Simultaneous Equations Models

Econometrics

J. Eduardo Vera-Valdés

eduardo@math.aau.dk

Department of Mathematical Sciences
Aalborg University
Denmark

AALBORG UNIVERSITY
DENMARK

Outline

Introduction

An Example

Simultaneous Equations

- 2SLS: IV
- 3SLS: (F)GLS
- GMM

Two Special Cases

Tests

Summing Up

J. Eduardo Vera-Valdés
eduardo@math.aau.

Department of Mathematical Sciences
Aalborg University
Denmark

Outline

J. Eduardo Vera-Valdés
eduardo@math.aau.

Department of Mathematical Sciences
Aalborg University
Denmark

Introduction

Multivariate models allow us to jointly determine the values of two or more dependent variables.

In SUR models we assumed that all of the regressors were independent from the regressands. Thus the “seemingly unrelated”.

We abstracted from the possibility of **simultaneity**. That is, one or more of the explanatory variables is jointly determined with the dependent variable.

Classical example of simultaneity are supply-demand equations, housing-savings relations, etc.

J. Eduardo Vera-Valdés
eduardo@math.aau.

Department of Mathematical Sciences
Aalborg University
Denmark

Outline

J. Eduardo Vera-Valdés
eduardo@math.aau.

Department of Mathematical Sciences
Aalborg University
Denmark

We start with the classical supply-demand example.

A simple version of the model is given by

$$\begin{aligned} q_{d,t} &= \alpha_1 p_t + \alpha_2 x_t + \varepsilon_{d,t}, \\ q_{s,t} &= \beta_1 p_t + \varepsilon_{s,t}, \\ q_{d,t} &= q_{s,t}, \end{aligned}$$

Written like this, the equations are called **structural equations**.

J. Eduardo Vera-Valdés
eduardo@math.aau.

Department of Mathematical Sciences
Aalborg University
Denmark

Structural equations are derived from theory and describe a particular aspect of the economy.

Exogenous variables are those determined outside of the model.

Endogenous variables are jointly determined inside the model.

The system is called **complete** since the number of equations is the same as the number of endogenous variables.

J. Eduardo Vera-Valdés
eduardo@math.aau.

Department of Mathematical Sciences
Aalborg University
Denmark

Solving the equations for p and q, we obtain the **reduced form equations**

Note that, $Cov[p, \varepsilon_d] \neq 0$ and $Cov[p, \varepsilon_s] \neq 0$, so neither equation satisfies the OLS assumptions.

Simultaneous Equations Models
J. Eduardo Vera-Valdés
eduardo@math.aau.

Department of Mathematical Sciences
Aalborg University
Denmark

Simultaneous Equations
An Example

Since OLS is inconsistent, we could use IV.

Which requires us to find instruments.

One option is to use the exogenous variables as instruments.

Identification problems may arise depending on the number of exogenous/endogeoues variables.

J. Eduardo Vera-Valdés
eduardo@math.aau.

Department of Mathematical Sciences
Aalborg University
Denmark

Outline

J. Eduardo Vera-Valdés
eduardo@math.aau.

Department of Mathematical Sciences
Aalborg University
Denmark

Simultaneous Equations

The general formulation can be written as

$$y_i = X_i\beta_i + u_i = Z_i\beta_{1i} + Y_i\beta_{2i} + u_i, \quad i = 1, \cdots, g;$$

where X_i is an $n \times k_i$ matrix of explanatory variables partitioned as $X_i = [Z_i \;\; Y_i]$, Z_i is an $n \times k_{1i}$ matrix of exogenous variables, and Y_i is an $n \times k_{2i}$ matrix of endogenous variables, $k_{1i} + k_{2i} = k_i$.

Note that we have as many equations as endogenous variables: **completeness condition**.

We assume the errors are contemporaneously correlated but show no serial correlation and are homoskedastic.

J. Eduardo Vera-Valdés
eduardo@math.aau.

Department of Mathematical Sciences
Aalborg University
Denmark

Simultaneous Equations

To estimate the model, we write the system as

$$Y_\bullet = X_\bullet \beta_\bullet + U_\bullet,$$

where we have stacked the observations vertically like $Y_\bullet = [Y_1^T, \cdots, Y_g^T]^T$.

