Part 2： Simultaneous Equation Models (SEM)

Chapter 20. How to Estimate SEM ?

20.1 Approaches to Estimation

Approaches to Estimation

In order to estimate the structural SEM, two approaches can be adopted:

- **Single equation method**, also known as **limited information methods**.

> estimate each equation in SEM one by one, considering only the constraints in that equation

- **System method** , also known as **full information method**

> estimate all the equations in the model simultaneously, taking into account all the constraints in the SEM

Approaches to Estimation

Instrumental variables are often used to estimate simultaneous equation problems, mainly including three IV techniques for **System method** :

- **Three-stage least squares (3SLS)**: Applicable in a few cases
- **Generalized moment method(GMM)**: It is commonly used for dynamic model problems
- **Full information maximum likelihood(FIML)**: It has much theoretical value and it brings no advantage over 3SLS, but is much more complicated to compute.

Consider the following SEM:

$$
\begin{cases}
Y_{t1} - \gamma_{21}Y_{t2} - \gamma_{31}Y_{t3} - \beta_{01} - \beta_{11}X_{t1} = u_{t1} \\
Y_{t2} - \gamma_{32}Y_{3t} - \beta_{02} - \beta_{12}X_{1t} - \beta_{22}X_{2t} = u_{t2} \\
-\gamma_{13}Y_{t1} + Y_{t3} - \beta_{03} - \beta_{13}X_{1t} - \beta_{23}X_{2t} = u_{t3} \\
-\gamma_{14}Y_{t1} - \gamma_{24}Y_{t2} + Y_{t4} - \beta_{04} - \beta_{34}X_{t3} = u_{t4}
\end{cases}
$$

- If you focus only on estimating the third equation, we can use the **single equation method** , which the variables Y_2, Y_4, X_3 were excluded from the estimation.
- If you want to estimate all four equations **simultaneously**, you should use the **system method**, and it will take into account all the constraints on multiple equations in the system.

Approaches to Estimation

In order to use all information of SEM, it is most desirable to apply **system method**, such as **full information maximum likelihood** (FIML).

In practice, however, **systems method** are not commonly used for the following main reasons:

1. The computational burden is too great.
2. Systematic methods such as FIML often bring highly nonlinearity on parameters , which are difficult to determine and caculate.
3. If there is one or more **specification error** in SEM (eg. an incorrect functional form or missing variables), the error will be passed to the remaining equations. As a result, the system method becomes very sensitive to the specification errors.

20.2 Least squares approach (LS)

OLS Approach with recursive model

Recursive model : also known as the **triangle model** or **causality model**.

> The simultaneous disturbance terms in different equations are unrelated, and each equation exhibits a one-way causal dependence.

Consider the following structural SEM:

$$
\begin{cases}
Y_{t1} = & +\beta_{01} + \beta_{11}X_{t1}+\beta_{21}X_{t2}+u_{t1} \\
Y_{t2} = +\gamma_{12}Y_{1t} & +\beta_{02} + \beta_{12}X_{t1}+\beta_{22}X_{t2}+u_{t2} \\
Y_{t3} = +\gamma_{13}Y_{t1}+\gamma_{23}Y_{t2} & +\beta_{03} + \beta_{13}X_{t1}+\beta_{23}X_{t2}+u_{t3}
\end{cases}
$$

OLS Approarch with recursive model

$$\begin{cases} Y_{t1} = & +\beta_{01} + \beta_{11}X_{t1} + \beta_{21}X_{t2} + u_{t1} \\ Y_{t2} = +\gamma_{12}Y_{1t} & +\beta_{02} + \beta_{12}X_{t1} + \beta_{22}X_{t2} + u_{t2} \\ Y_{t3} = +\gamma_{13}Y_{t1} + \gamma_{23}Y_{t2} & +\beta_{03} + \beta_{13}X_{t1} + \beta_{23}X_{t2} + u_{t3} \end{cases}$$

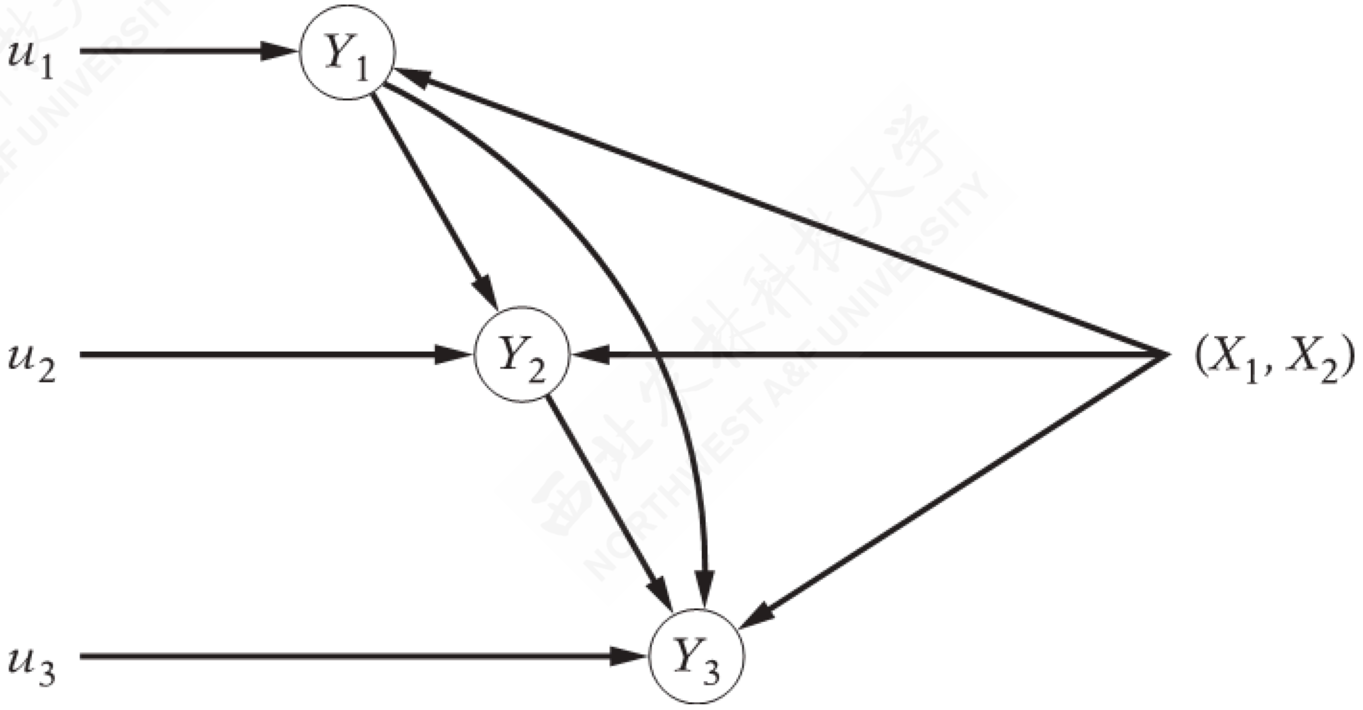
It is easy to find that the contemporaneous disturbance terms in different equations are irrelevant (namely **zero contemporaneous correlation**) :

$$cov(u_{t1}, u_{t2}) = cov(u_{t1}, u_{t3}) = cov(u_{t2}, u_{t3}) = 0$$

- Since the first equation' right-hand side only contains exogenous variables, and are not correlated with disturbance terms, so this equation satisfies the CLRM and OLS can be applied directly to it.
- because $cov(u_{t1}, u_{t2}) = 0$, and $cov(Y_{t1}, u_{t2}) = 0$. Thus OLS can be applied directly to it.
- because $cov(u_{t1}, u_{t3}) = 0$, and $cov(Y_{t1}, u_{t3}) = 0$. Also $cov(u_{t1}, u_{t3}) = 0$, and $cov(Y_{t2}, u_{t3}) = 0$. Thus OLS can be applied directly to it.

OLS Approarch with recursive model

We can also visualize it graphically:



SEM in recursive form

OLS Approach with recursive model

Let's look at the **wage-price model**:

$$\begin{cases} P_t = \beta_0 + \beta_1 UN_t + \beta_2 R_t + \beta_3 M_t + u_{t2} & \text{(price equation)} \\ W_t = \alpha_0 + \alpha_1 UN_t + \alpha_2 P_t + u_{t1} & \text{(wage equation)} \end{cases}$$

Where:

- W, the money wage rate;
- UN, unemployment, %;
- P, price rate;
- R, the cost of capital rate;
- M, import price change rate of raw materials.

20.3 Indirect least squares (ILS)

ILS approach with Just Identification model

For a just or exactly identified structural equation, the method of obtaining the estimates of the structural coefficients from the OLS estimates of the reduced-form coefficients is known as the method of **Indirect Least Squares** (ILS), and the estimates thus obtained are known as the **indirect least squares estimates**.

