,813.0	340,060.0	1989	786,725.0	1,029,178.0
1962	302,706.0	347,253.0	1990	821,863.0	1,083,572.0
1963	313,966.0	364,834.0	1991	839,593.0	1,099,072.0
1964	332,722.0	392,503.0	1992	861,245.0	1,158,025.0
1965	333,017.0	378,157.0	1993	898,682.0	1,223,816.0
1966	337,344.0	382,006.0	1994	942,359.0	1,302,076.0
1967	356,429.0	413,094.0	1995	999,729.0	1,396,974.0
1968	365,792.0	423,874.0	1996	1,077,445.0	1,508,378.0
1969	379,378.0	451,496.0	1997	1,109,656.0	1,573,263.0
1970	392,262.0	474,131.0	1998	1,181,797.0	1,678,410.0
1971	399,894.0	478,918.0	1999	1,253,643.0	1,786,526.0
1972	402,573.0	477,392.0	2000	1,292,986.0	1,864,773.0
1973	412,452.0	499,120.0	2001	1,367,758.0	1,972,912.0
1974	412,141.0	504,914.0	2002	1,397,069.0	2,047,733.0
1975	435,546.0	550,379.0	2003	1,493,871.0	2,222,591.0
1976	444,231.0	557,258.0	2004	1,579,255.0	2,389,660.0
1977	480,455.0	598,885.0	2005	1,689,861.0	2,604,532.0
			2006	1,800,874.0	2,871,118.0

Data on Y: Private Final Consumption Expenditure and X (Gross Domestic Product), both in 1999-2000 prices measured in rupee crore.

Data Source: National Accounts Statistics (2000, 2007, 2009), Central Statistical Organization, Ministry of Statistics and Programme Implementation, GOI.

5. Estimation of the Econometric Model:

Now that we have the data, our next task is to estimate the parameters of the consumption function:

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	103736.049	6587.369		15.748	.000
	GDP	.630	.006	.998	105.022	.000

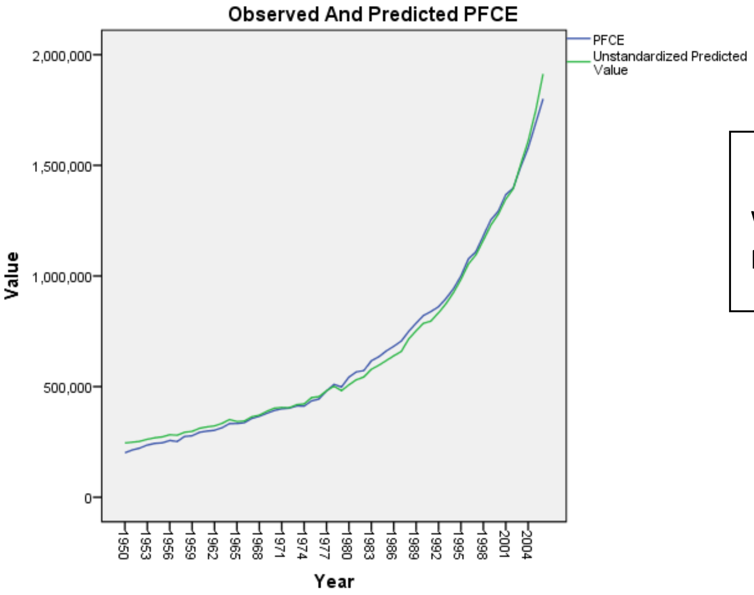
a. Dependent Variable: PFCE

The numerical estimates of the parameters give empirical content to the consumption function. We have used the statistical technique of regression analysis to obtain the estimates. Using this technique and the data given in Table we obtain the following estimates of β_1 and β_2

Thus, the estimated consumption function is:

$$\hat{Y} = 103736.049 + 0.63X + u$$

The hat on the Y indicates that it is an estimate. The representation is given as below.



