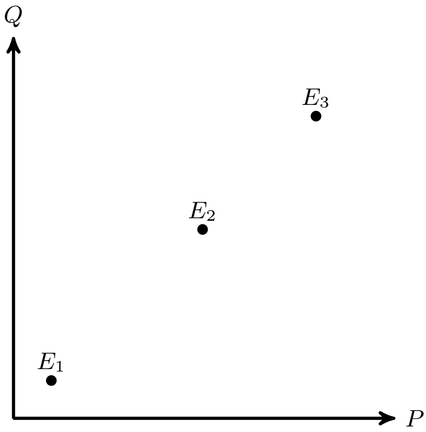
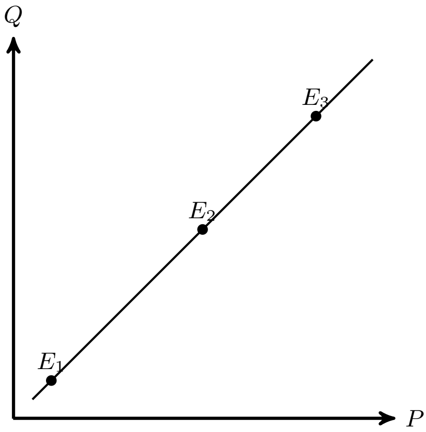
Simultaneous equations model. Parameter identification problem. Estimation method for SEM.

Jakub Mućk
SGH Warsaw School of Economics

Introduction



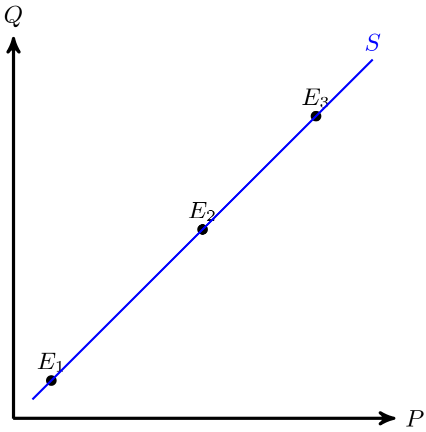
Q is the amount of good
P is the price of good
E_1, E_2 and E_3 are observed values.



Q is the amount of good
P is the price of good
E_1, E_2 and E_3 are observed values.

Relationship :

$$Q = \beta_0 + \beta_1 P + \varepsilon \quad (1)$$

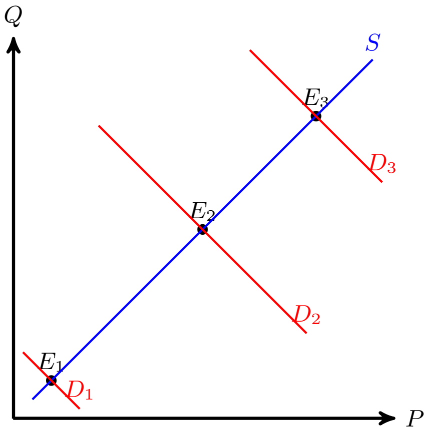


Q is the amount of good
P is the price of good
E_1, E_2 and E_3 are observed values.

The supply curve (S)

$$Q_S = \beta_0 + \beta_1 P + \varepsilon_s \qquad (1)$$

where ε_s is the **unobservable supply shifter**



Q is the amount of good
P is the price of good
E_1, E_2 and E_3 are observed values.

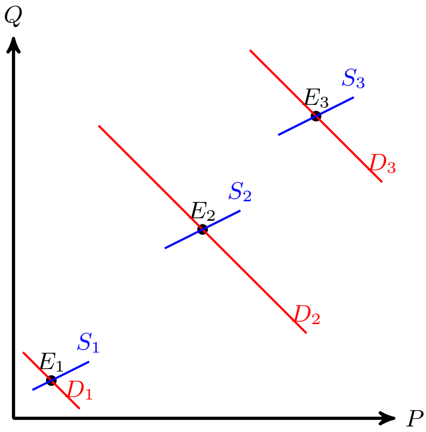
The supply curve (S)

$$Q_S = \beta_0 + \beta_1 P + \varepsilon_s \quad (1)$$

where ε_s is the **unobservable supply shifter**
The demand curves (D_i) for different values of Z

$$Q_D = \alpha_0 + \alpha_1 P + \alpha_2 Z + \varepsilon_d \quad (2)$$

where Z is the **observable demand shifter**, and ε_d is the **unobservable demand shifter**.