The errors' covariance matrix is given by

$$E[U_\bullet U_\bullet^T] = \Sigma \otimes I_n$$

Note that the model has endogenous variables and correlated error terms.

Outline

J. Eduardo Vera-Valdés
eduardo@math.aau.

Department of Mathematical Sciences
Aalborg University
Denmark

Simultaneous Equations
Instruments and Order Conditions

Recall that the i-th equation in the system is given by

$$y_i = X_i\beta_i + u_i = Z_i\beta_{1i} + Y_i\beta_{2i} + u_i.$$

For equation i, we use the exogenous variables not included in the equation as instruments for Y_i.

This imposes a restriction, the **order condition**: we need at least the same number of exogenous variables not included in the equation as the number of endogenous variables included.

An easy rule of thumb is that the system satisfies the conditions if every equation has its own unique exogenous regressor.

J. Eduardo Vera-Valdés
eduardo@math.aau.

Department of Mathematical Sciences
Aalborg University
Denmark

Assuming the order conditions are satisfied, define the matrix of instruments as

$$W = [Z_1 \; Z_2 \; \cdots Z_g].$$

If we ignore the contemporaneous correlation, we can estimate equation i by IV as

$$\hat{\beta}_i^{2SLS} = \left(X_i^T P_W X_i\right)^{-1} X_i^T P_W y_i.$$

Alternatively, recall that $\hat{\beta}_i^{2SLS}$ can also be computed by first running the regression

$$Y_i = W\delta + V_i,$$

retrieving the fitted values, $\hat{Y}_i$, and running a second regression,

$$\hat{\beta}_i^{2SLS} = \left([Z_i \; \hat{Y}_i]^T [Z_i \; \hat{Y}_i]\right)^{-1} [Z_i \; \hat{Y}_i]^T y_i.$$

Simultaneous Equations Models

J. Eduardo Vera-Valdés
eduardo@math.aau.

Department of Mathematical Sciences
Aalborg University
Denmark

Its covariance matrix is given by

$$Var(\hat{\beta}_i^{2SLS}) = \sigma_{ii} \left(X_i P_W X_i \right)^{-1}.$$

Where we can obtain a consistent estimator of σ_{ii} by

$$\hat{\sigma}_{ii} = (1/T)(y_i - X_i \hat{\beta}_i^{2SLS})^T (y_i - X_i \hat{\beta}_i^{2SLS}).$$

Nonetheless, note that we have ignored the contemporaneous correlation and thus missed efficiency.

J. Eduardo Vera-Valdés
eduardo@math.aau.

Department of Mathematical Sciences
Aalborg University
Denmark

Outline

J. Eduardo Vera-Valdés
eduardo@math.aau.

Department of Mathematical Sciences
Aalborg University
Denmark

Simultaneous Equations
3SLS

Assuming we know Σ completely, we can use GLS to recover efficiency,

$$\hat{\beta}_{\bullet}^{3SLS} = ((P_W X_{\bullet})^T (\Sigma^{-1} \otimes I) P_W X_{\bullet})^{-1} ((P_W X_{\bullet})^T (\Sigma^{-1} \otimes I) P_W Y_{\bullet}).$$

Its covariance matrix is

$$Var(\hat{\beta}_{\bullet}^{3SLS}) = ((P_W X_{\bullet})^T (\Sigma^{-1} \otimes I) P_W X_{\bullet})^{-1}.$$

J. Eduardo Vera-Valdés
eduardo@math.aau.

Department of Mathematical Sciences
Aalborg University
Denmark

Alternatively, assume that we know Δ where $\Sigma = \sigma^2\Delta$; that is, we know the covariance matrix up to a constant.

The estimator is the same as before, and we estimate its covariance matrix by

$$Var(\hat{\beta}_{\bullet}^{3SLS}) = \hat{\sigma}^2((P_W X_{\bullet})^T(\Delta^{-1} \otimes I)P_W X_{\bullet})^{-1},$$

where

$$\hat{\sigma}^2 = \frac{1}{gn}\hat{U}_{\bullet}^T(\Delta^{-1} \otimes I)\hat{U}_{\bullet}.$$

Simultaneous Equations Models
J. Eduardo Vera-Valdés
eduardo@math.aau.