Conclusion

We see prediction is close to observed PFCE. And it can be seen that $0 < \beta_2 < 1$

The coefficients of determination R square and adjusted R square are given as follows:

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.998[a]	.995	.995	30184.53297

a. Predictors: (Constant), GDP

b. Dependent Variable: PFCE

Conclusion: The Adjusted R square is 0.995 , which is very high. This clearly shows that the model is a good fit.

6. Hypothesis Testing:

Assuming that the fitted model is a reasonably good approximation of reality, we have to develop suitable criteria to find out whether the estimates obtained are in accord with the expectations of the theory that is being tested. According to "positive" economists like Milton

Friedman, a theory or hypothesis that is not verifiable by appeal to empirical evidence may not be admissible as a part of scientific enquiry. As noted earlier, Keynes expected the MPC to be positive but less than 1.

The confirmation or refutation of economic theories on the basis of sample evidence is based on a branch of statistical theory known as statistical inference (hypothesis testing).

We define

$$H_o: \beta_1 = MPC = 0$$

$$H_1: \beta_1 \neq MPC \neq 0$$

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	10049171516259.693	1	10049171516259.693	11029.640	.000[b]
	Residual	50110831669.812	55	911106030.360		
	Total	10099282347929.506	56			

a. Dependent Variable: PFCE

b. Predictors: (Constant), GDP

Conclusion: The p value < 0.05. Hence we reject null hypothesis and conclude that MPC is significant.

7. Forecasting or Prediction:

If the chosen model does not refute the hypothesis or theory under consideration, we may use it to predict the future value(s) of the dependent, or forecast, variable Y on the basis of known or expected future value(s) of the explanatory, or predictor, variable X.

We would use the model:

$$\hat{Y} = 103736.049 + 0.63X + u$$

To forecast the values of Y for corresponding values of X.

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

$$M = \frac{1}{1 - MPC}$$

The critical value in this computation is MPC, for the multiplier depends on it. And this estimate of the MPC can be obtained from regression models such as our estimated model. Thus, a quantitative estimate of MPC provides valuable information for policy purposes. Knowing MPC, one can predict the future course of income, consumption expenditure, and employment following a change in the government's fiscal policies.

8. **Use of the Model for Control or Policy Purposes:**

A. The sample regression function

$$\hat{Y} = 103736.049 + 0.63X + u$$

can be used for policy formulation. Consider the following examples.

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.

$$25{,}00{,}000 = 103736.049 + 0.63X$$

which gives X = 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.

B. 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

$$M = \frac{1}{1 - MPC}$$

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. Labor 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 (R^2, p-values etc.) should be provided.

Solution:

The Cobb–Douglas production function, in its stochastic form, may be expressed as:

$$Y_i = \beta_i X_{2i}^{\beta_2} X_{3i}^{\beta_3} e^{u_i}; \qquad \ldots (2.1)$$

Where $Y = \ total\ production$

$X_2 = labour\ input$

$X_3 = capital\ input$

$u = stochastic\ disturbance\ term$

$e = base\ of\ natural\ logarithm$

From Eq. (2.1) it is clear that the relationship between output and the two inputs is nonlinear. However, if we log-transform this model, we obtain:

$$lnY_i = ln\beta_1 + \beta_2 lnX_{2i} + \beta_3 lnX_{3i} + u_i$$
$$lnY_i = \beta_0 + \beta_2 lnX_{2i} + \beta_3 lnX_{3i} + u_i$$

Where $\beta_0 = ln\beta_1$

Thus written, the model is linear in the parameters β_0, β_2, and β_3 and is therefore a linear regression model. Notice, though, it is nonlinear in the variables Y and X but linear in the logs of these variables.

1. The properties of the Cobb–Douglas production function are quite well known:
2. β_2 is the (partial) elasticity of output with respect to the labor input, that is, it measures the percentage change in output for, say, a 1 percent change in the labor input, holding the capital input constant .
3. Likewise, β_3 is the (partial) elasticity of output with respect to the capital input, holding the labour input constant.

