MSc Economics - Econometrics II ([redacted])

Week 3: Simultaneous Equations

[redacted]

You are given the following system of equations

$$\begin{aligned} y_{1,t} &= \gamma_{10} + \beta_{12} y_{2,t} + u_{1,t} \\ y_{2,t} &= \gamma_{20} + \beta_{21} y_{1,t} + \gamma_{21} x_{1,t} + u_{2,t} \end{aligned} \tag{1}$$

where: $\mathbb{E}(u_{i,t}) = 0,\ \forall t$ and $i = 1,2$; $\mathbb{E}(u_{i,t}u_{j,s}) = 0,\ \forall t, s$ and $i \neq j$; $\mathbb{E}(X'u_i) = 0,\ i = 1,2$.

Notation You are used to systems of regression equations $y = X\beta + u$ where the set of regressors X is also the set of exogenous variables in the system, which can therefore be estimated consistently by OLS. In these notes the set of regressors includes other *endogenous* variables, $y_{i,t}$, so the system (1) cannot be estimated consistently by OLS. Why not?

Expectations Make sure you can translate between the verbal description on the question sheet and the symbolic statement above. If not *ask me*, this stuff will make no sense if you don't know what is exogenous and what is endogenous etc., and what that implies for estimation. *Hint:* think about the Gauss-Markov assumptions!

Intercept In a given equation an intercept is a coefficient on a 1. We denote a vector of ones $\iota = [1, 1, \ldots, 1]'$ which is a $T \times 1$ vector if there are T observations on each equation. ι is a non-stochastic, and therefore exogenous regressor; it goes in the set, X, of exogenous regressors even though you cannot see a 1 in any of the equations above.

Order Cond The order condition tells us that there must be at least as many restrictions per equation as there are endogenous variables in the model. Equivalently, the number of excluded variables, of both sorts, in equation i must be at least as big as $n - 1$, where there are n *endogenous* variables in the entire model. It is only a *necessary* condition. However, it can distinguish between just- and over-identified equations.

Rank Cond The rank condition is a *sufficient* condition for the identification of equation i. However, it will not tell you if an identified equation is just- or over-identified. It tells us if there is a unique solution to the i^{th} equation in the system, in which case the i^{th} equation is identified; or if there is a linear combination of other system equations which is indestinguishable from the i^{th} equation, in which case it is not identified. Following Walter's notes, p 8, the rank condition can be stated by collecting vectors of coefficients associated with variables excluded from the i^{th} equation and testing the rank of the resulting matirx.

Question a) Discuss the above equations in terms of the order and rank conditions:

Begin by re-arranging the equations so that each is written in terms of $u_{i,t}$. We want to write the system in matrix form, with errors in terms of endogenous and exogenous variables:

$$\begin{aligned} u_{1,t} &= y_{1,t} - \beta_{12}y_{2,t} - \gamma_{10} \\ u_{2,t} &= -\beta_{21}y_{1,t} + y_{2,1} - \gamma_{20} + \gamma_{21}x_{1,t} \\ \mathbf{u}_t &= \mathbf{B}\mathbf{y}_t + \mathbf{\Gamma}\mathbf{x}_t \end{aligned} \tag{2}$$

where

$$\mathbf{B} = \begin{bmatrix} 1 & -\beta_{12} \\ -\beta_{21} & 1 \end{bmatrix}, \mathbf{\Gamma} = \begin{bmatrix} -\gamma_{10} & 0 \\ -\gamma 20 & -\gamma_{21} \end{bmatrix}, \mathbf{x}_t = \begin{bmatrix} 1 \\ x_{1,t} \end{bmatrix}$$

rank condition: check that $rk \begin{bmatrix} B_i & \Gamma_i \end{bmatrix} = n - 1 = 1$

equation 1: $B_i = []$ (i.e. empty) as both endogenous variables are included

$\Gamma_i = \begin{bmatrix} 0 \\ \gamma_{21} \end{bmatrix}$ as $x_{1,t}$ is excluded from the first eqn

hence $rk \begin{bmatrix} B_i & \Gamma_i \end{bmatrix} = rk(\Gamma_i) = 1$, provided $\gamma_{21} \neq 0$ $\quad\Rightarrow$ eq. 1 is *identified*

equation 2: $B_i = \Gamma_i = []$ $\quad\Rightarrow$ eqaution is not identified

order condition: $n = 2 \Rightarrow n - 1 = 1$ so we need at least 1 exclusion per equation for the order condition to hold

equation 1: we exclude $x_{1,t}$, note $\mathbf{\Gamma}(1,2) = 0$, so there is one exclusion in this equation $\Rightarrow$ eq. 1 is *just identified*

equation 2: there are no exclusion restrictions, the equation is *not identified*

Question b) Derive reduced form equations for $y_{1,t}$, $y_{2,t}$.