Department of Mathematical Sciences
Aalborg University
Denmark

Simultaneous Equations
FGLS

In practice, Σ is not known, so we have to estimate it.

Analogous to previous discussions, we use the residuals from the inefficient estimator to construct estimates of the covariance matrix,

$$\hat{\Sigma}_{2SLS} = \frac{1}{n}\hat{U}_{2SLS}^T\hat{U}_{2SLS},$$

where $\hat{U}_{2SLS}$ is a $n \times g$ matrix with the 2SLS residuals.

Thus, the FGLS estimator, or 3SLS, has a similar expression as before replacing Σ with $\hat{\Sigma}_{2SLS}$.

J. Eduardo Vera-Valdés
eduardo@math.aau.

Department of Mathematical Sciences
Aalborg University
Denmark

Outline

J. Eduardo Vera-Valdés

eduardo@math.aau.

Department of Mathematical Sciences Aalborg University Denmark

Another way to estimate the model is to look for an estimator that satisfies the moment conditions imposed by IV and GLS.

The conditions are met by GMM defined by

$$\hat{\beta}_{\bullet}^{GMM} = (X_{\bullet}^{T}(\Sigma^{-1} \otimes P_W)X_{\bullet})^{-1}(X_{\bullet}^{T}(\Sigma^{-1} \otimes P_W)Y_{\bullet}).$$

Its covariance matrix is

$$Var(\hat{\beta}_{\bullet}^{GMM}) = (X_{\bullet}^{T}(\Sigma^{-1} \otimes P_W)X_{\bullet})^{-1}.$$

Defined like this, GMM is consistent and efficient. (More next week.)

Moreover, GMM is equivalent to 3SLS.

J. Eduardo Vera-Valdés
eduardo@math.aau.

Department of Mathematical Sciences
Aalborg University
Denmark

Outline

J. Eduardo Vera-Valdés
eduardo@math.aau.

Department of Mathematical Sciences
Aalborg University
Denmark

Similar to the SUR, there are two special cases where estimating a model using 3SLS is not more efficient as just estimating each equation by 2SLS:

- If the matrix Σ is diagonal.
- When each equation is just identified; that is, the number of instruments is equal to the number of explanatory variables.

J. Eduardo Vera-Valdés
eduardo@math.aau.

Department of Mathematical Sciences
Aalborg University
Denmark

Outline

J. Eduardo Vera-Valdés
eduardo@math.aau.

Department of Mathematical Sciences
Aalborg University
Denmark

Given that we are using IV for estimation, we can/should test for overidentifying restrictions.

Since the diagonality of Σ was used to improve efficiency using 3SLS, it is important to test if whether it is diagonal.

J. Eduardo Vera-Valdés
eduardo@math.aau.

Department of Mathematical Sciences Aalborg University Denmark

Simultaneous Equations

Test

Another test of interest is whether some parameters between equations are the same.

We can use a Wald-type test. Define test statistic by

$$W = (R\hat{\beta}_{\bullet} - r)^T [R Var(\hat{\beta}_{\bullet}) R^T]^{-1} (R\hat{\beta}_{\bullet} - r),$$

which follows a chi-square distribution with degrees of freedom equal to the number of restrictions. Where R and r are matrices for the linear restrictions.

J. Eduardo Vera-Valdés
eduardo@math.aau.

Department of Mathematical Sciences
Aalborg University
Denmark

Outline

J. Eduardo Vera-Valdés
eduardo@math.aau.

Department of Mathematical Sciences
Aalborg University
Denmark

Summing Up

- We have studied models for simultaneously determined multiple dependent variables.
- We show that simultaneity makes OLS inconsistent.
- We can recover consistency by using IV (2SLS) with the exogenous regressors as instruments.
- A necessary condition for estimation is that in each equation there are more exogenous regressors excluded that endogenous included.
- 2SLS is not efficient, we can gain efficiency by using (F)GLS (3SLS).
- We can make hypothesis tests on whether some parameters should be the same across equations.

J. Eduardo Vera-Valdés
eduardo@math.aau.

Department of Mathematical Sciences
Aalborg University
Denmark