ILS involves the following three steps:

- Step 1. We first obtain the reduced-form SEM.
- Step 2. We apply **OLS** to the reduced-form SEM individually.
- Step 3. We obtain estimates of the original structural coefficients from the estimated reduced-form coefficients obtained in Step 2.

> If an equation is **exactly identified**, there is one-to-one mapping between the structural and reduced coefficients.

Case demo: US crop supply and demand

The variables in the US crop supply and demand case are illustrated below

variables in the model

vars	label	note
Q	Crop yield index	(1996=100)
P	Agricultural products purchasing prices index	(1990-1992=100)
X	Capital personal consumption expenditure	(In 2007 dollars)

Case demo: data set

The data for US crop supply and demand case show here:

sample data (n=30)

year	Q	P	X
1975	66	88	4789
1976	67	87	5282
1977	71	83	5804
1978	73	89	6417
1979	78	98	7073
1980	75	107	7716
1981	81	111	8439

Showing 1 to 7 of 30 entries

Previous 1 2 3 4 5 Next

Case demo: structural SEM

So we can construct the following structural SEM:

$$\begin{cases} Q = \alpha_0 + \alpha_1 P_t + \alpha_2 X_t + u_{t1} & (\alpha_1 < 0, \alpha_2 > 0) \quad \text{(demand function)} \\ Q = \beta_0 + \beta_1 P_t + u_{t2} & (\beta_1 > 0) \quad \text{(supply function)} \end{cases}$$

where:

- Q =Crop yield index;
- P =Agricultural products purchasing prices index;
- X =Capital personal consumption expenditure.

Case demo: reduced SEM

Thus we can obtain the reduced SEM:

$$\begin{cases} P_t = \pi_{11} + \pi_{21} X_t + w_t & \text{(eq1)} \\ Q_t = \pi_{12} + \pi_{22} X_t + v_t & \text{(eq2)} \end{cases}$$

and the relationship between structural and reduced coefficients is:

$$\begin{cases} \pi_{11} = \dfrac{\beta_0 - \alpha_0}{\alpha_1 - \beta_1} \\ \pi_{21} = -\dfrac{\alpha_2}{\alpha_1 - \beta_1} \\ w_t = \dfrac{u_{2t} - u_{t1}}{\alpha_1 - \beta_1} \end{cases} \qquad \begin{cases} \pi_{12} = \dfrac{\alpha_1 \beta_0 - \alpha_0 \beta_1}{\alpha_1 - \beta_1} \\ \pi_{22} = -\dfrac{\alpha_2 \beta_1}{\alpha_1 - \beta_1} \\ v_t = \dfrac{\alpha_1 u_{t2} - \beta_1 u_{1t}}{\alpha_1 - \beta_1} \end{cases}$$

Case demo: reduced coefficients

For the above reduced SEM, we can use OLS method to obtain the estimated coefficients:

$$
\begin{cases}
\widehat{\pi}_{21} = \dfrac{\sum p_t x_t}{\sum x_t^2} & \text{(slope of the reduced price eq)} \\
\widehat{\pi}_{11} = \overline{P} - \widehat{\pi}_1 \cdot \overline{X} & \text{(intercept of the reduced price eq)} \\
\widehat{\pi}_{22} = \dfrac{\sum q_t x_t}{\sum x_t^2} & \text{(slope of the reduced quantaty eq)} \\
\widehat{\pi}_{12} = \overline{Q} - \widehat{\pi}_3 \cdot \overline{X} & \text{(intercept of the reduced quantaty eq)}
\end{cases}
$$

Case demo: structural coefficients

Because we already know that **the supply equation** in the structural SEM is **Just identification** (please review the order and rank conditions) , hence the structural coefficients of the supply equation can be calculated uniquely with the reduced coefficients.

$$\begin{cases} \beta_0 = \pi_{12} + \beta_1 \pi_{11} \\ \beta_1 = \dfrac{\pi_{22}}{\pi_{21}} \end{cases}$$

which is:

$$\begin{cases} \hat{\beta}_0 = \widehat{\pi}_{12} + \hat{\beta}_1 \widehat{\pi}_{11} \\ \hat{\beta}_1 = \dfrac{\widehat{\pi}_{22}}{\widehat{\pi}_{21}} \end{cases}$$

Case demo: OLS estimates for reduced equation

Next, we carry out OLS regression for the reduced equation.

$$\begin{cases} P_t = \pi_{11} + \pi_{21} X_t + w_t & \text{(reduced eq1)} \\ Q_t = \pi_{12} + \pi_{22} X_t + v_t & \text{(reduced eq2)} \end{cases}$$

The regression result of the reduced price equation is:

$$\begin{array}{lll} \widehat{P} = & +90.96 & +0.00X \\ \text{(t)} & (22.4499) & (3.0060) \\ \text{(se)} & (4.0517) & (0.0002) \\ \text{(fitness)} & R^2 = 0.2440; \bar{R}^2 = 0.2170 & \\ & F^* = 9.04; \quad p = 0.0055 & \end{array}$$

The regression result of the reduced quantity equation is:

$$\begin{array}{lll} \widehat{Q} = & +59.76 & +0.00X \\ \text{(t)} & (38.3080) & (20.9273) \\ \text{(se)} & (1.5600) & (0.0001) \\ \text{(fitness)} & R^2 = 0.9399; \bar{R}^2 = 0.9378 & \\ & F^* = 437.95; p = 0.0000 & \end{array}$$

Case demo: obtain structural coefficients

we can obtain the **reduced coefficients**:

- $\widehat{\pi}_{21} = 0.00074$
- $\widehat{\pi}_{11} = 90.96007$
- $\widehat{\pi}_{22} = 0.00197$
- $\widehat{\pi}_{12} = 59.76183$

Because **supply equation** in structural SEM is **Just identification**, so the structural coefficients of **supply equation** can be calculated by using the estimated reduced coefficients.

$$\hat{\beta}_1 = \frac{\widehat{\pi}_{22}}{\widehat{\pi}_{12}} = 0.00197 / 0.00074 = 2.68052$$

$$\hat{\beta}_0 = \widehat{\pi}_{12} + \hat{\beta}_1 \widehat{\pi}_{11} = 59.76183 - 2.68052 \cdot 90.96007 = -184.05874$$

Therefore, the ILS estimators of **supply equation** parameters are:

$$\hat{Q}_t = -184.05874 + 2.68052\ P_t$$

Case demo: result comparison

As comparison, we will show a **"biased"** estimation method, which use OLS directly for both quantity and price equation.

- Estimation of the **supply equation** based on the ILS approach:

$$\hat{Q}_t = \text{-}184.05874 + 2.68052\ P_t$$

- Estimation of the **supply equation** based on the **biased** OLS approach:

$$\begin{array}{llll} \widehat{Q} = & +20.89 & +0.67P \\ \text{(t)} & (0.9067) & (2.9940) \\ \text{(se)} & (23.0396) & (0.2246) \\ \text{(fitness)} & R^2 = 0.2425; & \bar{R}^2 = 0.2154 \\ & F^* = 8.96; & p = 0.0057 \end{array}$$

20.4 Two-stage least square method (2SLS)

Overidentification: structural SEM

Consider the following structural SEM:

$$
\begin{cases}
Y_{t1} = \quad + \gamma_{21} Y_{2t} + \beta_{01} + \beta_{11} X_{t1} + \beta_{21} X_{t2} + u_{t1} & \text{(income eq)} \\
Y_{t2} = + \gamma_{12} Y_{t1} \quad + \beta_{02} \quad + u_{t2} & \text{(monetary supply eq)}
\end{cases}
$$

where: Y_1 =Income; Y_2 =Monetary stock; X_1 =Government expenditure; X_2 = Government spending on goods and services

Using **order condition rules** and **rank condition rules** (Review), we can know:

- The **Income equation** is **underidentification**
- if you don't change the model specification, then god can't help you!
- The **monetary supply equation** is **overidentification** it's easy to prove that if we apply the ILS approach we will obtain two estimates on γ_{21}. Hence it is impossible to determin the exact value.

Overidentification: Instrument variables

Looking for **Instrument variables approach** to crack the **overidentification** problems:

- In practice, people might want to use OLS to estimate the monetary supply equation, but it will get the biased estimators, because there exist correlationship between Y_1 and u_2.

> **Instrument Variable**: An agent variable which is highly correlated with Y_1 but have no relationship with u_2.

- if we can find an **instrument variable**, then we can apply OLS approach directly to estimate the structural monetary supply eqution.

But how does one obtain such an instrumental variable?

One answer is provided by the **two-stageleast squares** (2SLS), developed independently by Henri Theil and Robert Basmann.