The sum (β_2 + β_3) gives information about the returns to scale, that is, the response of output to a proportionate change in the inputs. If this sum is 1, then there are constant returns to scale, that is, doubling the inputs will double the output, tripling the inputs will triple the output, and so on. If the sum is less than 1, there are decreasing returns to scale—doubling the inputs will less than double the output. Finally, if the sum is greater than 1, there are increasing returns to scale—doubling the inputs will more than double the output.

Our data under consideration is:

Real gross product (millions of New Taiwan $)	Labor days (millions of days)	Real capital input (millions of New Taiwan $)
Y	**L**	**K**
16607.7	275.5	17803.7
17511.3	274.4	18096.8
20171.2	269.7	18271.8
20932.9	267	19167.3
20406	267.8	19647.6
20831.6	275	20803.5
24806.3	283	22076.6
26465.8	300.7	23445.2
27403	307.5	24939
28628.7	303.7	26713.7
29904.5	304.7	29957.8
27508.2	298.6	31575.9
29035.5	295.5	33474.5
29281.5	299	34821.8
31535.8	288.1	41794.3

Running the regression model, our estimated Cobb Douglas function is:

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients		
		B	Std. Error	Beta	t	Sig.
1	(Constant)	-3.339	2.449		-1.363	.198
	Ln_L	1.499	.540	.373	2.777	.017
	Ln_K	.490	.102	.644	4.801	.000

a. Dependent Variable: Ln_Y

$$ln\hat{Y}_i = -3.339 + 1.499\ \ln X_{2i} + 0.490 lnX_{3i}$$

We see that in the Taiwanese agricultural sector for the period 1958–1972 the output elasticities of labor and capital were 1.4988 and 0.4899, respectively. In other words, over the period of study, holding the capital input constant, a 1 percent increase in the labour input led on the average to about a 1.5 percent increase in the output. Similarly, holding the labour input constant, a 1 percent increase in the capital input led on the average to about a 0.5 percent increase in the output. Adding the two output elasticities, we obtain 1.9887, which gives the value of the returns to scale parameter. As is evident, over the period of the study, the Taiwanese agricultural sector was characterized by increasing returns to scale.

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.943[a]	.889	.871	.07480

a. Predictors: (Constant), Ln_K, Ln_L

b. Dependent Variable: Ln_Y

Conclusion: From a purely statistical viewpoint, the estimated regression line fits the data quite well. The R^2 value of 0.889 means that about 89 percent of the variation in the (log of) output is explained by the (logs of) labour and capital.

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:

$$Q_d = exp(\beta_0 + \beta_1 lnP)$$

where Q_d is the quantity demanded of the given commodity and P is the price of the given commodity.

Solution:

Price elasticity of demand is a measure used to show the responsiveness, or elasticity, of the quantity demanded of a good or service to a change in its price. More precisely, it gives the percentage change in quantity demanded in response to a one percent change in price (holding constant all the other determinants of demand, such as income).

Price elasticities are almost always negative, although analysts tend to ignore the sign even though this can lead to ambiguity. Only goods which do not conform to the law of demand, such as Giffen goods, have a positive price elasticities of demand. In general, the demand for a good is said to be inelastic (or relatively inelastic) when the price elasticity of demand is less than one (in absolute value): that is, changes in price have a relatively small effect on the quantity of the good demanded. The demand for a good is said to be elastic (or relatively elastic) when its price elasticity of demand is greater than one (in absolute value): that is, changes in price have a relatively large effect on the quantity of a good demanded.

