Linear_regression using instrumental variables

Contents

1. Instrumental variables

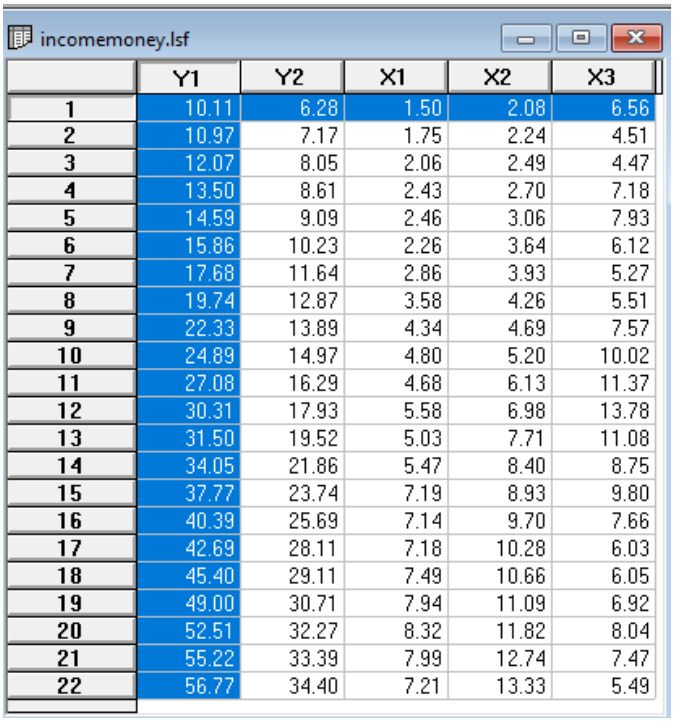
Instrumental variable procedures are needed when some regressors are endogenous (correlated with the error term). An example of instrumental variables is when wages and education jointly depend on ability which is not directly observable, but we can use available test scores to proxy for ability.

In this example (Gujarati, 1995) we examine the relationship between income and money supply. Information is also available on investment expenditure and government spending on goods and services. The data given in **incomemoney.lsf** are selected macro-economic data for the USA for the period 1970 to 1991. The data and syntax files can be found in the **MVABOOK examples\Chapter2** folder.

The variables are:

- y_1 = income
- y_2 = money supply
- x_1 = investment expenditure
- x_2 = government spending on goods and services
- x_3 = interest rate on 6-month Treasury bills in %.

All variables except for x_3 are measured in billions of dollars.



incomemoney.lsf

	Y1	Y2	X1	X2	X3
1	10.11	6.28	1.50	2.08	6.56
2	10.97	7.17	1.75	2.24	4.51
3	12.07	8.05	2.06	2.49	4.47
4	13.50	8.61	2.43	2.70	7.18
5	14.59	9.09	2.46	3.06	7.93
6	15.86	10.23	2.26	3.64	6.12
7	17.68	11.64	2.86	3.93	5.27
8	19.74	12.87	3.58	4.26	5.51
9	22.33	13.89	4.34	4.69	7.57
10	24.89	14.97	4.80	5.20	10.02
11	27.08	16.29	4.68	6.13	11.37
12	30.31	17.93	5.58	6.98	13.78
13	31.50	19.52	5.03	7.71	11.08
14	34.05	21.86	5.47	8.40	8.75
15	37.77	23.74	7.19	8.93	9.80
16	40.39	25.69	7.14	9.70	7.66
17	42.69	28.11	7.18	10.28	6.03
18	45.40	29.11	7.49	10.66	6.05
19	49.00	30.71	7.94	11.09	6.92
20	52.51	32.27	8.32	11.82	8.04
21	55.22	33.39	7.99	12.74	7.47
22	56.77	34.40	7.21	13.33	5.49

We consider the model

$$y_1 = \alpha_1 + \beta_1 y_2 + \gamma_1 x_1 + \gamma_2 x_2 + u_1$$
$$y_2 = \alpha_2 + \beta_2 y_1 + u_2$$

Although not explicitly included in the model, we can use x_3 as an instrumental variable in addition to x_1 and x_2 to estimate the model.

2. Two-stage least squares estimation

To estimate the model, we use Two-Stage Least Squares (TSLS). Two-stage least-squares (TSLS) is particularly useful for estimating econometric models of the form

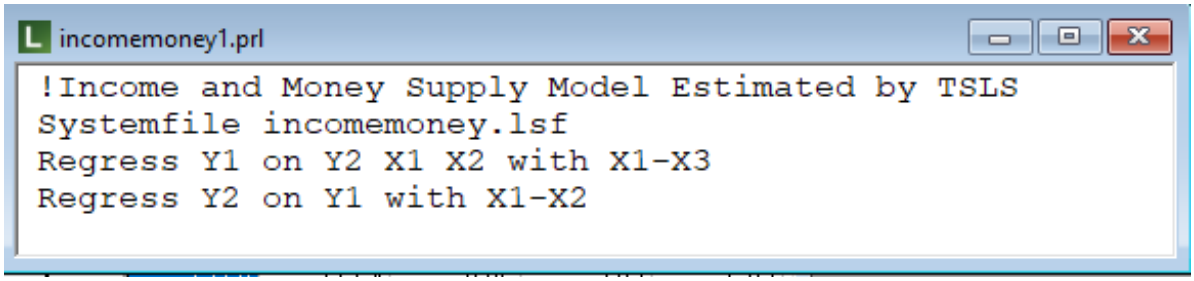
$$\mathbf{y} = \mathbf{B}\mathbf{y} + \mathbf{\Gamma}\mathbf{x} + \mathbf{u},$$

where $\mathbf{y} = (y_1, y_2, ..., y_p)$ is a set of endogenous or jointly dependent variables, $\mathbf{x} = (x_1, x_2, ..., x_q)$ is a set of exogenous or predetermined variables uncorrelated with the error terms $\mathbf{u} = (u_1, u_2, ..., u_p)$, and $\mathbf{B}$ and $\mathbf{\Gamma}$ are parameter matrices.

A typical feature of the above model is that not all y-variables and not all x-variables are included in each equation.

A necessary condition for identification of each equation is that, for every y-variable on the right side of the equation, there must be at least one x-variable excluded from that equation. There is also a sufficient condition for identification, the so-called rank condition, but this is often difficult to apply in practice. For further information on identification of interdependent systems, see, e.g., Goldberger (1964, pp. 313-318).

The syntax is given in **incomemoney1.prl**:



Partial results for this analysis are shown below.

```
Estimated Equations

      Y1 = 0.773 + 0.493*Y2 + 0.292*X1 + 2.698*X2 + Error, R² = 0.996
 Standerr  (0.700) (0.411)    (0.416)    (0.957)
 t-values   1.104   1.200      0.701      2.818
 P-values   0.283   0.245      0.492      0.011

 Error Variance = 1.076

 Instrumental Variables: X1 X2 X3

      Y2 = 0.346 + 0.614*Y1 + Error, R² = 0.996
 Standerr  (0.305) (0.00906)
 t-values   1.132   67.793
 P-values   0.270   0.000

 Error Variance = 0.404

 Instrumental Variables: X1 X2
```

As expected, the estimate of β_2 is highly significant. In the income equation, only the estimate of γ_2 (government spending) is statistically significant at a 5% level.