2SLS method involves two successive applications of OLS. The process is as follows:

Stage 1. To get rid of the likely correlation between Y_1 and u_2,apply regresssion Y_1 on all the predetermined variables in the whole system, not just that equation.

$$\begin{aligned} Y_{t1} &= \widehat{\pi}_{01} + \widehat{\pi}_{11} X_{t1} + \widehat{\pi}_{21} X_{t2} + \hat{v}_{t1} \\ &= \hat{Y_{t1}} + \hat{v}_{t1} \\ \hat{Y_{t1}} &= \widehat{\pi}_{01} + \widehat{\pi}_{11} X_{t1} + \widehat{\pi}_{21} X_{t2} \end{aligned}$$

Indicates that the random Y_1 is composed of two parts:

- a linear combination of the nonstochastic X
- random component $\hat{u}_t$

> according to OLS theory, $\hat{Y}_{t1}$ is not related to $\hat{v}_{t1}$ (Why?).

Overidentification: stage 2 of 2SLS

stage 2. Now retransform the **overidentification** supply equation as follow:

$$\begin{aligned}
Y_{t2} &= \beta_{02} + \gamma_{12} Y_{t1} + u_{t2} \\
&= \beta_{02} + \gamma_{12} (\hat{Y}_{t1} + \hat{v}_{t1}) + u_{t2} \\
&= \beta_{02} + \gamma_{12} \hat{Y}_{t1} + (\gamma_{12} \hat{v}_{t1} + u_{t2}) \\
&= \beta_{02} + \gamma_{12} \hat{Y}_{t1} + u^*_{t2}
\end{aligned}$$

We can prove that:

- the variable Y_{t1} may be relative with the disturbance term u_{t2}, which will invalid the OLS approach.
- Meanwhile, $\hat{Y}_{1t}$is uncorrelated with u^*_{t2} asymptotically, that is, in the large sample (or more accurately, as the sample size increases indefinitely).

> As a result, OLS can be applied to monetary Eq, which will give consistent estimates of the parameters of the monetary supply function.

2SLS approach: Features

Note the following features of 2SLS:

- It can be applied to an individual equation in the system without directly taking into account any other equation(s) in the system. Hence, for solving econometric models involving a large number of equations, 2SLS offers an economical method.
- Unlike ILS, which provides multiple estimates of parameters in the overidentified equations, 2SLS provides only one estimate per parameter.
- It is easy to apply because all one needs to know is the total number of exogenous or pre-determined variables in the system without knowing any other variables in the system.
- Although specially designed to handle overidentified equations, the method can also be applied to exactly identified equations. But then ILS and 2SLS will give identical estimates. (Why?)

Note the following features of 2SLS (continue):

- If the R^2 values in the reduced-form regressions (that is, Stage 1 regressions) are very high, say, in excess of 0.8, the classical OLS estimates and 2SLS estimates will be very close.

> But this result should not be surprising because if the R^2 value in the first stage is very high, it means that the estimated values of the endogenous variables are very close to their actual values.
> And hence the latter are less likely to be correlated with the stochastic disturbances in the original structural SEM. (Why?)

2SLS approach: Features

Note the following features of 2SLS (continue):

- Notice that in reporting the ILS regression we did not state the **standard errors** of the estimated coefficients . But we can do this for the 2SLS estimates because the structural coefficients are directly estimated from the second-stage (OLS) regressions.

  > The estimated standard errors in the second-stage regressions need to be modified because the error term u_t^* is, in fact, equal to
  > $u_{2t} + \beta_{21}\hat{u}_t$.

> Hence, the variance of u_t^* is not exactly equal to the variance of the original u_{2t}.

2SLS approach: Features

Note the following features of 2SLS (continue):

- Remarks from `Henri Theil`:

> - The statistical justification of the 2SLS is of the large-sample type.
> - When the equation system contains lagged endogenous variables, the consistency and large-sample normality of the 2SLS coefficient estimators require an additional condition.
> - Take cautions when lagged endogenous variables are not really predetermined.

Standard error correction: why

In the regression report of ILS method, we do not give the **standard error** of the estimated coefficient, but we can give these **standard error** for the estimator of 2SLS.

- Remind $u_{t2}^* = u_{t2} + \gamma_{12}\hat{v}_{t1}$
- It will imply $u_{t2}^* \neq u_{t2}$, and then we need to calculate the "correct" standard error for the purpose of inference.

> For the specific method of error correction, please refer to appendix 20a.2 of the textbook (Damodar Gujarati).

- In the following cases illurtration, we will show the 2SLS estimates **without error correction** and the 2SLS estimates **with error correction** respectively.

Standard error correction: focus stage 2

The process for error correction show as below.

- **stage 2:** The regression form of the supply equation is

$$
\begin{aligned}
Y_{t2} &= \beta_{02} + \gamma_{12} Y_{1t} + u_{2t} \\
&= \beta_{02} + \gamma_{12}(\hat{Y_{1t}} + \hat{v}_{t1}) + u_{2t} \\
&= \beta_{02} + \gamma_{12}\hat{Y_{1t}} + (\gamma_{12}\hat{v}_{t1} + u_{t2}) \\
&= \beta_{02} + \gamma_{12}\hat{Y_{1t}} + u^*_{t2}
\end{aligned}
$$

> Where:
>
> $u^*_{t2} = u_{t2} + \gamma_{12}\hat{v}_{t1}$

Standard error correction: focus stage 2

- **stage 2:** the estimation for the parameter γ_{12} is $\hat{\gamma}_{12}$, and its standard error $s.e_{\hat{\gamma}_{12}}$ can be calculated as below.

$$Y_{t2} = \beta_{02} + \gamma_{12}\hat{Y_{1t}} + u^*_{t2}$$

$$s.e_{\hat{\gamma}_{21}} = \frac{\hat{\sigma}^2_{u^*_{t2}}}{\sum \hat{y}^2_{t1}}$$

$$\hat{\sigma}^2_{u^*_{t2}} = \frac{\sum \left(u^*_{t2}\right)^2}{n-2} = \frac{\sum \left(Y_{t2} - \hat{\beta}_{02} - \hat{\gamma}_{12}\hat{Y}_{1t}\right)^2}{n-2}$$

Standard error correction: results

- In fact, we know $u_{t2}^{*} \neq u_{t2}$, which means $\hat{\sigma}_{u_{t2}^{*}} \neq \hat{\sigma}_{u_{t2}}$.
- Thus we can obtain $\hat{\sigma}_{u_{t2}}$.

$$\hat{u}_{t2} = Y_{t2} - \hat{\beta}_{02} - \hat{\gamma}_{12} Y_{t1}$$

$$\hat{\sigma}_{u_{2t}}^{2} = \frac{\sum (u_{2t})^{2}}{n-2} = \frac{\sum \left(Y_{t2} - \hat{\beta}_{02} - \hat{\gamma}_{12} Y_{1t}\right)^{2}}{n-2}$$

Standard error correction: for coefficients

Therefore, in order to correct the standard error of the coefficients estimated by **stage 2** regression, it is necessary to multiply the standard error of each coefficient by the following **error correction factor**.

$$\eta = \frac{\hat{\sigma}^2_{u_{t2}}}{\hat{\sigma}^2_{u^*_{t2}}}$$

$$s.e^*_{\hat{\gamma}_{12}} = s.e_{\hat{\gamma}_{12}} \cdot \eta = s.e_{\hat{\gamma}_{12}} \cdot \frac{\hat{\sigma}^2_{u_{t2}}}{\hat{\sigma}^2_{u^*_{t2}}}$$

$$s.e^*_{\hat{\beta}_{02}} = s.e_{\hat{\beta}_{02}} \cdot \eta = s.e_{\hat{\beta}_{20}} \cdot \frac{\hat{\sigma}^2_{u_{t2}}}{\hat{\sigma}^2_{u^*_{t2}}}$$

Case study and application for 2SLS approach

variable description

Variables description

vars	label	note
Y1	GDP: gross domestic product	($1 billion in 2000)
Y2	M2:money supply	($1 billion)
X1	GDPI: Total private domestic investment	($1 billion in 2000)
X2	FEDEXP: Federal expenditure	($1 billion)
Y1.l1	GDP_t-1: The gross domestic product of the previous period	($1 billion in 2000)
Y2.l1	M2_t-1: The money supply in the previous period	($1 billion)
X3	TB6: 6 month Treasury bond interest rate	(%)

data set

Sample data (n=36)

year	Y1	Y2	X1	X2	Y1.l1	Y2.l1
1970	3771.9	626.5	427.1	201.1		
1971	3898.6	710.3	475.7	220	3771.9	626.5
1972	4105	802.3	532.1	244.4	3898.6	710.3
1973	4341.5	855.5	594.4	261.7	4105	802.3
1974	4319.6	902.1	550.6	293.3	4341.5	855.5
1975	4311.2	1016.2	453.1	346.2	4319.6	902.1
1976	4540.9	1152	544.7	374.3	4311.2	1016.2
1977	4750.5	1270.3	627	407.5	4540.9	1152