Q is the amount of good
P is the price of good
E_1, E_2 and E_3 are observed values.

The demand curves (S_i) for different values of X:

$$Q_S = \beta_0 + \beta_1 P + \beta_2 X + \varepsilon_s \quad (1)$$

where X is the **observable demand shifter**, and ε_s is the **unobservable supply shifter**

The demand curves (D_i) for different values of Z

$$Q_D = \alpha_0 + \alpha_1 P + \alpha_2 Z + \varepsilon_d \quad (2)$$

where Z is the **observable demand shifter**, and ε_d is the **unobservable demand shifter**.

Simultaneous Equations Models (SEM)

- **Structural form** is the the form derived from an economic theory.
- System of M linear equations:

$$\begin{array}{llllllll} \gamma_{11}y_1+ & \ldots+ & \gamma_{1M}y_M+ & \beta_{11}x_1+ & \ldots+ & \beta_{1K}x_K & = & \varepsilon_1 \\ \gamma_{21}y_2+ & \ldots+ & \gamma_{2M}y_M+ & \beta_{21}x_2+ & \ldots+ & \beta_{2K}x_K & = & \varepsilon_2 \\ & & & & & & \vdots & \\ \gamma_{M1}y_1+ & \ldots+ & \gamma_{MM}y_M+ & \beta_{M1}x_M+ & \ldots+ & \beta_{MK}x_K & = & \varepsilon_M \end{array}$$

- M **endogenous:** y_1, y_2, ..., y_M.
- K **exogenous variables:** x_1, x_2, ..., x_K.
- M **structural error terms/innovations:** ε_{1t}, ε_2, ..., ε_M.
- Assuming that $\varepsilon = [\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_M]$

$$\mathbb{E}(\varepsilon) = 0 \qquad \text{oraz} \qquad \mathbb{E}(\varepsilon\varepsilon^T) = \Sigma,$$

where Σ is the variance-covariance matrix of structural innovations.

- **Structural form** – matrix notation:

$$\mathbf{Y}\Gamma + \mathbf{X}\mathbf{B} = \varepsilon, \tag{3}$$

where (N denotes number of observations):
  - $\mathbf{Y}$ is a $N \times M$ matrix of endogenous variables.
  - $\mathbf{X}$ is a $N \times K$ matrix of exogenous variables.
  - ε is a $N \times M$ matrix of structural disturbances.
  - Γ is a $M \times M$ matrix.
  - $\mathbf{B}$ is $K \times M$ a matrix.
- In addition:

$$\mathbb{E}(\varepsilon) = 0 \qquad \text{oraz} \qquad \mathbb{E}(\varepsilon\varepsilon') = \Sigma,$$

where Σ is the variance-covariance matrix of the error term
- The structural form consists of: Γ, $\mathbf{B}$, and Σ.

- **Reduced form**:

$$\mathbf{Y} = \mathbf{X}\Pi + \nu, \tag{4}$$

- where

$$\Pi = -\mathbf{B}\Gamma^{-1} \tag{5}$$

and ν is a matrix of disturbances in the reduced form:

$$\nu = \varepsilon\Gamma^{-1} \tag{6}$$

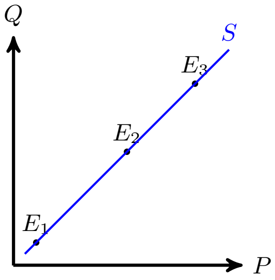
and

$$\mathbb{E}(\nu\nu^T) = \left(\Gamma^{-1}\right)^T \Sigma\Gamma^{-1} = \Omega. \tag{7}$$