It is a measure of responsiveness of the quantity of a raw good or service demanded to changes in its price. The formula for the coefficient of price elasticity of demand for a good is:

$$e_p = \frac{\frac{dQ}{Q}}{\frac{dP}{P}}$$

The above formula usually yields a negative value, due to the inverse nature of the relationship between price and quantity demanded, as described by the "law of demand". For example, if the price increases by 5% and quantity demanded decreases by 5%, then the elasticity at the initial price and quantity = −5%/5% = −1. Although the price elasticity of demand is negative for the vast majority of goods and services, economists often refer to price elasticity of demand as a positive value (i.e., in absolute value terms).

For a regression model, the price elasticity of demand is given as:

$$e_p = \frac{\partial Y}{\partial X} . \frac{\bar{X}}{\bar{Y}}$$

In our case, the model could be written as:

$$lnQ_d = \beta_0 + \beta_1 ln\, P$$

$$\frac{1}{Q_d}\frac{\partial Q_d}{\partial P} = \frac{\beta_1}{P}$$

$$\frac{\partial Q_d}{\partial P} = \frac{\beta_1 Q_d}{P}$$

$$e_p = \beta_1$$

The data in use is given as follows:

Price (in INR)	Quantity (in Kilograms)
10	43.80
11.5	40.67
12	39.77
12.8	38.43
12.9	38.27
13	38.11
13.3	37.66
15	35.33
18	32.08
20	30.33
22	28.84
22.3	28.63
22.8	28.30
25	26.95

On running regression analysis on the model the estimates are:

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	5.000	.000		23074.058	.000
	Ln_Price	-.530	.000	-1.000	-6785.789	.000

a. Dependent Variable: Ln_Quantity

Conclusion: So $\beta_1 = -0.53$. Taking absolute value, the price elasticity of demand is 0.53.

We conclude that for $\mathbf{1}\%$ change in price there is $\mathbf{53}\%$ change in quantity demanded.

Case 4: Indirect Least Squares Estimation:

Consider the following price mechanism model for some commodity:

$$\text{Demand Function: } Q_t^d = \alpha_0 + \alpha_1 P_t + \alpha_2 Y_t + u_{1t}\ ;\ \alpha_1 < 0 \text{ and } \alpha_2 > 0$$

$$\text{Supply Function: } Q_t^s = \beta_0 + \beta_1 P_t + u_{2t}\ ;\quad \beta_1 > 0$$

$$\text{Equilibrium Function: } Q_t^s = Q_t^d = Q_t$$

where $Qt\ d$ 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:

From the given equations for demand, supply and Equilibrium, we get:

$$Q_t = \alpha_0 + \alpha_1 P_t + \alpha_2 Y_t + u_{1t}\ ;\ \alpha_1 < 0 \text{ and } \alpha_2 > 0 \ldots (1)$$

$$Q_t = \beta_0 + \beta_1 P_t + u_{2t}\ ;\quad \beta_1 > 0 \ldots (2)$$

Now, multiply equation (1) by λ and equation (2) by (1-λ) for some λ ϵ [0, 1] and adding them together we get,

$$\lambda Q_t + (1-\lambda) Q_t = [\lambda\alpha_0 + (1-\lambda)\beta_0] + [\lambda\alpha_1 + (1-\lambda)\beta_1] P_t + \lambda\alpha_2 Y_t + [\lambda u_{1t} + (1-\lambda) u_{2t}]$$

$$Q_t = \gamma_0 + \gamma_1 P_t + \gamma_2 Y_t + u_t \ldots (3)$$

We can see that neither equation (1) nor (3) can be the supply equation as income doesn't affect the supply. Hence (2) is the supply equation. But still no way to identify the true demand equation between (1) and (3). This situation is called Partial Identification. So we will estimate parameters only for supply equation i.e. we will estimate $\beta_0$and β_1.