Showing 1 to 8 of 36 entries

Previous 1 2 3 4 5 Next

Modeling scenario 1

Only the money supply equation is overidentifiable

Structural SEM and identification problems

Therefore, we can construct the following structural SEM:

$$\begin{cases} Y_{t1} = \beta_{01} \quad + \gamma_{21} Y_{t2} + \beta_{11} X_{t1} + \beta_{21} X_{t2} + u_{t1} & \text{(income eq)} \\ Y_{t2} = \beta_{02} + \gamma_{12} Y_{1t} \quad + u_{t2} & \text{(money supply eq)} \end{cases}$$

Where:

- $Y_1 = GDP$ (gross domestic product GDP);
- $Y_2 = M2$ (money supply);
- $X_1 = GDPI$ (Private domestic investment);
- $X_2 = FEDEXP$ (Federal expenditure)

2SLS approach 1: without error correction

$$\begin{cases} Y_{t1} = \beta_{01} \quad + \gamma_{21} Y_{t2} + \beta_{11} X_{t1} + \beta_{21} X_{t2} + u_{t1} & \text{(income eq)} \\ Y_{t2} = \beta_{02} + \gamma_{12} Y_{1t} \quad + u_{t2} & \text{(money supply eq)} \end{cases}$$

stage 1: Estimate the regression of Y_1 to all **predetermined variables** in the structural SEM (not only in the equation under consideration), and obtain $\hat{Y}_{t1}; \hat{n}_{t1}$.

That is:

$$\begin{aligned} Y_{1t} &= \widehat{\pi_0} + \widehat{\pi_1} X_{1t} + \widehat{\pi_2} X_{2t} + \hat{v}_{t1} \\ &= \hat{Y}_{1t} + \hat{v}_{t1} \end{aligned}$$

Regression results of **stage 1**:

$$\begin{array}{llll} \widehat{Y1} = & +2689.85 & +1.87X1 & +2.03X2 \\ \text{(t)} & (39.5639) & (10.8938) & (18.9295) \\ \text{(se)} & (67.9874) & (0.1717) & (0.1075) \\ \text{(fitness)} R^2 = 0.9964; & \bar{R}^2 = 0.9962 & & \\ & F^* = 4534.36; p = 0.0000 & & \end{array}$$

2SLS approach 1: without error correction

At the same time, we can obtain $\hat{Y}_{t1}; \hat{v}_{t1}$:

New variable Yt1.hat and vt1.hat after regression of stage I

year	Y1	Y2	X1	X2	Y1.l1	Y2.l1	Y1.hat	v1.hat
1970	3771.9	626.5	427.1	201.1			3897.6136	-125.7136
1971	3898.6	710.3	475.7	220	3771.9	626.5	4026.9427	-128.3427
1972	4105	802.3	532.1	244.4	3898.6	710.3	4182.0464	-77.0464
1973	4341.5	855.5	594.4	261.7	4105	802.3	4333.7391	7.7609
1974	4319.6	902.1	550.6	293.3	4341.5	855.5	4316.1193	3.4807
1975	4311.2	1016.2	453.1	346.2	4319.6	902.1	4241.4135	69.7865

Showing 1 to 6 of 36 entries

Previous 1 2 3 4 5 6 Next

2SLS approach 1: without error correction

$$\begin{cases} Y_{t1} = \beta_{01} \quad + \gamma_{21} Y_{t2} + \beta_{11} X_{t1} + \beta_{21} X_{t2} + u_{t1} & \text{(income eq)} \\ Y_{t2} = \beta_{02} + \gamma_{12} Y_{1t} \quad + u_{t2} & \text{(money supply eq)} \end{cases}$$

stage 2: Now transform the **overidentification** supply equation as follows:

$$Y_{t2} = \beta_{02} + \gamma_{12} \hat{Y}_{1t} + u^*_{t2}$$

> Using the new variables in **stage 1** results, and apply OLS estimation to obtain:

$$\begin{aligned}
\widehat{Y2} = &\quad -2440.20 \quad + 0.79 Y1.hat \\
\text{(t)} &\quad (-19.1578) \quad (44.5241) \\
\text{(se)} &\quad (127.3738) \quad (0.0178) \\
\text{(fitness)} &\, R^2 = 0.9831; \;\; \bar{R}^2 = 0.9826 \\
&\, F^* = 1982.40; p = 0.0000
\end{aligned}$$

Comparison 1: OLS approach with biased estimation

As comparison, we give a **"biased"** estimation with OLS method directly to the money supply equation, and obtain following result:

$$
\begin{array}{llll}
\widehat{Y2} = & -2430.34 & +0.79Y1 \\
(\text{t}) & (-19.1042) & (44.5059) \\
(\text{se}) & (127.2148) & (0.0178) \\
(\text{fitness}) & R^2 = 0.9831; & \bar{R}^2 = 0.9826 \\
 & F^* = 1980.77; p = 0.0000 &
\end{array}
$$

Comparison 1: OLS approach with biased estimation

The R raw report for the biased regression show as:

```
Call:
lm(formula = models_money$mod.ols, data = us_money_new)

Residuals:
   Min     1Q Median     3Q    Max 
-418.3 -151.6   40.2  143.5  380.8 

Coefficients:
             Estimate Std. Error t value Pr(>|t|)    
(Intercept) -2.43e+03   1.27e+02   -19.1   <2e-16 ***
Y1           7.90e-01   1.78e-02    44.5   <2e-16 ***
---
Signif. codes:  
0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 229 on 34 degrees of freedom
Multiple R-squared:  0.983,	Adjusted R-squared:  0.983 
F-statistic: 1.98e+03 on 1 and 34 DF,  p-value: <2e-16
```

Modeling scenario 2

both income equation and money supply equation are over-identifiable

Improved structural SEM and identification problems

Different from the former structural SEM, we can construct the improved one:

$$\begin{cases} Y_{t1} = \beta_{01} \qquad + \gamma_{12} Y_{t2} + \beta_{11} X_{t1} + \beta_{21} X_{t2} \qquad + u_{t1} & \text{(income eq)} \\ Y_{t2} = \beta_{02} + \gamma_{12} Y_{1t} \qquad + \beta_{12} Y_{1,t-1} + \beta_{22} Y_{2,t-1} + u_{t2} & \text{(money supply eq)} \end{cases}$$

Next, we judge the identification problem according to **order condition rules 2** :

- The number of predetermineed variables in the structural SEM is $K = 5$.
- The first equation: the number of predetermined variables is $k = 3$, and $(K - k) = 2$. Also the number of endogenous variables is $m = 2$, and $(m - 1) = 1$. We will see $(K - k) > (m - 1)$. So the first equation is **overidentification**.
- The second equation: the number of predetermined variables is $k = 3$, and $(K - k) = 2$; Also the number of endogenous variables is $m = 2$, and $(m - 1) = 1$. We will see $(K - k) > (m - 1)$. Thus it is also **overidentification**.

2SLS approach 2: without error correction

Now, we will use two-stage least squares (2SLS) to get **consistent estimates** for both **income equation** and **money supply equation**.

Stage 1:

- Estimate the regression of Y_1 to all **predetermined variables** in the structural SEM (not only in the equation under consideration), and obtain $\hat{Y}_{t1}; \hat{v}^*_{t1}$.
- Meanwhile, estimate the regression of Y_2 to all **predetermined variables** in the structural SEM (not only in the equation under consideration), and obtain $\hat{Y}_{t2}; \hat{v}^*_{t2}$:

$$
\begin{aligned}
Y_{t1} &= \widehat{\pi}_{01} + \widehat{\pi}_{11} X_{t1} + \widehat{\pi}_{21} X_{t2} + \widehat{\pi}_{31} Y_{t-1,1} + \widehat{\pi}_{41} Y_{t-1,2} + \hat{v}_{t1} \\
&= \hat{Y}_{1t} + \hat{v}_{t1} \\
Y_{t2} &= \widehat{\pi}_{02} + \widehat{\pi}_{12} X_{t1} + \widehat{\pi}_{22} X_{t2} + \widehat{\pi}_{32} Y_{t-1,1} + \widehat{\pi}_{42} Y_{t-1,2} + \hat{v}_{t2} \\
&= \hat{Y}_{t2} + \hat{v}_{t2}
\end{aligned}
$$