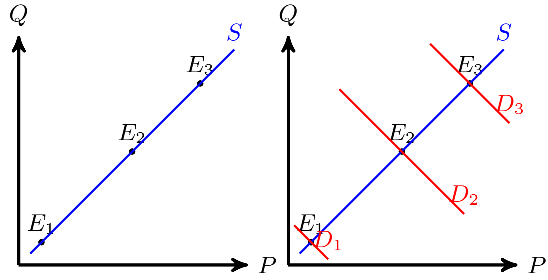
The Problem of Identification

 SGH

Empirical observations: E_1, E_2 and E_3.

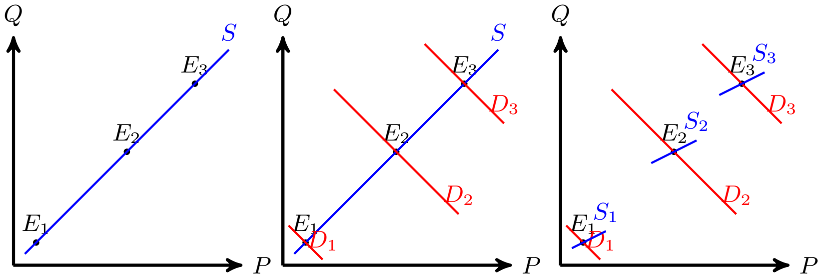


Empirical observations: E_1, E_2 and E_3.



SGH

Empirical observations: E_1, E_2 and E_3.



The Problem of Identification I

- **Structural form:**

$$\mathbf{Y}\Gamma + \mathbf{X}\mathbf{B} = \varepsilon \tag{8}$$

- **Reduced form:**

$$\mathbf{Y} = \mathbf{X}\Pi + \nu \tag{9}$$

$$\Pi = -\mathbf{B}\Gamma^{-1} \tag{10}$$

- It is importantly to have reduced form that has exactly one corresponding structural form.
- If there are several structural forms that correspond to reduced form $\Longrightarrow$ **lack od identification**
- For instance, let $\mathbf{F}$ denote some transformation (matrix). Then, alternative structural form, which will be some transformation of **true** structural form:

$$\mathbf{Y}\tilde{\Gamma} + \mathbf{X}\tilde{\mathbf{B}} = \tilde{\varepsilon} = \mathbf{Y}\Gamma\mathbf{F} + \mathbf{X}\mathbf{B}\mathbf{F} = \varepsilon\mathbf{F} \tag{11}$$

Then reduced form corresponding to *new/alternative* structura form will be the same as in the case of original structural form.

$$\tilde{\Pi} = -\mathbf{B}\mathbf{F}\mathbf{F}^{-\mathbf{1}}\Gamma^{-1} = \mathbf{B}\Gamma^{-1} = \Pi \tag{12}$$

The Problem of Identification II

- It implies that both structural form are empirically equivalent.
- **A popular strategy to identify parameters is to use non-sample information**. This refers to restrictions that can be derived from an economic theory.

- Example (q_d – demand; q_s – supply, p – price; z – exogenous variable):

$$q_d = \alpha_0 + \alpha_1 p + \alpha_2 z + \varepsilon_d \quad (13)$$
$$q_s = \beta_0 + \beta_1 p + \beta_2 z + \varepsilon_s \quad (14)$$
$$q_d = q_s \quad (15)$$

After manipulation and assuming that $\alpha_1 \neq \beta_1$ the reduced form is as follows:

$$q = \frac{\alpha_1\beta_0-\alpha_0\beta_1}{\alpha_1-\beta_1} + \frac{\alpha_1\beta_2-\alpha_2\beta_1}{\alpha_1-\beta_1}z + \frac{\alpha_1\varepsilon_s-\alpha_2\varepsilon_d}{\alpha_1-\beta_1} = \pi_{11} + \pi_{21}z + \nu_q, \quad (16)$$
$$p = \frac{\beta_0-\alpha_0}{\alpha_1-\beta_1} + \frac{\beta_1-\alpha 2}{\alpha_1-\beta_2}z + \frac{\varepsilon_s-\varepsilon_d}{\alpha_1-\beta_1} = \pi_{12} + \pi_{22}z + \nu_p, \quad (17)$$

Reduced form consists of **four parameters π_{11}, π_{12}, π_{21} i π_{22})**, while in the structural form we have **six parameters (α_0, α_1, α_2, β_0, β_1 i β_2).** Obviously, **there is no unique solution of system of six equations (parameters in structural form) with four variables (parameters in reduced form)**.