Reduced Form: Using **equilibrium condition** we can obtain the reduced form of price mechanism as follows:

$$\alpha_0 + \alpha_1 P_t + \alpha_2 Y_t + u_{1t} = \beta_0 + \beta_1 P_t + u_{2t}$$

$$P_t = \frac{(\beta_0 - \alpha_0)}{(\alpha_1 - \beta_1)} - \frac{\alpha_2}{(\alpha_1 - \beta_1)} Y_t + \frac{(u_{2t} - u_{1t})}{(\alpha_1 - \beta_1)}$$

$$P_t = \pi_0 + \pi_1 Y_t + v_{1t} \ldots (4)$$

Regressing $P_t$on Y_t i.e. Price on income, we get:

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.494[a]	.244	.217	9.10670

a. Predictors: (Constant), Income

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	90.961	4.051		22.456	.000
	Income	.001	.000	.494	3.007	.006

a. Dependent Variable: Price

Conclusion: By fitting the linear regression, we get p-value (.006 & 0.00) < 0.05, hence we conclude that coefficients come out to be significant. We can see that $\pi_0 = 90.961$ *and* $\pi_1 = 0.001$. Hence, the reduced form as given in eq. (4) can be written as:

$$P_t = 90.961 + 0.001\, Y_t + v_{1t}$$

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	749.770	1	749.770	9.041	.006[b]
	Residual	2322.096	28	82.932		
	Total	3071.867	29			

a. Dependent Variable: Price

b. Predictors: (Constant), Income

Conclusion: Since p-value (0.006) < 0.05, hence the results are significant at the 5% level of significance

And using the **demand function and equilibrium condition**, we have,

$$Q_t = \alpha_0 + \alpha_1\{\frac{(\beta_0 - \alpha_0)}{(\alpha_1 - \beta_1)} + \frac{\alpha_2}{(\alpha_1 - \beta_1)} Y_t + \frac{(u_{2t} - u_{1t})}{(\alpha_1 - \beta_1)}\} + \alpha_2 Y_t + u_{1t}$$

$$Q_t = \frac{(\alpha_1\beta_0 - \alpha_{0\beta_1})}{(\alpha_1 - \beta_1)} - \frac{\alpha_2\beta_1}{(\alpha_1 - \beta_1)} Y_t + \frac{(u_{2t}\alpha_1 - u_{1t}\beta_1)}{(\alpha_1 - \beta_1)}$$

$$Q_t = \pi_2 + \pi_3 Y_t + v_{2t} \quad \dots (5)$$

Equation (4) and (5) are the reduced forms.

Regressing Q_t on Y_t i.e. Quantity on income, we get:

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.969[a]	.940	.938	3.50870

a. Predictors: (Constant), Income

Conclusion: 96.9% of response variable can be explained by the explanatory variable

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	5384.760	1	5384.760	437.395	.000[b]
	Residual	344.707	28	12.311		
	Total	5729.467	29			

a. Dependent Variable: Quantity

b. Predictors: (Constant), Income

Conclusion: Since p-value (0.000) < o.o5, hence the results are significant at the 5% level of significance.

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	59.771	1.561		38.299	.000
	Income	.002	.000	.969	20.914	.000

a. Dependent Variable: Quantity

Conclusion: We see p-value (.000 & 0.00) < 0.05, hence we conclude that coefficients come out to be significant. We can see that $\pi_2 = 59.771\ and\ \pi_1$ =.002. Hence, the reduced form as given in eq. (5) can be written as:

$$Q_t = 59.771 + 0.002Y_t + v_{2t}$$

Reduced form model can be analyzed by OLSE, thus we get estimates:

$$\pi_0 = \frac{(\beta_0 - \alpha_0)}{(\alpha_1 - \beta_1)};\ \pi_1 = -\frac{\alpha_2}{(\alpha_1 - \beta_1)}$$

$$\pi_2 = \frac{(a_1\beta_0 - \alpha_{0\beta_1})}{(\alpha_1 - \beta_1)};\ \pi_3 = -\frac{\alpha_2\beta_1}{(\alpha_1 - \beta_1)}$$

Indirect Least Squares method (ILS):

Obtain OLS estimates of $\pi_0, \pi_1, \pi_2\ and\ \pi_3$ using the reduced form equations.