2SLS approach 2: without error correction

- OLS regression results of **new income equation** at **stage 1**:

$$
\begin{array}{llllll}
\widehat{Y1} = & +1098.90 & +0.98X1 & +0.77X2 & +0.59Y1.l1 & -0.01Y2.l1 \\
\text{(t)} & (5.9222) & (7.4954) & (4.1895) & (8.8729) & (-0.1024) \\
\text{(se)} & (185.5566) & (0.1308) & (0.1831) & (0.0667) & (0.0721) \\
\text{(fitness)} & R^2 = 0.9990; & \bar{R}^2 = 0.9989 & & & \\
 & F^* = 7857.58; & p = 0.0000 & & &
\end{array}
$$

2SLS approach 2: without error correction

- OLS regression results of **new money supply equation** at **stage 1**:

$$
\begin{aligned}
\widehat{Y2} = &\; -207.14 && +0.20X1 && -0.35X2 && +0.06Y1.l1 && +1.06Y2.l1 \\
(t) &\; (-1.1202) && (1.5298) && (-1.9455) && (0.9352) && (14.7779) \\
(se) &\; (184.9121) && (0.1303) && (0.1824) && (0.0665) && (0.0718) \\
(fitness) &\; R^2 = 0.9985; \; \bar{R}^2 = 0.9983 \\
&\; F^* = 5050.57; p = 0.0000
\end{aligned}
$$

2SLS approach 2: without error correction

Hence, we can obtain new variables from the two former regressions results respectively: $\hat{Y}_{t1}; \hat{v}_{t1}$, and $\hat{Y}_{t2}; \hat{v}_{t2}$.

New variables from regression of stage I

Y1	Y1.l1	Y1.hat	v1.hat	Y2	Y2.l1	Y2.hat	v2.hat
3,771.9				626.5			
3,898.6	3,771.9	3,963.2	-64.5622	710.3	626.5	709.5	0.8424
4,105.0	3,898.6	4,111.6	-6.5777	802.3	710.3	808.9	-6.6001
4,341.5	4,105.0	4,307.5	34.0275	855.5	802.3	925.7	-70.2030
4,319.6	4,341.5	4,428.4	-108.8458	902.1	855.5	977.0	-74.8543
4,311.2	4,319.6	4,360.1	-48.9302	1,016.2	902.1	986.9	29.3416

Showing 1 to 6 of 36 entries

Previous 1 2 3 4 5 6 Next

Stage 2:

- The re-transformed new income equation and the money supply equation are:

$$Y_{t1} = \beta_{01} + \gamma_{21}\hat{Y}_{t2} + \beta_{11}X_{t1} + \beta_{21}X_{t2} + u_{t1}^*$$
$$Y_{t2} = \beta_{20} + \gamma_{12}\hat{Y}_{1t} + \beta_{12}Y_{t-1,1} + \beta_{22}Y_{t-1,2} + u_{t2}^*$$

- And then, we can conduct these new equations by using the former variables frome stage 1.

2SLS approach 2: without error correction

- OLS estimation results of **new** income equation in **stage 2**:

$$
\begin{array}{lllll}
\widehat{Y1} = & +2723.68 & +0.22Y2.hat & +1.71X1 & +1.57X2 \\
(\text{t}) & (40.3310) & (1.8961) & (9.2748) & (5.9811) \\
(\text{se}) & (67.5331) & (0.1156) & (0.1848) & (0.2623) \\
(\text{fitness}) & R^2 = 0.9966; & \bar{R}^2 = 0.9963 & & \\
 & F^* = 3073.97; p = 0.0000 & & &
\end{array}
$$

2SLS approach 2: without error correction

- OLS estimation results of the **new** money supply equation in **stage 2**:

$$
\begin{array}{lllll}
\widehat{Y2} = & -228.13 & +0.11Y1.hat & -0.03Y1.l1 & +0.93Y2.l1 \\
\text{(t)} & (-1.4843) & (0.7685) & (-0.1691) & (15.0961) \\
\text{(se)} & (153.6925) & (0.1431) & (0.1481) & (0.0618) \\
\text{(fitness)} & R^2 = 0.9981; & \bar{R}^2 = 0.9979 & & \\
 & F^* = 5461.60; p = 0.0000 & & &
\end{array}
$$

2SLS approach 2: without error correction

- The R raw report of OLS estimation for the new income equation in **stage 2**:

```
Call:
lm(formula = models_money2$mod2.stage2.1, data = us_money_new2)

Residuals:
   Min      1Q Median      3Q     Max 
-360.4  -66.8   35.7   81.2  186.2 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept) 2723.681     67.533   40.33  < 2e-16 ***
Y2.hat         0.219      0.116    1.90    0.067 .  
X1             1.714      0.185    9.27  1.9e-10 ***
X2             1.569      0.262    5.98  1.3e-06 ***
---
Signif. codes:  
0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 130 on 31 degrees of freedom
  (因为不存在，1个观察量被删除了)
Multiple R-squared:  0.997,    Adjusted R-squared:  0.996 
F-statistic: 3.07e+03 on 3 and 31 DF,  p-value: <2e-16
```

2SLS approach 2: without error correction

- The R raw report of OLS estimation for the new money supply equation in **stage 2**:

```
Call:
lm(formula = models_money2$mod2.stage2.2, data = us_money_new2)

Residuals:
    Min      1Q  Median      3Q     Max
-192.58  -42.96    7.05   42.93  218.68

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept) -228.1320   153.6925   -1.48     0.15
Y1.hat         0.1100     0.1431    0.77     0.45
Y1.l1         -0.0250     0.1481   -0.17     0.87
Y2.l1          0.9330     0.0618   15.10  7.8e-16 ***
---
Signif. codes:
0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 78 on 31 degrees of freedom
  (因为不存在，1个观察量被删除了)
Multiple R-squared:  0.998,    Adjusted R-squared:  0.998
F-statistic: 5.46e+03 on 3 and 31 DF,  p-value: <2e-16
```

2SLS approach 2: with error correction

By using R `systemfit` package, we can apply the two-stage least square method with **"error correction"**, and the report summarized as follows:

Result of 2SLS with error correction

eq	vars	Estimate	Std. Error	t value	Pr(>|t|)
eq1	(Intercept)	2,723.68	69.1017	39.4155	0.0000
eq1	Y2	0.22	0.1183	1.8530	0.0734
eq1	X1	1.71	0.1891	9.0642	0.0000
eq1	X2	1.57	0.2684	5.8453	0.0000
eq2	(Intercept)	-228.13	157.9455	-1.4444	0.1587
eq2	Y1	0.11	0.1471	0.7478	0.4602
eq2	Y1.11	-0.03	0.1522	-0.1645	0.8704
eq2	Y2.11	0.93	0.0635	14.6896	0.0000

2SLS approach 2: with error correction

By using R `systemfit` package, we can apply the two-stage least square method with **"error correction"**, and the detail report show as follows:

```
systemfit results
method: 2SLS

        N DF     SSR  detRCov OLS-R2 McElroy-R2
system 70 62 749260 1.07e+08  0.997      0.998

     N DF    SSR   MSE  RMSE    R2 Adj R2
eq1 35 31 549669 17731 133.2 0.996  0.996
eq2 35 31 199592  6438  80.2 0.998  0.998

The covariance matrix of the residuals
      eq1   eq2
eq1 17731 -2604
eq2 -2604  6438

The correlations of the residuals
       eq1    eq2
eq1  1.000 -0.244
eq2 -0.244  1.000
```

Comparison 2: OLS approach with biased estimation

- **"biased"** OLS estimation results of the income equation:

$$
\begin{aligned}
\widehat{Y1} = & \quad +2706.39 & +0.17Y2 & \quad +1.75X1 + 1.68X2 \\
(\text{t}) & \quad (40.0265) & (1.5062) & \quad (9.3862) \quad (6.6008) \\
(\text{se}) & \quad (67.6150) & (0.1115) & \quad (0.1864) \quad (0.2552) \\
(\text{fitness}) & R^2 = 0.9966; \quad \bar{R}^2 = 0.9963 \\
& F^* = 3139.86; p = 0.0000
\end{aligned}
$$

- **"biased"** OLS estimates of the money supply equation:

$$
\begin{aligned}
\widehat{Y2} = & \quad -206.53 & -0.02Y1 & \quad +0.10Y1.l1 + 0.94Y2.l1 \\
(\text{t}) & \quad (-1.3373) & (-0.1377) & \quad (0.7870) \quad (15.1363) \\
(\text{se}) & \quad (154.4428) & (0.1178) & \quad (0.1252) \quad (0.0622) \\
(\text{fitness}) & R^2 = 0.9981; \quad \bar{R}^2 = 0.9979 \\
& F^* = 5362.58; p = 0.0000
\end{aligned}
$$

20.5 Truffle supply and demand

Variables description

variables description

vars	label	measure
P	market price of truffles	dollar/ounce
Q	market quantity of truffles	ounce
PS	market price of substitute	dollar/ounce
DI	disposable income	dollar/person, monthly
PF	rental price of truffle-pigs	dollar/hour