Reduced Form (RF) refers to an arrangement of the system such that the endogenous variables are expressed in terms of the exogenous variables and some relevant errors, $\mathbf{v}_t$. The best way to find the RF is to solve the matrix expression of the Structural Form (SF), given in (2), in terms of the endogenous variables:

$$\begin{aligned} \mathbf{u}_t &= \mathbf{B}\mathbf{y}_t + \mathbf{\Gamma}\mathbf{x}_t \\ \mathbf{y}_t &= -\mathbf{B}^{-1}\mathbf{\Gamma}\mathbf{x}_t + \mathbf{B}^{-1}\mathbf{u}_t \\ &= \mathbf{\Pi}\mathbf{x}_t + \mathbf{v}_t \end{aligned} \tag{3}$$

where it is useful to note for later that

$$\mathbf{\Pi} = \begin{bmatrix} \gamma_{10} + \beta_{12}\gamma_{20} & \beta_{12}\gamma_{21} \\ \gamma_{20} + \beta_{21}\gamma_{10} & \gamma_{21} \end{bmatrix} \frac{1}{1 - \beta_{12}\beta_{21}} \tag{4}$$

Question c) Show that the ILS and 2SLS estimators for β_{12} coincide

ILS Run OLS on the reduced form equations of the system and recover (identified) the parameters of interest using restrictions that you have placed on the relationship between the coefficients of the RF and SF models.

2SLS Split the endogenous variables on the r.h.s of equation i into part that lives in $Col(X)$ and part in $Col^{\perp}(X)$. Re-express the r.h.s. of each equation in terms of a set of regressors $\hat{Z} \in Col(X)$ and a residual, $w \in Col^{\perp}(X)$. The 2SLS estimator is standard OLS estimator associated with this second stage regression, $(\hat{Z}'\hat{Z})^{-1}\hat{Z}'y_i$ Where y_i is a $T \times 1$ vector of observations on the i^{th} endogenousvariable.

The construction of $\hat{\beta}_{12}^{ILS}$:

From (4) we can solve for

$$\hat{\beta}_{12}^{ILS} = \frac{\hat{\Pi}_{12}}{\hat{\Pi}_{22}}$$

Now recall from partitioned regression, Week 1, any OLS fit can be considered a two part procedure. As in week 1, let X_1 and X_2 denote two disjoint subsets of the (exogenous) regressors from the general regression $y = \begin{bmatrix} X_1 & X_2 \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} + u$. Then we can always express $\hat{\beta}_1$ as the coefficient from the regression of the part of y orthogonal to X_2 on the part of X_1 orthogonal to X_2:

$$\begin{aligned} \hat{\beta}_1 &= (X'_{1\perp 2}X_{1\perp 2})^{-1}X'_{1\perp 2}y_{\perp 2} \\ \text{where } X_{1\perp 2} &= [I - P_{X_2}]X_1 \\ y_{\perp 2} &= [I - P_{X_2}]y \end{aligned} \tag{5}$$

Each equation of the reduced form system (3) is just an OLS regression of y_i on $\mathbf{x}$, therefore we can find $\hat{\Pi}_{i2}$ by considering the partitioned regression formulation of equation i of our reduced form system. Let y_i denote a $T \times 1$ stack of observations on $y_{i,t}$ and $\mathbf{x}$ denote the $T \times 1$ stack of observations on $x_{1,t}$ then we can restsate each line of 3:

$$\begin{aligned} y_i &= \begin{bmatrix} \iota & \mathbf{x} \end{bmatrix} \begin{bmatrix} \Pi_{i1} \\ \Pi_{i2} \end{bmatrix} \\ &= \begin{bmatrix} \mathbf{x} & \iota \end{bmatrix} \begin{bmatrix} \Pi_{i2} \\ \Pi_{i1} \end{bmatrix} \end{aligned} \tag{6}$$

where, as in the introduction, ι denotes a $T \times 1$ stack of ones and the second equality is demonstrated to make the notation more obviously consistent with the notes on partitioned regression, and in P.A. Ruud, Ch3.