- Example (q_d – demand; q_s – supply, p – price; z and x – exogenous variables):

$$q_d = \alpha_0 + \alpha_1 p + \alpha_2 x + \varepsilon_d \quad (18)$$

$$q_s = \beta_0 + \beta_1 p + \beta_2 z + \varepsilon_s \quad (19)$$

$$q_d = q_s \quad (20)$$

Structural form:

$$[q \quad p]\begin{bmatrix} 1 & 1 \\ -\alpha_1 & -\beta_1 \end{bmatrix} + [1 \quad x \quad z]\begin{bmatrix} -\alpha_0 & -\beta_0 \\ -\alpha_2 & 0 \\ 0 & -\beta_2 \end{bmatrix} = [\varepsilon_d \quad \varepsilon_s] \quad (21)$$

Reduced form:

$$[q \quad p] = [1 \quad x \quad z]\begin{bmatrix} (\alpha_1\beta_0 - \alpha_0\beta_1)/\gamma & (\beta_0 - \alpha_0)/\gamma \\ -\alpha_2\beta_1/\gamma & -\alpha_2/\gamma \\ \alpha_1\beta_2/\gamma & \beta_2/\gamma \end{bmatrix} + [\nu_1 \quad \nu_2] \quad (22)$$

where $\gamma = \alpha_1 - \beta_1$
All structural forms, including false ones, can be described as follows:

$$\tilde{\mathbf{B}} = \mathbf{BF} = \begin{bmatrix} \alpha_0 f_{11} + \beta_0 f_{12} & \alpha_0 + \beta_0 f_{22} \\ \alpha_2 f_{11} & \alpha_2 f_{12} \\ \beta_2 f_{21} & \beta_2 f_{22} \end{bmatrix},$$

and looking at the **F** components

- when $f_{12} \neq 0 \implies$ then x appears in the supply equation. This is inconsistent with the assumed theory (structural form).
- when $f_{21} \neq 0 \implies$ then z appears in the demand equation. This is inconsistent with the assumed theory (structural form).

Summing up, if $f_{12} = f_{21} = 0$ then reduced form is consistent with structural form. $\mathbf{F} = \mathbf{I}$ allows to get original structural form.
Unique solution:

$$\alpha_0 = \pi_{11} - \pi_{12}\left(\frac{\pi_{31}}{\pi_{32}}\right), \quad \alpha_1 = \frac{\pi_{31}}{\pi_{32}}, \quad \alpha_2 = \pi_{22}\left(\frac{\pi_{21}}{\pi_{22}} - \frac{\pi_{31}}{\pi_{32}}\right),$$

$$\beta_0 = \pi_{11} - \pi_{12}\left(\frac{\pi_{21}}{\pi_{22}}\right), \quad \beta_1 = \frac{\pi_{21}}{\pi_{22}}, \quad \beta_2 = \pi_{32}\left(\frac{\pi_{31}}{\pi_{32}} - \frac{\pi_{21}}{\pi_{22}}\right).$$

- **Structural form:**
  - Γ is an $M \times M$ nonsingular matrix $\implies$ M^2 parameters.
  - $\mathbf{B}$ is an $K \times M$ nonsingular matrix $\implies$ KM parameters.
  - Σ is $M \times M$ symetric matrix $\implies$ $\frac{1}{2}M(M+1)$ parameters.