Data set

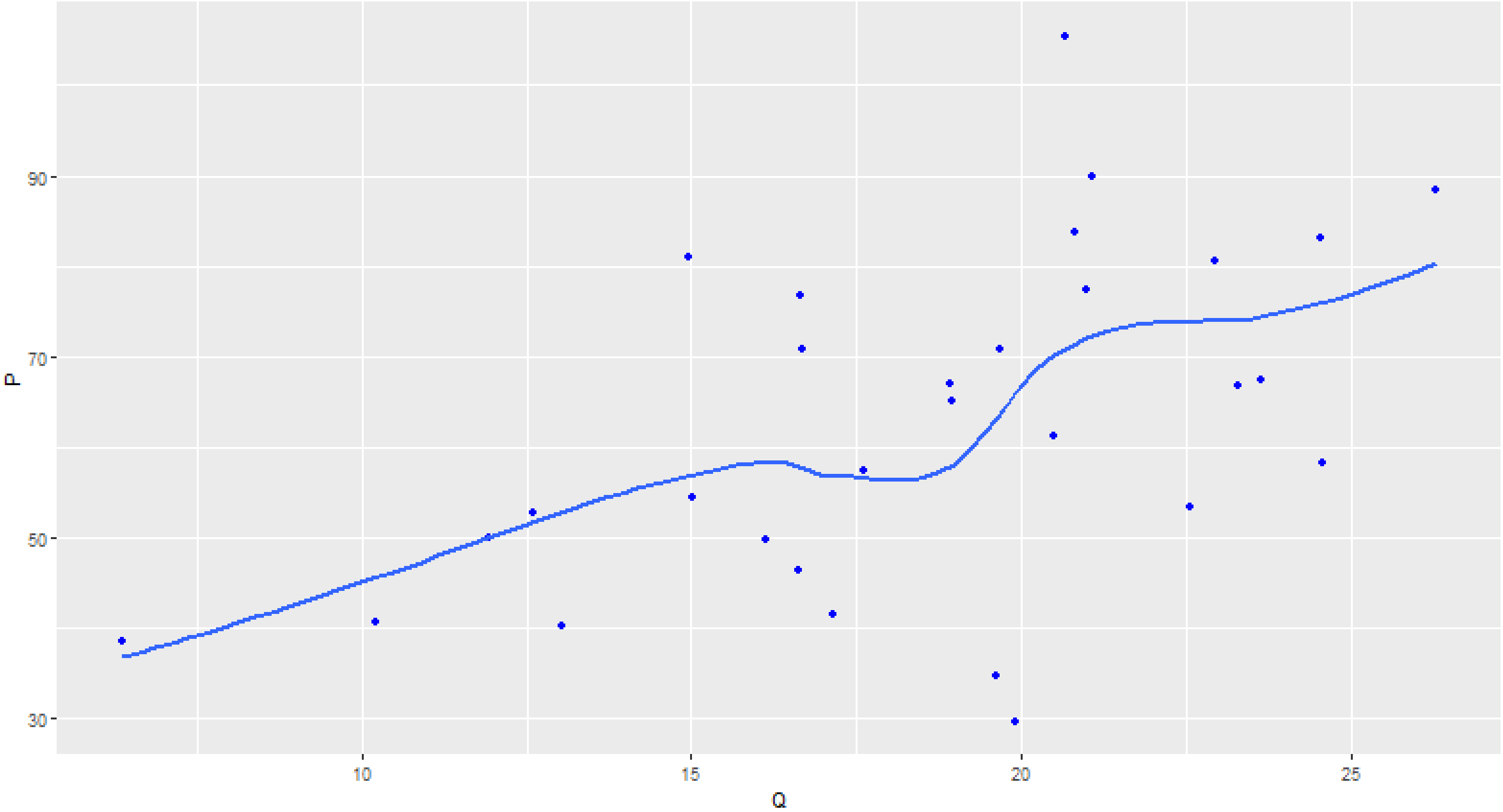
Case data set (n=30)

P	Q	PS	DI	PF
29.64	19.89	19.97	2.103	10.52
40.23	13.04	18.04	2.043	19.67
34.71	19.61	22.36	1.87	13.74
41.43	17.13	20.87	1.525	17.95
53.37	22.55	19.79	2.709	13.71
38.52	6.37	15.98	2.489	24.95
54.33	15.02	17.94	2.294	24.17
40.56	10.22	17.09	2.196	23.61

Showing 1 to 8 of 30 entries

Previous 1 2 3 4 Next

The scatter plot on truffle quantity Q and truffle market price P is given below:



The structural SEM

Given the structural SEM:

$$
\begin{cases}
Q_i = \alpha_0 + \alpha_1 P_i + \alpha_2 PS_i + \alpha_3 DI_i + u_{i1} & \text{(demand function)} \\
Q_i = \beta_0 + \beta_1 P_i + \beta_2 PF_i + u_{i2} & \text{(supply function)}
\end{cases}
$$

We can get the reduced SEM:

$$\begin{cases} P_i = \pi_{01} + \pi_{11}PS_i + \pi_{21}DI_i + \pi_{31}PF_i + v_{t1} \\ Q_i = \pi_{02} + \pi_{12}PS_t + \pi_{22}DI_i + \pi_{32}PF_i + v_{t2} \end{cases}$$

Also we obtain the relationship between structural and reduced coefficients:

$$\begin{cases} \pi_{01} = \dfrac{\beta_0 - \alpha_0}{\alpha_1 - \beta_1} \\ \pi_{11} = -\dfrac{\alpha_2}{\alpha_1 - \beta_1} \\ \pi_{21} = -\dfrac{\alpha_3}{\alpha_1 - \beta_1} \\ \pi_{31} = \dfrac{\beta_2}{\alpha_1 - \beta_1} \\ v_{t1} = \dfrac{u_{2t} - u_{1t}}{\alpha_1 - \beta_1} \end{cases}$$

$$\begin{cases} \pi_{02} = -\dfrac{\alpha_1\beta_0 - \alpha_0\beta_1}{\alpha_1 - \beta_1} \\ \pi_{12} = -\dfrac{\alpha_2\beta_1}{\alpha_1 - \beta_1} \\ \pi_{22} = -\dfrac{\alpha_3\beta_1}{\alpha_1 - \beta_1} \\ \pi_{32} = \dfrac{\alpha_1\beta_2}{\alpha_1 - \beta_1} \\ v_{t2} = \dfrac{\alpha_1 u_{2t} - \beta_1 u_{1t}}{\alpha_1 - \beta_1} \end{cases}$$

Simple OLS solution: results

We can apply OLS method directly. Of course estimation results will be biased.

- tidy results of **bias** OLS estimation for the demand equation:

$$
\begin{aligned}
\widehat{Q} = &\quad +1.09 &\quad +0.02P &\quad +0.71PS + 0.08DI \\
\text{(t)} &\quad (0.2940) &\quad (0.3032) &\quad (3.3129) \quad (0.0642) \\
\text{(se)} &\quad (3.7116) &\quad (0.0768) &\quad (0.2143) \quad (1.1909) \\
\text{(fitness)} & R^2 = 0.4957; \bar{R}^2 = 0.4375 \\
& F^* = 8.52; \quad p = 0.0004
\end{aligned}
$$

- tidy results of **bias** OLS estimation for the supply equation:

$$
\begin{aligned}
\widehat{Q} = &\quad +20.03 &\quad +0.34P &\quad -1.00PF \\
\text{(t)} &\quad (16.3938) &\quad (15.5436) &\quad (-13.1028) \\
\text{(se)} &\quad (1.2220) &\quad (0.0217) &\quad (0.0764) \\
\text{(fitness)} & R^2 = 0.9019; \bar{R}^2 = 0.8946 \\
& F^* = 124.08; p = 0.0000
\end{aligned}
$$

Simple OLS solution: R code (`lm`)

```r
# set equation systems
eq.D <- Q~P+PS+DI
eq.S <- Q~P+PF

# fit using direct `OLS` method
ols.D <- lm(formula = eq.D, data = truffles)
ols.S <- lm(formula = eq.S, data = truffles)
# report
smry.olsD <- summary(ols.D)
smry.olsS <- summary(ols.S)
```