When $X_2 = \iota$ and $X_1 = \mathbf{x}$,

$$\begin{aligned} P_{X_2}X_1 &= \iota(\iota'\iota)^{-1}\iota'\mathbf{x} \\ &= \begin{bmatrix} \bar{x_1} \\ \vdots \\ \bar{x_1} \end{bmatrix} \\ (I - P_{X_2})X_1 &= \begin{bmatrix} x_{1,1} - \bar{x_1} \\ \vdots \\ x_{1,T} - \bar{x_1} \end{bmatrix} \end{aligned}$$

and similarly

$$(I - P_{X_2})y_i = \begin{bmatrix} y_{i,1} - \bar{y}_i \\ \vdots \\ y_{i,T} - \bar{y}_i \end{bmatrix} \tag{7}$$

Deriving (7) is left as an exercise (in using the dot-product). Substituting (7) into (6) and (6) into (5) we can write

$$\begin{aligned} \hat{\Pi}_{12} &= ((x_{1,t} - \bar{x_1})'(x_{1,t} - \bar{x_1}))^{-1}(x_{1,t} - \bar{x_1})'(y_{1,t} - \bar{y}) \\ \hat{\Pi}_{22} &= ((x_{1,t} - \bar{x_1})'(x_{1,t} - \bar{x_1}))^{-1}(x_{1,t} - \bar{x_1})'(y_{2,t} - \bar{y}) \end{aligned}$$

where $(x_{1,t} - \bar{x_1})$ etc denote the $T \times 1$ vectors from (7). Hence,

$$\begin{aligned} \hat{\beta}_{12}^{ILS} &= \frac{\hat{\Pi}_{12}}{\hat{\Pi}_{22}} \\ &= \frac{\sum_{t=1}^{T}(y_{1,t} - \bar{y_1})}{\sum_{t=1}^{T}(y_{2,t} - \bar{y_2})} \end{aligned} \tag{8}$$

The 2 Stage Least Squares solution:

In Walter's notation, p.14, the 2SLS procedure requires

$$\begin{aligned} \hat{Y}_1 &= P_X Y_1 \\ \hat{Z} &= \begin{bmatrix} \hat{Y}_1 & X_1 \end{bmatrix} \\ y_{1,t} &= \hat{Z}\delta_{2SLS} + w_t \\ \hat{\delta}_{2SLS} &= (\hat{Z}'\hat{Z})^{-1}\hat{Z}y_1 \end{aligned} \tag{9}$$

where we consider the first equation of an arbitrary system and; the first line is the regession of the endogenous right hand side variables in the first equation on *all* the exogenous variables in the sysem; the second line concatenates (puts next to eachother in a matrix) the part of $Y_1 \in Col(X)$ and the exogenous variables *appearing in equation 1*; the third line is the second stage regression equation and; the final line defines the 2SLS estimator.

In terms of the system (1) considered in this question, the definitions in (9) relate to the following objects:

$$\begin{aligned} X &= \begin{bmatrix} x_1 & \iota \end{bmatrix} \\ Y_1 = y_2 &\Rightarrow \hat{Y_1} = \hat{y_2} = P_X y_2 \\ X_1 = \iota &\Rightarrow \hat{Z} = \begin{bmatrix} \hat{y_2} & \iota \end{bmatrix} \end{aligned}$$

Notice that the last line of (9) implies

$$\hat{\delta}_{2SLS} = (\begin{bmatrix} \hat{y_2} & \iota \end{bmatrix}' \begin{bmatrix} \hat{y_2} & \iota \end{bmatrix})^{-1} \begin{bmatrix} \hat{y_2} & \iota \end{bmatrix}' y_1$$

Also notice both X and $\hat{Z}$ have the same structure that we encountered in the above derivation of the ILS estimator, i.e. they imply regression on a constant and a single other variable. We can use this to derive an expression for $\hat{y_2}$ and from there we can find $\hat{\delta}_{2SLS}$.

Consider the first regression of our 2-stage least squares process:

$$\begin{aligned} \hat{y_2} &= X\hat{\beta}_{y_2} \\ \hat{y_2} &= \begin{bmatrix} x_1 & \iota \end{bmatrix} \hat{\beta}_{y_2} \\ \hat{y_2} &= x_1\hat{\beta}_{y_2,1} + \iota\hat{\beta}_{y_2,2} \\ (\iota'\iota)^{-1}\iota'\hat{y_2} &= (\iota'\iota)^{-1}\iota' x_1\hat{\beta}_{y_2,1} + \hat{\beta}_{y_2,2} \\ \hat{\beta}_{y_2,2} &= \frac{1}{T}\sum_t \hat{y_{2,t}} - \frac{1}{T}\sum_t x_{1,t} \times \hat{\beta}_{y_2,1} \\ \hat{\beta}_{y_2,2} &= \bar{y_2} - \bar{x}_1\hat{\beta}_{y_2,1} \end{aligned} \tag{10}$$

where, inuitively we are just saying that the intercept (i.e. the estimator $\hat{\beta}_{y_2,2}$ that multiplies the vector of ones in line 3 of (10)) is equal to the mean of y_2 less the mean effect of the regressor, $\bar{x}_1\hat{\beta}_{y_2,1}$. (We saw this in the partitioned regression example too).