As supply equation is exactly identified. Indirect Least Squares estimators of the structural coefficients in it viz., $\beta_1\ and\ \beta_2$ are obtained as follows:

$$\widehat{\beta_{1,ILS}} = \frac{\widehat{\pi_3}}{\widehat{\pi_1}} and\ \widehat{\beta_{0,ILS}} = \widehat{\pi_2} - \frac{\widehat{\pi_3}}{\widehat{\pi_1}}\widehat{\pi_0}$$

Conclusion: $\widehat{\beta_{1,ILS}} = 2\ and\ \widehat{\beta_{0,ILS}} = -122.151$

Thus, we get Supply Function as : $Q_t^s = -122.151 + 2 * P_t + u_{2t}$; $\beta_1 > 0$

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.

$$\text{Income Function: } Y_t = \beta_{10} + \beta_{11}M_t + \gamma_{11}I_t + \gamma_{12}G_t + u_{1t}$$

$$\text{Money Supply Function: } M_t = \beta_{20} + \beta_{21}Y_t + u_{2t}$$

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:

Suppose there are G endogenous variables $y_1, y_2, \dots, y_G$ and K exogenous varaibles $x_1, x_2, \dots, x_K$ following G structural equations represent SEM: for $t = 1, \dots, n$

$$\beta_{11}y_{1t} + \beta_{12}y_{2t} + \dots + \beta_{1G}y_{Gt} + \gamma_{11}x_{1t} + \dots + \gamma_{1K}x_{kt} = \epsilon_{1t}$$

$$\beta_{21}y_{1t} + \beta_{22}y_{2t} + \dots + \beta_{2G}y_{Gt} + \gamma_{21}x_{1t} + \dots + \gamma_{2K}x_{kt} = \epsilon_{2t}$$

$$\begin{array}{cccccc} . & . & . & . & . & . \\ . & . & . & . & . & . \\ . & . & . & . & . & . \end{array}$$

$$\beta_{G1}y_{1t} + \beta_{G2}y_{2t} + \dots + \beta_{GG}y_{Gt} + \gamma_{G1}x_{1t} + \dots + \gamma_{GK}x_{kt} = \epsilon_{Gt}$$

Equivalently, $\quad S: By_t + \Gamma x_t = \epsilon_t \quad ; t = 1, \dots, n$

Reduced form: $\quad R: y_t = \Gamma x_t + v_t \quad ; t = 1, \dots, n$

$$\pi = -B^{-1}\Gamma$$

$$B\pi = -\Gamma$$

Without loss of generality let us consider first equation of SEM. Suppose there are $G_\Delta$endogenous variables and $K_\Delta$exogenous variables in 1^{st} equation and the rest $(G - G_\Delta)$ endogenous and $(k - k_\Delta)$ exogenous variables have coefficients as zero. For the 1^{st} equation we have:

$$(\beta_\Delta \;\; 0_\Delta) = -(\gamma_* \;\; 0_{**}) \qquad \text{where} \quad \pi = \begin{bmatrix} \pi_{\Delta *} & \pi_{\Delta **} \\ \pi_{\Delta\Delta *} & \pi_{\Delta\Delta **} \end{bmatrix}$$

- ⇨ $\beta_\Delta \pi_{\Delta **} = 0_{**}$
- ⇨ $(1, \beta_{12}, \dots, \beta_{1G\Delta})\pi_{\Delta **} = 0_{**}$

Thus $(G_\Delta - 1)$ unknown coefficients in $\beta_\Delta$can be uniquely obtained if $Rank(\Pi_{\Delta **}) = (G_\Delta - 1)$.