  The structural form consists of $M^2 + KM + \frac{1}{2}M(M+1)$. parameters

- **Structural form:**
  - Γ is an $M \times M$ nonsingular matrix $\implies$ M^2 parameters.
  - $\mathbf{B}$ is an $K \times M$ nonsingular matrix $\implies$ KM parameters.
  - Σ is $M \times M$ symetric matrix $\implies$ $\frac{1}{2}M(M+1)$ parameters.

  The structural form consists of $M^2 + KM + \frac{1}{2}M(M+1)$. parameters
- **Reduced form:**
  - Π is an $K \times M$ nonsingular matrix $\implies$ KM parameters.
  - Ω is $M \times M$ symetric matrix $\implies$ $\frac{1}{2}M(M+1)$ parameters.

  The structural form consists of $KM + \frac{1}{2}M(M+1)$ parameters.
- The difference between structural and reduced form equals M^2. Therefore, it is essential to use **non-sample information**.

1. **Normalization** – in each equation one endogenous variable has coefficient 1.
   - Then the number of additional parameters equals $M(M-1)$.
2. **Identities**.
3. **Exclusion** – zeros in matrices Γ and $\mathbf{B}$.
4. **Linear restrictions**.
   - But it can lead to false structure. For example, assuming constant returns to scale in production function lead to situation, in which effects of economies of scale is measured by technical change.
5. **Restrictions on Σ**.
   - In modern macroeconometrics, the common approach is to assumed uncorrelated structural shocks.

SGH

Order condition

The number of exogenous variables excluded from given equation must be not smaller than the number of endogenous variables included in given equation.

- The order condition should be checked for each equation.
- The order condition can be understood by comparing with IV.
- The order condition is not sufficient.

- Practically, we can use method with tableau which allows to check both conditions.
- [Step #1] Rewrite parameters of structural forms in the following way:

y_1	y_2	$\dots$	y_M	1	x_1	x_2	$\dots$	x_k
		Γ'				$\mathbf{B}'$		

- [Krok #2] For each equation (denoted by j) we consider submatrix from which we exclude:
  - j-th row;
  - columns which have nonzero elmenents in the $j$row.

j-th equation is identified when the above submatrix has full rank.

The equation is:

- **Underidentified** when the number of excluded exogenous variables is smaller than the number of endogenous variables or rank condition fails.
- **Exactly identified** when the number of excluded exogenous variables equals the number of endogenous variables and rank condition is satisfied.
- **Overidentified** when the number of excluded exogenous variables is greater than the number of endogenous variables and rank condition is satisfied.

$$
\begin{aligned}
q_d &= \alpha_0 + \alpha_1 p + \varepsilon_d \\
q_s &= \beta_0 + \beta_1 p + \varepsilon_s \\
q_d &= q_s
\end{aligned}
$$

q_d	q_s	1	p
1	0	$-\alpha_0$	$-\alpha_1$
0	1	$-\beta_0$	$-\beta_1$
1	-1	0	0

$$\begin{aligned} q_d &= \alpha_0 + \alpha_1 p + \varepsilon_d \\ q_s &= \beta_0 + \beta_1 p + \varepsilon_s \\ q_d &= q_s \end{aligned}$$

q_d	q_s	1	p
1	0	$-\alpha_0$	$-\alpha_1$
0	1	$-\beta_0$	$-\beta_1$
1	-1	0	0

first equation: $[1 \quad -1]'$ $\implies$ equation is underidentified.

SGH

$$
\begin{aligned}
q_d &= \alpha_0 + \alpha_1 p + \varepsilon_d \\
q_s &= \beta_0 + \beta_1 p + \varepsilon_s \\
q_d &= q_s
\end{aligned}
$$

q_d	q_s	1	p
1	0	$-\alpha_0$	$-\alpha_1$
0	1	$-\beta_0$	$-\beta_1$
1	-1	0	0

first equation: $[1 \quad -1]' \implies$ equation is underidentified.

second equation: $[1 \quad 1]' \implies$ equation is underidentified.

