Econometrics I

Workshop IX

Mar 18, 2023

Simultaneous Equation

What came first?

The hen or the egg?

Real world and real problems have feedback effects and bidirectional causal effects that require the application of simultaneous equations

Single equation models can be represented as:

$$y_x = b_0 + b_1 X_1 + b_2 X_2 + U_t$$

A simultaneous equation system is the one in which Y has effect over at least one of the X besides the effect that the rest of X has over Y

This type of models distinguish variables that are simultaneously determined (Ys), referring them as endogenous, from variables that are not, referring them as exogenous (Xs)

$$y_{1t} = \alpha_0 + \alpha_1 Y_{2t} + \alpha_2 X_{1t} + \alpha_3 X_{2t} + U_{1t} \tag{1.1}$$

$$y_{1t} = \beta_0 + \beta_1 Y_{1t} + \beta_2 X_{3t} + \beta_3 X_{2t} + U_{1t} \tag{1.2}$$

$$y_{1t} = \alpha_0 + \alpha_1 Y_{2t} + \alpha_2 X_{1t} + \alpha_3 X_{2t} + U_{1t}$$

$$y_{1t} = \beta_0 + \beta_1 Y_{1t} + \beta_2 X_{3t} + \beta_3 X_{2t} + U_{1t}$$

Structural Equations

Structural equations are inherent to economic theory that relies upon every endogenous variable, expressing it in terms of exogenous and endogenous variables.

$$y_{1t} = \boxed{\alpha_0} + \boxed{\alpha_1} Y_{2t} + \boxed{\alpha_2} X_{1t} + \boxed{\alpha_3} X_{2t} + U_{1t}$$

Structural Coefficients

$$y_{1t} = \boxed{\beta_0} + \boxed{\beta_1} Y_{1t} + \boxed{\beta_2} X_{3t} + \boxed{\beta_3} X_{2t} + U_{1t}$$

Note: Delayed endogenous variables may appear in models

They are called *predetermined variables*, and can be considered as exogenous variables

Let

$$Q_{Dt} = \alpha_0 + \alpha_1 P_t + \alpha_2 X_{1t} + \alpha_3 X_{2t} + U_{Dt} \quad (1.3)$$

$$Q_{St} = \beta_0 + \beta_1 P_t + \beta_2 X_{3t} + \qquad + U_{St} \quad (1.4)$$

$$Q_{dt} = Q_{st} \quad (1.5)$$

Where:

$Q_{Dt} = demanded\ quantity$

$Q_{St} = supplied\ quantity$

$P = price\ (own)$

$X_1 = Substitute\ good$

$X_2 = Income$

$X_3 = price\ for\ one\ factor\ of\ production$

$$Q_{Dt} = \alpha_0 + \alpha_1 P_t + \alpha_2 X_{1t} + \alpha_3 X_{2t} + U_{Dt}$$

$$Q_{St} = \beta_0 + \beta_1 P_t + \beta_2 X_{3t} + \qquad\qquad + U_{St}$$

$$Q_{dt} = Q_{st}$$

$$Q_{Dt} = \alpha_0 + \alpha_1 P_t + \alpha_2 X_{1t} + \alpha_3 X_{2t} + U_{Dt}$$

$$Q_{St} = \beta_0 + \beta_1 P_t + \beta_2 X_{3t} + \qquad + U_{St}$$

Q: Can these equation be estimated in isolation?