- ⇨ $(G_\Delta - 1) \le min((G_\Delta, K_{**})$
- ⇨ $(G_\Delta - 1) \le K_{**}$
- ⇨ $(G_\Delta - 1) \le (k - K_*)$

This is called Order condition of identifiability. It is only necessary condition not sufficient.

$$(k - K_*) < (G_\Delta - 1)\text{: equation is unidentified}$$

$$(k - K_*) = (G_\Delta - 1)\text{: equation is exactly identified}$$

$$(k - K_*) > (G_\Delta - 1)\text{: equation is over} - \text{identified}$$

Working rule for rank condition can be described as follows:

Step1: Write model in a tabular form by putting crosses if variable is present and 0 if it is absent.

Step2: For equation under study, mark the 0's and pick up the corresponding columns suppressing that row.

Step3: If we can choose $(G - 1)$ rows and $(G - 1)$ columns that are not at all zero then it can be identified.

In the given set of equations we have $Y_t$and $M_t$as the endogenous variables:

$$Y_t = \beta_{10} + \beta_{11} M_t + \gamma_{11} I_t + \gamma_{12} G_t + u_{1t}$$

$$M_t = \beta_{20} + \beta_{21} Y_t + u_{2t}$$

Hence $G = 2$ and $K = 2$

Step1:

Equation	Y_t	M_t	I_t	G_t	$G_\Delta - 1$	$k - K_*$	Identification
1	X	X	X	X	$2 - 1 = 1$	$2 - 2 = 0$	Unidentificatied
2	X	X	0	0	$2 - 1 = 1$	$2 - 0 = 2$	Over-identified

Step2:

Equation	I_t	G_t
1	X	X

Conclusion:

1. In step1 using Order conditions of identifiability, we can see that Income function is unidentified whereas Money supply function over-identified.
2. In step2 using rules for rank condition, we need to verify if we can choose G-1=2-1=1 rows and columns with not all zeros. Clearly it is possible. Hence equation 2 is identified

Two Stage Least Square Estimation (2-SLS):

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

Step1: 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

$$Y_t = a_0 + a_1 \mathrm{I}_t + a_2 G_t + \epsilon_t$$

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.996[a]	.992	.991	124.14796

a. Predictors: (Constant), Govt_Exp, Investment

b. Dependent Variable: Income

Conclusion: R^2 is 0.992, 99.2% variability is explained by the model.

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	26127940.697	2	13063970.349	847.610	.000[b]
	Residual	200365.303	13	15412.716		
	Total	26328306.000	15			

a. Dependent Variable: Income

b. Predictors: (Constant), Govt_Exp, Investment

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	3091.934	157.835		19.590	.000
	Investment	1.670	.190	.475	8.786	.000
	Govt_Exp	2.069	.204	.548	10.133	.000

a. Dependent Variable: Income

Conclusion: We get $\widehat{Y_t} = 3091.934 + 1.670 * \mathrm{I}_t + 2.069 * G_t$

Step2: Replace $Y\ t$ by $Y\hat{}\ t$ in the money supply equation and build the following linear regression model.

$$M_t = \beta_{20} + \beta_{21}\widehat{Y_t} + u_{2t}$$

Using OLS estimate β_{20} and β_{21}.

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.957[a]	.916	.910	352.37750

a. Predictors: (Constant), Unstandardized Predicted Value

Conclusion: R^2 is 0.916, 91.6% variability is explained by the model.

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	19070135.151	1	19070135.151	153.581	.000[b]
	Residual	1738378.599	14	124169.900		
	Total	20808513.750	15			

a. Dependent Variable: Money

b. Predictors: (Constant), Unstandardized Predicted Value

Conclusion: p-value is 0.0 < 0.05, which is significant, hence the overall regression is significant.

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	-3046.617	619.495		-4.918	.000
	Unstandardized Predicted Value	.854	.069	.957	12.393	.000

a. Dependent Variable: Money

Conclusion: Thus we get $\beta_{20} = -3047.082$ and $\beta_{21} = .854$. Hence, Money supply function as:

$$M_t = -3047.082 + 0.854 * \widehat{Y_t} + u_{2t}$$

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.