SGH

$$
\begin{aligned}
q_d &= \alpha_0 + \alpha_1 p + \alpha_2 z + \varepsilon_d \\
q_s &= \beta_0 + \beta_1 p + \varepsilon_s \\
q_d &= q_s
\end{aligned}
$$

q_d	q_s	1	p	z
1	0	$-\alpha_0$	$-\alpha_1$	$-\alpha_2$
0	1	$-\beta_0$	$-\beta_1$	0
1	-1	0	0	0

$$
\begin{aligned}
q_d &= \alpha_0 + \alpha_1 p + \alpha_2 z + \varepsilon_d \\
q_s &= \beta_0 + \beta_1 p + \varepsilon_s \\
q_d &= q_s
\end{aligned}
$$

q_d	q_s	1	p	z
1	0	$-\alpha_0$	$-\alpha_1$	$-\alpha_2$
0	1	$-\beta_0$	$-\beta_1$	0
1	-1	0	0	0

first equation: $[1 \quad -1]'$ $\implies$ equation is underidentified.

Example II

$$
\begin{aligned}
q_d &= \alpha_0 + \alpha_1 p + \alpha_2 z + \varepsilon_d \\
q_s &= \beta_0 + \beta_1 p + \varepsilon_s \\
q_d &= q_s
\end{aligned}
$$

q_d	q_s	1	p	z
1	0	$-\alpha_0$	$-\alpha_1$	$-\alpha_2$
0	1	$-\beta_0$	$-\beta_1$	0
1	-1	0	0	0

first equation: $[1 \quad -1]'$ $\implies$ equation is underidentified.

second equation: $[1 \quad 1]'$ and $[-\alpha_2 \quad 0]'$. In addition,

$$\text{rank}\begin{bmatrix} 1 & -\alpha_2 \\ 1 & 0 \end{bmatrix} = 2$$

$\implies$ equation is exactly identified.

$$q_d = \alpha_0 + \alpha_1 p + \alpha_2 z + \varepsilon_d$$
$$q_s = \beta_0 + \beta_1 p + \beta_2 x + \varepsilon_s$$
$$q_d = q_s$$

q_d	q_s	1	p	z	x
1	0	$-\alpha_0$	$-\alpha_1$	$-\alpha_2$	0
0	1	$-\beta_0$	$-\beta_1$	0	$-\beta_2$
1	-1	0	0	0	0

$$\begin{aligned} q_d &= \alpha_0 + \alpha_1 p + \alpha_2 z + \varepsilon_d \\ q_s &= \beta_0 + \beta_1 p + \beta_2 x + \varepsilon_s \\ q_d &= q_s \end{aligned}$$

q_d	q_s	1	p	z	x
1	0	$-\alpha_0$	$-\alpha_1$	$-\alpha_2$	0
0	1	$-\beta_0$	$-\beta_1$	0	$-\beta_2$
1	-1	0	0	0	0

first equation: $[1 \quad -1]'$ and $[-\beta_2 \quad 0]'$. Importantly,

$$\text{rank}\begin{bmatrix} 1 & -\beta_2 \\ -1 & 0 \end{bmatrix} = 2$$

$\implies$ equation is exactly identified.

$$\begin{aligned} q_d &= \alpha_0 + \alpha_1 p + \alpha_2 z + \varepsilon_d \\ q_s &= \beta_0 + \beta_1 p + \beta_2 x + \varepsilon_s \\ q_d &= q_s \end{aligned}$$

q_d	q_s	1	p	z	x
1	0	$-\alpha_0$	$-\alpha_1$	$-\alpha_2$	0
0	1	$-\beta_0$	$-\beta_1$	0	$-\beta_2$
1	-1	0	0	0	0

first equation: $[1 \quad -1]'$ and $[-\beta_2 \quad 0]'$. Importantly,

$$\text{rank} \begin{bmatrix} 1 & -\beta_2 \\ -1 & 0 \end{bmatrix} = 2$$

$\Longrightarrow$ equation is exactly identified.

second equation: $[1 \quad 1]'$ and $[-\alpha_2 \quad 0]'$. In addition,

$$\text{rank} \begin{bmatrix} 1 & -\alpha_2 \\ 1 & 0 \end{bmatrix} = 2$$

$\Longrightarrow$ equation is exactly identified.