A: NO, estimations may be biased and inconsistent

Estimation strategy should be another

Reduced form equations: is the one that comes from expressing endogenous depending on exogenous and predetermined variables only

$$Y_{1t} = \boxed{\pi_0} + \boxed{\pi_1} X_t + \boxed{\pi_2} X_{2t} + \boxed{\pi_3} X_{3t} + V_{1t}$$

Reduced form coefficients

$$Y_{2t} = \boxed{\pi_4} + \boxed{\pi_5} X_{1t} + \boxed{\pi_6} X_{2t} + \boxed{\pi_7} X_{3t} + V_{1t}$$

$$Y_{1t} = \pi_0 + \boxed{\pi_1} X_t + \pi_2 X_{2t} + \pi_3 X_{3t} + V_{1t}$$

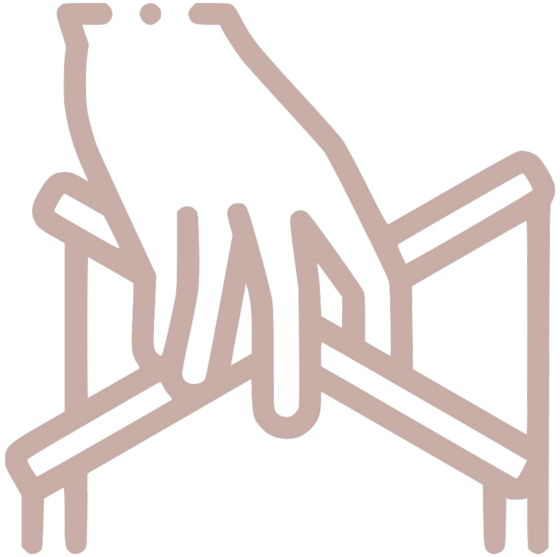
Shock multipliers

$$Y_{2t} = \pi_4 + \boxed{\pi_5} X_{1t} + \pi_6 X_{2t} + \pi_7 X_{3t} + V_{1t}$$

Shock multipliers quantify the impact from endogenous variable for a unitary change in the value of a predetermined variable after allowing the feedback effects in a complete system

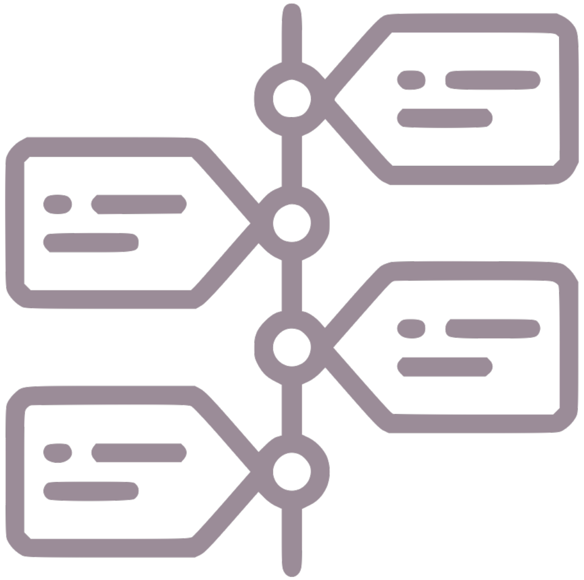
Four reasons to use reduced form equations:

1. Reduced form equations does not have any intrinsic simultaneity, they do not attempt against the $Cov(x, u) = 0$ assumption. Thus, they can be estimated under OLS



2. Reduced form coefficients can be mathematically manipulated in order to obtain structural coefficients. In other words, estimations from equations 1.6 and 1.7 can be used to solve the original equations

Four reasons to use reduced form equations:



3. Interpretation from reduced form coefficients can have an economic meaning

4. They play an important role in one of the most important techniques for simultaneous equations models: two-stage least-squares

Problem of identification

Identification is a previous condition for the use of two-stage least squares to equations in a simultaneous equation model

A structural equation is identified if and only if an enough number of predetermined variables in the system is greater or equal to the number of coefficients (slopes) of the equation we need to identify. (Notice that an equation from a simultaneous equation system can be identified but another from the same system may not)

Order condition is a systematic method to determine if a particular equation in a simultaneous equation system has the potential to be identified

If an equation comply with the order condition, it is feasible to be identified but we cannot ensure it

It is said that order condition is a necessary condition but not sufficient for identification

We must recognize the type of variables in a system:

Endogenous variables

They are determined inside the system in current period

Exogenous variables

They are determined outside the system

Predetermined variables

Exogenous and lagged endogenous Inside the model

Thus, for each equation in the system we need to determine…

1. …the number of predetermined variables (exogenous and lagged endogenous

2. …the number of slope estimated coefficients for an equation

$$\underbrace{Number\ of\ predetermined\ variables}_{(in\ model)} \geq \underbrace{Number\ of\ slope\ coefficients}_{(in\ equation)}$$

Identification:

For the supply-demand model:

$$Q_{Dt} = \alpha_0 + \alpha_1 P_t + \alpha_2 X_{1t} + \alpha_3 X_{2t} + U_{Dt}$$

$$Q_{St} = \beta_0 + \beta_1 P_t + \beta_2 X_{3t} + \quad + U_{St}$$

$$Q_{dt} = Q_{st}$$

Precisely identification:

$$Q_{Dt} = \alpha_0 + \alpha_1 P_t + \alpha_2 X_{1t} + \alpha_3 X_{2t} + U_{Dt}$$

This equation is identified for the order condition given that predetermined variables in the model X_1, X_2, X_3 is equal to the number of slope coefficients in the model $\alpha_1, \alpha_2, \alpha_3$

Thus, this equation is precisely identified for the order condition

Over-identification:

$$Q_{St} = \beta_0 + \beta_1 P_t + \beta_2 X_{3t} + \qquad + U_{St}$$

On the other hand, this equation is also identified for the order condition. There is three predetermined variables in the system (exogenous and lagged endogenous) but there is only two slope-coefficients in equation.

This implies that this equation is over-identified

A macroeconomic example:

$$Y_t = CO_t + I_1 + G_t + NX_t \qquad (1.6)$$

$$CO_t = \beta_0 + \beta_1 YD_1 + \beta_2 CO_{t-1} + e_{it} \qquad (1.7)$$

$$YD_t = Y_t + T_t \qquad (1.8)$$

$$I_t = 3 + \beta_4 Y_t + \beta_5 r_{t-1} + e_{2t} \qquad (1.9)$$

1. We notice five predetermined variables (exogenous and lagged endogenous) in the model (G_t, NX_t, CO_{t-1}, T_t and r_{t-1})
2. Equation (1.7) has two slope coefficients (β_1, β_2), so this equation is over-identified ($5 > 2$) and meet the identification order
3. We can verify that equation (1.9) is over-identified. The two-staged least squares method does not require to verify the characteristics of identification of entities.

```
// Dataset: Mroz.dta
// Run Regression
reg lwage educ age kidslt6 nwifeinc

// Storage
estimates store mco_e1
```

```
// Determine second equation
reg lwage hours educ exper expers

// Storage it
estimates store mco_e2

// Visualize it
estimates table mco_e1 mco_2, star
```

```
// Statistics
var dep: hours
var ind: educ kidslt6 nwifeinc

// Endogenous covariables
var endogena:lwage

// Uniequational regression with VI
var instrumentales:educ age kidslt6
nwifeinc exper expersq
```

```
estimates store mc2e_1

lwage educ exer expersq (hours=…)

estimates store mce2_e

// Compare models
estimates table mco_e1 mco_e2 mc2e_2

mc2e_2, star
```

```
// EC1
Dep:hours

// Independiente (exogenous y endogenous)
Independiente: lwage educ age kidslt6 nwifeinc

// EC2
Dep:lwage

// Independiente (exogenous y endogenous)
Independiente: hours edic exper expersq

// Endogenous
Educ age kidslt6 nwifeinc exper expersq
```

```
// Store
estimates store mc3e

// Compare all models
estimates table mco_e1 mco_e2 mc2e_1 mc2e_2 mc3e, star
```

References

- **Salvatore, D., & Sarmiento, J. C.** (1983). *Econometría* (No. HB141 S39). McGraw-Hill.
- **Gujarati, D. N.** (2009). Basic econometrics. Tata McGraw-Hill Education.
- **Wooldridge, J.M.** (2016). *Introductory Econometrics*, Cengage Learning, 6th edition.