Simple OLS solution: R report (`lm`)

```
Call:
lm(formula = eq.D, data = truffles)

Residuals:
   Min      1Q  Median      3Q     Max
-7.155 -1.936 -0.374  2.396  6.335

Coefficients:
            Estimate Std. Error t value
(Intercept)   1.0910     3.7116    0.29
P             0.0233     0.0768    0.30
PS            0.7100     0.2143    3.31
DI            0.0764     1.1909    0.06
---
Signif. codes:
0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.

Residual standard error: 3.5 on 26 degr
Multiple R-squared:  0.496,    Adjusted
F-statistic: 8.52 on 3 and 26 DF,  p-va
```

```
Call:
lm(formula = eq.S, data = truffles)

Residuals:
   Min      1Q  Median      3Q     Max
-3.783 -0.853  0.227  0.758  3.347

Coefficients:
            Estimate Std. Error t value
(Intercept)  20.0328     1.2220    16.4
P             0.3380     0.0217    15.5
PF           -1.0009     0.0764   -13.1
---
Signif. codes:
0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.

Residual standard error: 1.5 on 27 degr
Multiple R-squared:  0.902,    Adjusted
F-statistic:  124 on 2 and 27 DF,  p-va
```

Simple OLS solution: R code (`symtemfit`)

Simple OLS solution: R report (`symtemfit`)

```
systemfit results
method: OLS

        N DF SSR detRCov OLS-R2 McElroy-R2
system 60 53 372    23.6  0.699      0.809

     N DF   SSR   MSE RMSE    R2 Adj R2
eq1 30 26 311.2 11.97 3.46 0.496  0.438
eq2 30 27  60.6  2.24 1.50 0.902  0.895

The covariance matrix of the residuals
      eq1  eq2
eq1 11.97 1.81
eq2  1.81 2.24

The correlations of the residuals
      eq1   eq2
eq1 1.000 0.349
eq2 0.349 1.000
```

IV-2SLS Solution: results

IV 2SLS result(using $systemfit::systemfit()$

eq	vars	Estimate	Std. Error	t value	Pr(>|t|)
eq1	(Intercept)	-4.2795	5.5439	-0.7719	0.4471
eq1	P	-0.3745	0.1648	-2.2729	0.0315
eq1	PS	1.2960	0.3552	3.6488	0.0012
eq1	DI	5.0140	2.2836	2.1957	0.0372
eq2	(Intercept)	20.0328	1.2231	16.3785	0.0000
eq2	P	0.3380	0.0249	13.5629	0.0000
eq2	PF	-1.0009	0.0825	-12.1281	0.0000

IV-2SLS Solution: R code (`symtemfit`)

```r
# load pkg
require(systemfit)
# set equation systems
eq.D <- Q~P+PS+DI
eq.S <- Q~P+PF
eq.sys <- list(eq.D, eq.S)
# set instruments
instr <- ~PS+DI+PF
# system fit using `2SLS` method
system.iv <-systemfit(
  formula = eq.sys, inst = instr,
  method="2SLS",
  data=truffles)
# report
smry.iv <- summary(system.iv)
```

IV-2SLS Solution: R report (`symtemfit`)

```
systemfit results
method: 2SLS

        N DF SSR detRCov OLS-R2 McElroy-R2
system 60 53 692    49.8  0.439      0.807

     N DF   SSR   MSE RMSE     R2 Adj R2
eq1 30 26 631.9 24.30 4.93 -0.024 -0.142
eq2 30 27  60.6  2.24 1.50  0.902  0.895

The covariance matrix of the residuals
      eq1  eq2
eq1 24.30 2.17
eq2  2.17 2.24

The correlations of the residuals
      eq1   eq2
eq1 1.000 0.294
eq2 0.294 1.000
```

IV-2SLS Solution: results of the reduced SEM

The OLS regression results of **reduced price equation**:

$$
\begin{array}{lllll}
\widehat{P} = & -32.51 & +1.71PS & +7.60DI & +1.35PF \\
\text{(t)} & (-4.0721) & (4.8682) & (4.4089) & (4.5356) \\
\text{(se)} & (7.9842) & (0.3509) & (1.7243) & (0.2985) \\
\multicolumn{5}{l}{\text{(fitness)} R^2 = 0.8887; \bar{R}^2 = 0.8758} \\
\multicolumn{5}{l}{\qquad F^* = 69.19;\ p = 0.0000}
\end{array}
$$

The OLS regression results of **reduced quantity equation**:

$$
\begin{array}{lllll}
\widehat{Q} = & +7.90 & +0.66PS & +2.17DI & -0.51PF \\
\text{(t)} & (2.4342) & (4.6051) & (3.0938) & (-4.1809) \\
\text{(se)} & (3.2434) & (0.1425) & (0.7005) & (0.1213) \\
\multicolumn{5}{l}{\text{(fitness)} R^2 = 0.6974; \bar{R}^2 = 0.6625} \\
\multicolumn{5}{l}{\qquad F^* = 19.97;\ p = 0.0000}
\end{array}
$$

Comparison: the biased OLS estimation

We can apply OLS method directly. Of course estimation results will be biased.

- tidy results of **bias** OLS estimation for the demand equation:

$$
\begin{array}{lllll}
\widehat{Q} = & +1.09 & +0.02P & +0.71PS & +0.08DI \\
\text{(t)} & (0.2940) & (0.3032) & (3.3129) & (0.0642) \\
\text{(se)} & (3.7116) & (0.0768) & (0.2143) & (1.1909) \\
\text{(fitness)} R^2 = 0.4957; \bar{R}^2 = 0.4375 & & & & \\
F^* = 8.52; \quad p = 0.0004 & & & &
\end{array}
$$

- tidy results of **bias** OLS estimation for the supply equation:

$$
\begin{array}{llll}
\widehat{Q} = & +20.03 & +0.34P & -1.00PF \\
\text{(t)} & (16.3938) & (15.5436) & (-13.1028) \\
\text{(se)} & (1.2220) & (0.0217) & (0.0764) \\
\text{(fitness)} R^2 = 0.9019; \bar{R}^2 = 0.8946 & & & \\
F^* = 124.08; p = 0.0000 & & &
\end{array}
$$

20.6 Cod supply and demand


Moran's Marina

Variables description

Variables description of the cod case

vars	label	note
lprice	the logarithmic market price of cod	continuous variable
lquan	the logarithmic quantity of cod	continuous variable
mon	monday	dummy variable 0/1
tue	tuesday	dummy variable 0/1
wen	wensday	dummy variable 0/1
thu	thursday	dummy variable 0/1
stormy	Stormy	dummy variable 0/1

Showing 1 to 7 of 7 entries

Previous 1 Next

Sample data set

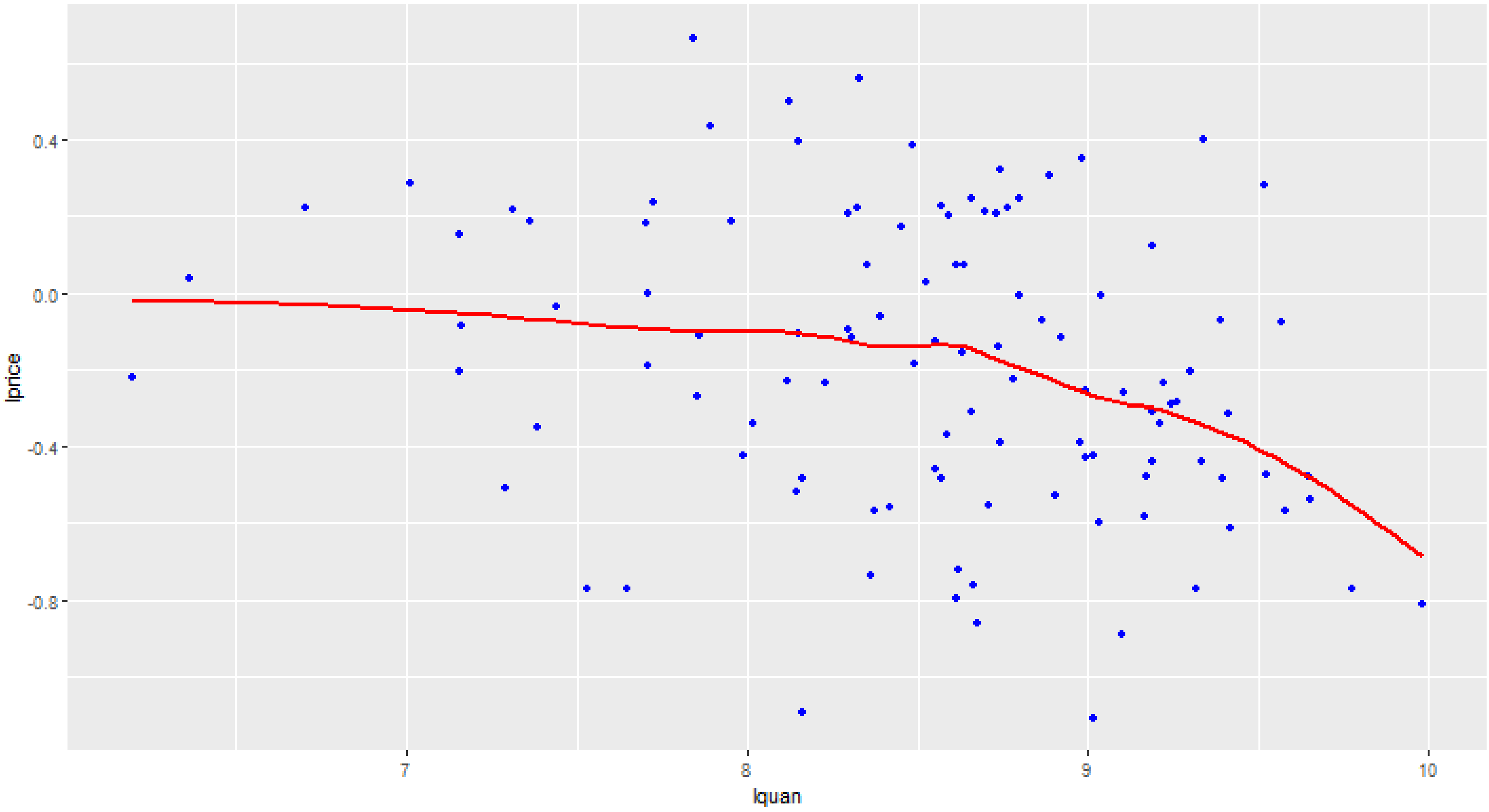
Data set of cod case(obs. n=111)