Therefore:

$$\begin{aligned}
\hat{y_2} &= x_1\hat{\beta}_{y_2,1} + \iota(\bar{y_2} - \bar{x_1}\hat{\beta}_{y_2,1}) \\
&= \iota\bar{y_2} + (x_1 - \iota\bar{x})\hat{\beta}_{y_2,1} \\
\text{or} \quad \hat{y_{2,t}} &= \bar{y_2} - (x_{1,t} - \bar{x_1})\hat{\beta}_{y_2,1} \qquad (11)
\end{aligned}$$

where we substitue the last line of (10) into the third line to derive our expression for $\hat{y_{2,t}}$. Also note, by the argument in the ILS section above, that $\hat{\beta}_{y_2,1}$ is just the usual OLS estimator, constructed using the de-meaned data (partitioned reg argument again!):

$$\hat{\beta}_{y_2,1} = \frac{\sum_t (x_{1,t} - \bar{x_1})(y_{2,t} - \bar{y_2})}{\sum_t (x_{1,t} - \bar{x_1})^2}$$

Recall from above that $\hat{Z} = \begin{bmatrix} \hat{y_2} & \iota \end{bmatrix}$ so we can now use the penultimate line of (9) together with our definition of $\hat{y_2}$ and what we know about regressions against a variable and a constant to solve for our 2SLS estimate of the effect of y_2 on y_1.

Consider the second-stage regression:

$$y_1 = \begin{bmatrix} \hat{y_2} & \iota \end{bmatrix} \begin{bmatrix} \delta_1 \\ \delta_2 \end{bmatrix} + w_1$$

Clearly, $\hat{\delta_1}$ is our 2SLS estimate of the effect of y_2 on y_1. We want to see if this is the same as our ILS estimate of the same thing, i.e. $\hat{\beta}_{12}^{ILS}$. Once again using the partitioned regression argument, we can write:

$$\hat{\delta_1} = \frac{\sum_t (y_{1,t} - \bar{y_1})(\hat{y_{2,t}} - \bar{\hat{y_2}})}{\sum_t (\hat{y_{2,t}} - \bar{\hat{y_2}})^2}$$

Using the Law of Iterated Expectations and the third line of (11) we re-write

$$
\begin{aligned}
\hat{y_{2,t}} - \bar{\hat{y_2}} &= \hat{y_{2,t}} - \bar{y_2} \\
&= -(x_{1,t} - \bar{x_1})\hat{\beta}_{y_2,1}
\end{aligned}
$$

$$
\begin{aligned}
\Rightarrow \hat{\delta}_1 &= \frac{\hat{\beta}_{y_2,1}\sum_t(y_{1,t} - \bar{y_1})(x_{1,t} - \bar{x_1})}{\hat{\beta}^2_{y_2,1}\sum_t(x_{1,t} - \bar{x_1})^2} \\
&= \frac{\sum_t(y_{1,t} - \bar{y_1})(x_{1,t} - \bar{x_1})}{\frac{\sum_t(x_{1,t}-\bar{x_1})(y_{2,t}-\bar{y_2})}{\sum_t(x_{1,t}-\bar{x_1})^2}\sum_t(x_{1,t} - \bar{x_1})^2} \\
&= \frac{\sum_t(y_{1,t} - \bar{y_1})(x_{1,t} - \bar{x_1})}{\sum_t(x_{1,t} - \bar{x_1})(y_{2,t} - \bar{y_2})} \\
&= \frac{\sum_t(y_{1,t} - \bar{y_1})}{\sum_t(y_{2,t} - \bar{y_2})} \\
&= \hat{\beta}^{ILS}_{12}
\end{aligned}
$$

so we are done! While this is all good practice, the most important thing to remember about 2SLS is that it deals with the edogeneity problem by spliting the endogenous variables on the r.h.s of equation i into a part linearly dependent on the *exogenous* variables X, and therefore uncorrelated with the (unobserved) errors $u_{i,t}$ and a part which is correlated with the (unobserved) errors and so is pushed into the 2SLS regression residual $w_{i,t}$. Thus estimates $\hat{\delta}_i$ are consistent.