Estimation Methods

Estimation Methods

SGH

Single equation

- OLS/LS (*ordinary least squares*);
- 2SLS (*two-stages least squares*;
- GMM (*generalized methods of moments*);
- LIML (*limited information maximum likelihood*).

System estimation

- 3SLS (*three-stages least squares*);
- GMM (*generalized methods of moments*);
- FIML (*full-information maximum likelihood*).

- **Structural form of j-th equation:**

$$\begin{aligned} \mathbf{y}_j &= \mathbf{Y}_j\gamma_j + \mathbf{X}_j\beta_j + \varepsilon_j \\ &= \mathbf{Z}_j\delta_j + \varepsilon_j \end{aligned}$$

where $\mathbf{Z}_j$ is the matrix containing both endogenous and exogenous variables that appears in the j-th equation, i.e. $\mathbf{Z}_j = [\mathbf{Y}_j \quad \mathbf{X}_j]$.

- **Reduced form of j-th equation:**

$$\mathbf{Y}_j = \Pi_j\mathbf{X}_j + \nu_j$$

where Π_j is the j-th column of the matrix Π and $\nu = \varepsilon\Gamma^{-1}$.

- **Structural form of j-th equation:**

$$\begin{aligned}\mathbf{y}_j &= \mathbf{Y}_j\gamma_j + \mathbf{X}_j\beta_j + \varepsilon_j \\ &= \mathbf{Z}_j\delta_j + \varepsilon_j\end{aligned}$$

  where $\mathbf{Z}_j$ is the matrix containing both endogenous and exogenous variables that appears in the j-th equation, i.e. $\mathbf{Z}_j = [\mathbf{Y}_j \quad \mathbf{X}_j]$.

- **Reduced form of j-th equation:**

$$\mathbf{Y}_j = \Pi_j\mathbf{X}_j + \nu_j$$

  where Π_j is the j-th column of the matrix Π and $\nu = \varepsilon\Gamma^{-1}$.

- **The OLS estimator is inconsistent:**

$$\hat{\delta}_j^{OLS} = \left[\mathbf{Z}_j'\mathbf{Z}_j\right]^{-1}\mathbf{Z}_j'\mathbf{y}_j = \delta_j + \begin{bmatrix} \mathbf{Y}_j'\mathbf{Y}_j & \mathbf{Y}_j'\mathbf{X}_j \\ \mathbf{X}_j'\mathbf{Y}_j & \mathbf{X}_j'\mathbf{X}_j \end{bmatrix}^{-1}\begin{bmatrix} \mathbf{Y}_j'\varepsilon_j \\ \mathbf{X}_j'\varepsilon_j \end{bmatrix}$$

  - $\text{plim}\frac{1}{N}\mathbf{X}_j^T\varepsilon_j$ tends to zero as $\mathbf{X}_j$ are non-random (or exogenous).
  - **The key problem is $\text{plim}\frac{1}{N}\mathbf{Y}_j^T\varepsilon_j$ which does not tend to 0** . This correlation leads to **simultaneous equations bias**.

- There are special cases when the OLS is consistent: recursive models with uncorrelated structural errors.

- **2SLS estimator (*two-stage least squares 2SLS*)** bases on IV approach.
- Key idea: use $\mathbf{X}$ as instruments for $\mathbf{Y}_j$ in the j-the equation.
- **[Step #1]** Regress all endogenous variables j ($\mathbf{Y}_j$) on exogenous variables. $\Longrightarrow$ Calculate fitted values $\hat{\mathbf{Y}}_j$.
- **[Step #2]** Regress $\mathbf{y}_j$ on $\hat{\mathbf{Y}}_j$ and $\mathbf{X}_j$.
  - Above strategy illustrate why the **order condition** is so important.
  - **The IV perspective**: if the order condition is not satisfied then the number of endogenous variables is greater than number of instruments.

