Lecture 3: System Estimation by IV

January 24, 2025

Motivation Example: Simultaneous Equations

- As before, write the population model as

$$y_1 = x_1\beta_1 + u_1 \tag{1}$$
$$y_2 = x_2\beta_2 + u_2 \tag{2}$$
$$\vdots$$
$$y_G = x_G\beta_G + u_G \tag{3}$$

where y_g is a response variable, $g = 1, \ldots, G$. The explanatory variables, x_g, can be different across equations.

- Now we want to allow $E(x_g' u_g) \neq 0$ for at least some equations g.

- **EXAMPLE:** Individual Labor Supply. Consider a labor supply function and a wage offer (inverse labor demand) function:

$$h^s(w) = \gamma_1 w + z_1\delta_1 + u_1 \tag{4}$$
$$w^o(h) = \gamma_2 h + z_2\delta_2 + u_2 \tag{5}$$

Equation 4 shows how much each unit in the population would work at any given wage, w. Once we hold fixed z_1 and u_1, we can trace out the (linear) supply curve as a function of w.

Motivation Example: Simultaneous Equations

- We bring in the wage offer function to recognize that, for retrospective data, we observe the pair (h_i, w_i), with w_i not being randomly assigned. What is a sensible assumption about how (h_i, w_i) are generated? A standard approach is to assume that we observe equilibrium hours and wages for each individual. That is, the data are generated as

$$h_i = \gamma_1 w_i + z_{i1}\delta_1 + u_{i1} \tag{6}$$

$$w_i = \gamma_2 h_i + z_{i2}\delta_2 + u_{i2} \tag{7}$$

Identification

- General two-equation structural system (in the population):

$$y_1 = \gamma_1 y_2 + z_1 \delta_1 + u_1 \tag{8}$$
$$y_2 = \gamma_2 y_1 + z_2 \delta_2 + u_2 \tag{9}$$

  where z_1 is $1 \times M_1$ and z_2 is $1 \times M_2$. Let z be $1 \times M$ contain all (nonredundant) exogenous variables

$$E(z'u_1) = E(z'u_2) = 0$$

  where in almost all applications z_1 and z_2 (and therefore z) include unity. We act as if that is true here, so that the structural errors u_1 and u_2 have zero means.

- $\gamma_1, \delta_1, \gamma_2, \delta_2$ are the structural parameters.

- The moment conditions imply that if a variable is exogenous in any equation, it is exogenous in all equations; this is the traditional starting point.

- Although we do not need them to study identification, we can obtain reduced forms for y_1 and y_2 if $\gamma_1 \gamma_2 \neq 1$.

Identification

- Generally, a reduced form expresses an endogenous variable as a function of exogenous variables and unobserved errors.
- In this case, solve the two equations for y_1 and y_2:

$$\begin{aligned} y_1 &= \gamma_1(\gamma_2 y_1 + z_2\delta_2 + u_2) + z_1\delta_1 + u_1 \\ &= \gamma_1\gamma_2 y_1 + z_1\delta_1 + z_2\gamma_1\delta_2 + u_1 + \gamma_1 u_2. \end{aligned}$$

- Therefore, if $\gamma_1\gamma_2 \neq 1$,

$$\begin{aligned} y_1 &= (1-\gamma_1\gamma_2)^{-1}(z_1\delta_1 + z_2\gamma_2\delta_2 + u_1 + \gamma_1 u_2) \\ &= \mathbf{z}\pi_1 + v_1 \end{aligned}$$

  where π_1 is the $M \times 1$ vector of reduced form parameters and $v_1 = (1-\gamma_1\gamma_2)^{-1}(u_1 + \gamma_1 u_2)$ is a reduced form error.
- We can do the same for y_2, so we have

$$\begin{aligned} y_1 &= \mathbf{z}\pi_1 + v_1 \\ y_2 &= \mathbf{z}\pi_2 + v_2. \end{aligned}$$

  Both reduced form errors satisfy $E(\mathbf{z}'v_1) = E(\mathbf{z}'v_2) = 0$, which means π_1 and π_2 can be consistently estimated by OLS.