date	lprice	quan	lquan	mon	tue	wed	thu	stormy
1991-12-02	-0.43	8,058	8.99	1	0	0	0	1
1991-12-03	0.00	2,224	7.71	0	1	0	0	1
1991-12-04	0.07	4,231	8.35	0	0	1	0	0
1991-12-05	0.25	5,750	8.66	0	0	0	1	1
1991-12-06	0.66	2,551	7.84	0	0	0	0	1
1991-12-09	-0.21	10,952	9.30	1	0	0	0	0
1991-12-10	-0.12	7,485	8.92	0	1	0	0	0

Showing 1 to 7 of 111 entries

Previous 1 2 3 4 5 … 16 Next

Scatter



The structural and reduced SEM

Given the structural SEM:

$$
\begin{cases}
lquan_t = \alpha_0 + \alpha_1 lprice_t + \alpha_2 mon_t + \alpha_3 tue_t + \alpha_4 wen_t + \alpha_5 thu_t + u_{1t} & \text{(demand eq)} \\
lquan_t = \beta_0 + \beta_1 lprice_t + \beta_3 stormy_t + u_{2t} & \text{(supply eq)}
\end{cases}
$$

We can obtain the reduced SEM:

$$
\begin{cases}
lquan_t = \pi_0 + \pi_1 mon_t + \pi_2 tue_t + \pi_3 wen_t + \pi_4 thu_t + \pi_5 stormy_t + v_t & \text{(reduced eq1)} \\
lprice_t = \pi_0 + \pi_1 mon_t + \pi_2 tue_t + \pi_3 wen_t + \pi_4 thu_t + \pi_5 stormy_t + v_t & \text{(reduced eq2)}
\end{cases}
$$

The regression results of reduced **quantity** equation show as follows:

$$
\begin{array}{lllllll}
\widehat{lquan} = & +8.81 & +0.10mon & -0.48tue & -0.55wed & +0.05thu & -0.39stormy \\
(\text{t}) & (59.9225) & (0.4891) & (-2.4097) & (-2.6875) & (0.2671) & (-2.6979) \\
(\text{se}) & (0.1470) & (0.2065) & (0.2011) & (0.2058) & (0.2010) & (0.1437) \\
(\text{fitness}) n = 111; & R^2 = 0.1934; \bar{R}^2 = 0.1550 \\
 & F^* = 5.03; p = 0.0004
\end{array}
$$

The regression results of reduced **price** equation show as follows:

$$
\begin{array}{lllllll}
\widehat{lprice} = & -0.27 & -0.11mon & -0.04tue & -0.01wed & +0.05thu & +0.35stormy \\
(\text{t}) & (-3.5569) & (-1.0525) & (-0.3937) & (-0.1106) & (0.4753) & (4.6387) \\
(\text{se}) & (0.0764) & (0.1073) & (0.1045) & (0.1069) & (0.1045) & (0.0747) \\
(\text{fitness}) n = 111; & R^2 = 0.1789; \bar{R}^2 = 0.1398 \\
 & F^* = 4.58; p = 0.0008
\end{array}
$$

Two-stage least squares (2SLS) regression results

Results of 2SLS with error correction

eq	vars	Estimate	Std. Error	t value	Pr(>|t|)
eq1	(Intercept)	8.5059	0.1662	51.1890	0.0000
eq1	lprice	-1.1194	0.4286	-2.6115	0.0103
eq1	mon	-0.0254	0.2148	-0.1183	0.9061
eq1	tue	-0.5308	0.2080	-2.5518	0.0122
eq1	wed	-0.5664	0.2128	-2.6620	0.0090
eq1	thu	0.1093	0.2088	0.5233	0.6018
eq2	(Intercept)	8.6284	0.3890	22.1826	0.0000
eq2	lprice	0.0011	1.3095	0.0008	0.9994
eq2	stormy	-0.3632	0.4649	-0.7813	0.4363

Two-stage least squares (2SLS) regression results

```
systemfit results
method: 2SLS

         N  DF SSR detRCov OLS-R2 McElroy-R2
system 222 213 110   0.107  0.094     -0.598

      N  DF  SSR   MSE  RMSE    R2 Adj R2
eq1 111 105 52.1 0.496 0.704 0.139  0.098
eq2 111 108 57.5 0.533 0.730 0.049  0.032

The covariance matrix of the residuals
      eq1   eq2
eq1 0.496 0.396
eq2 0.396 0.533

The correlations of the residuals
      eq1   eq2
eq1 1.000 0.771
eq2 0.771 1.000


2SLS estimates for 'eq1' (equation 1)
Model Formula: lquan ~ lprice + mon + tue + wed + thu
Instruments: ~mon + tue + wed + thu + stormy
```

Comparison: the biased OLS estimation

- tidy results of **bias** OLS estimation for the demand equation:

$$
\begin{aligned}
\widehat{lquan} =& +8.61 && -0.56lprice && +0.01mon && -0.52tue && -0.56wed && +0.08thu \\
\text{(t)} & (60.1698) && (-3.3443) && (0.0706) && (-2.6114) && (-2.7450) && (0.4126) \\
\text{(se)} & (0.1430) && (0.1682) && (0.2026) && (0.1977) && (0.2023) && (0.1978) \\
\text{(fitness)} & R^2 = 0.2205; \bar{R}^2 = 0.1834 \\
& F^* = 5.94; \quad p = 0.0001
\end{aligned}
$$

- tidy results of **bias** OLS estimation for the supply equation:

$$
\begin{aligned}
\widehat{lquan} =& +8.50 && -0.44lprice && -0.22stormy \\
\text{(t)} & (86.6914) && (-2.2560) && (-1.3253) \\
\text{(se)} & (0.0981) && (0.1942) && (0.1630) \\
\text{(fitness)} & R^2 = 0.0923; \bar{R}^2 = 0.0755 \\
& F^* = 5.49; \quad p = 0.0053
\end{aligned}
$$

Comparison: the biased OLS estimation

- raw R summry of **bias** OLS estimation for the demand equation:

```
Call:
lm(formula = fish.D, data = fultonfish)

Residuals:
    Min      1Q  Median      3Q     Max
-2.2384 -0.3674  0.0883  0.4230  1.2487

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)   8.6069     0.1430   60.17   <2e-16 ***
lprice       -0.5625     0.1682   -3.34   0.0011 **
mon           0.0143     0.2026    0.07   0.9438
tue          -0.5162     0.1977   -2.61   0.0103 *
wed          -0.5554     0.2023   -2.75   0.0071 **
thu           0.0816     0.1978    0.41   0.6807
---
Signif. codes:
0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.67 on 105 degrees of freedom
Multiple R-squared:  0.22,    Adjusted R-squared:  0.183
```

Comparison: the biased OLS estimation

- raw R summry of **bias** OLS estimation for the supply equation:

```
Call:
lm(formula = fish.S, data = fultonfish)

Residuals:
    Min      1Q  Median      3Q     Max
-2.4042 -0.3754  0.0734  0.5197  1.2267

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)   8.5009     0.0981   86.69   <2e-16 ***
lprice       -0.4381     0.1942   -2.26    0.026 *
stormy       -0.2160     0.1630   -1.33    0.188
---
Signif. codes:
0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.71 on 108 degrees of freedom
Multiple R-squared:  0.0923,    Adjusted R-squared:  0.0755
F-statistic: 5.49 on 2 and 108 DF,  p-value: 0.00534
```

End of this chapter!