- The 2SLS estimator for the j-th equation that exploits the fitted values from the first step ($\hat{\mathbf{Y}}_j$):

$$\hat{\delta}_j^{2SLS} = \begin{bmatrix} \hat{\mathbf{Y}}_j^T \mathbf{Y}_j & \hat{\mathbf{Y}}_j^T \mathbf{X}_j \\ \mathbf{X}_j^T \mathbf{Y}_j & \mathbf{X}_j^T \mathbf{X}_j \end{bmatrix}^{-1} \begin{bmatrix} \hat{\mathbf{Y}}_j^T \mathbf{y}_j \\ \mathbf{X}_j^T \mathbf{y}_j \end{bmatrix}. \quad (23)$$

- Asymptotic variance:

$$Var\left(\hat{\delta}_j^{2SLS}\right) = \hat{\sigma}_{jj}\left[\hat{\mathbf{Z}}_j^T \hat{\mathbf{Z}}_j\right]^{-1}, \quad (24)$$

where $\mathbf{Z}_j$ is matrix containing all variables in the j-th equation. The variance of the error term can be estimated with:

$$\hat{\sigma}_{jj} = \frac{\left(\mathbf{y}_j - \mathbf{Z}_j\hat{\delta}_j^{2SLS}\right)^T \left(\mathbf{y}_j - \mathbf{Z}_j\hat{\delta}_j^{2SLS}\right)}{N}, \quad (25)$$

and depends on observed variance ($\mathbf{Z}_j$).

- **3SLS estimator (*three-stage least squares*)** is system estimation method, i.e., parameters are estimated jointly.
- Key concept:
  - Use the 2SLS estimator for each equation.
  - In the next step, use feasible generalized least squares/ seemingly unrelated regression (FGLS/SUR) estimator that accounts for cross-equation correlation of the error term.

- Matrix notation:

$$\begin{bmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \\ \vdots \\ \mathbf{y}_M \end{bmatrix} \begin{bmatrix} \mathbf{Z}_1 & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{Z}_2 & \dots & \mathbf{0} \\ \vdots & \vdots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & \mathbf{Z}_M \end{bmatrix} \begin{bmatrix} \delta_1 \\ \delta_2 \\ \vdots \\ \delta_M \end{bmatrix} = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_M \end{bmatrix}, \tag{26}$$

  where $\mathbf{Z}_i$ is a set of exogenous and endogenous variable in the i-th equation and $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_M$ are the structural disturbances.
- Key assumption about the error term:
  - $\mathbb{E}(\varepsilon|\mathbf{X}) = 0$
  - $\mathbb{E}(\varepsilon\varepsilon'|\mathbf{X}) = \bar{\Sigma} = \Sigma \otimes I$.
- Variance-covariance of the error term:

$$\Sigma \otimes I = \begin{bmatrix} \sigma^2_{11} & \sigma^2_{12} & \dots & \sigma^2_{1M} \\ \sigma^2_{21} & \sigma^2_{22} & \dots & \sigma^2_{2M} \\ \vdots & \vdots & \vdots & \vdots \\ \sigma^2_{M1} & \sigma^2_{M2} & \dots & \sigma^2_{MM} \end{bmatrix} \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & 1 \end{bmatrix}, \tag{27}$$

- The cross-equation correlation can be estimated with the 2SLS estimates:

$$\hat{\sigma}_{ij} = \frac{\left(\mathbf{y}_i - \mathbf{Z}_i \hat{\delta}_i^{2SLS}\right)' \left(\mathbf{y}_i - \mathbf{Z}_j \hat{\delta}_j^{2SLS}\right)}{N}, \tag{28}$$

- Finally, the standard FGLS can be applied:

$$\hat{\delta}^{3SLS} = \left[\hat{\mathbf{Z}}' \left(\Sigma^{-1} \otimes I\right) \hat{\mathbf{Z}}\right]^{-1} \hat{\mathbf{Z}}' \left(\Sigma^{-1} \otimes I\right) \mathbf{y}. \tag{29}$$