Identification

- We can always consistently estimate the RF (Reduced Form) parameters. When are the structural parameters identified?
- Identification in the two-equation case is straightforward. Consider identification of the first structural equation. Write it with the RF of y_2:

$$
\begin{aligned}
y_1 &= \gamma_1 y_2 + z_1 \delta_1 + u_1 \\
y_2 &= \mathbf{z} \pi_2 + v_2
\end{aligned}
$$

- Because y_2 is the only endogenous explanatory variable, we need at least one instrument for it. That means we must have something in $\mathbf{z}$ in the RF with a nonzero coefficient that is not also in z_1.
- But $y_2 = \gamma_2 y_1 + z_2 \delta_2 + u_2$ and so π_2 has a nonzero coefficient on something not in z_1 if and only if there is at least one element of z_2 that is not also in z_1 with nonzero coefficient in δ_2.

Identification

- So, we can read identification of each equation off of the structural system:

$$y_1 = \gamma_1 y_2 + z_1\delta_1 + u_1$$
$$y_2 = \gamma_2 y_1 + z_2\delta_2 + u_2$$

  The first equation is identified if and only if there is at least one element in z_2 not in z_1 with a nonzero coefficient (element of δ_2) in the second equation. Similarly, the second equation is identified if and only if there is something in z_1 not in z_2 with a corresponding nonzero element in δ_1.

Identification

In the system

$$h = \gamma_1 w + \delta_1 \text{exper} + \delta_2 \text{exper}^2 + \delta_3 \text{othinc} + \delta_4 \text{kids} + u_1$$
$$w = \gamma_2 h + \delta_1 \text{exper} + \delta_2 \text{exper}^2 + \delta_3 \text{educ} + u_2,$$

the labor supply function is identified if and only if $\delta_3 \neq 0$. The wage offer function is identified if and only if at least one of δ_3 and δ_4 is different from zero.

- Important: Our imposing of exclusion restrictions means that it must be the case that *educ* is legitimately excluded from the supply equation and *othinc* and *kids* are properly excluded from the wage offer equation.

Example: Log-linear Supply-Demand System (Zoutman et al., 2018)

Assume log-linear supply-demand system.

$$y_{it}^S = \varepsilon^S p_{it} + \eta z_{it} + \Gamma^S x_{it} + v_{it}^S \quad \text{(supply)}$$
$$y_{it}^D = \varepsilon^D p_{it} + \gamma z_{it} + \Gamma^D x_{it} + v_{it}^D \quad \text{(demand)}$$

Can similarly write down reduced-form.

$$\begin{pmatrix} y_{it} \\ p_{it} \end{pmatrix} = \begin{pmatrix} \pi_{zy} \\ \pi_{zp} \end{pmatrix} z_{it} + \begin{pmatrix} \pi_{xy} \\ \pi_{xp} \end{pmatrix} x_{it} + \xi_{it}.$$

Without further restrictions.

$$\underbrace{\begin{pmatrix} \pi_{zy} \\ \pi_{zp} \end{pmatrix}}_{\substack{\text{Two} \\ \text{reduced-form} \\ \text{estimates}}} = \underbrace{\begin{pmatrix} \frac{\varepsilon^S \gamma - \varepsilon^D \eta}{\varepsilon^S - \varepsilon^D} \\ \frac{\gamma - \eta}{\varepsilon^S - \varepsilon^D} \end{pmatrix}}_{\substack{\text{Four unknown} \\ \text{structural} \\ \text{parameters}}} \qquad \textit{NOT identified.}$$

Example: Log-linear Supply-Demand System (Zoutman et al., 2018)

What additional exclusion restriction do we need?

- Interestingly, if z_{it} happens to represent an "ad valorem" tax that enters log price linearly, then it works!
- SER (Standard Exclusion Restriction): $\eta = 0$ i.e. tax is levied on demand side.
- RER (Ramsey Exclusion Restriction) $z_{it} = log(1 + \tau_{it})$, $\gamma = \varepsilon^{D}$ i.e. demand only depends on price after tax.

Example: Log-linear Supply-Demand System (Zoutman et al., 2018)

With these assumptions, we can write:

$$\begin{pmatrix} \pi_{zy} \\ \pi_{zp} \end{pmatrix} = \begin{pmatrix} \frac{\varepsilon^S \varepsilon^D}{\varepsilon^S - \varepsilon^D} \\ \frac{\varepsilon^D}{\varepsilon^S - \varepsilon^D} \end{pmatrix} \quad \text{Identified}$$

Obviously, we will still need **the relevance condition for IVs**:

- $\pi_{zp} \neq 0$ (Variation in tax rate does affect pre-tax price, o.w. entire incidence is on demand side.)
- $\pi_{zp} \neq -1$ (After-tax price is not independent of tax rate, o.w. entire incidence is on supply side.)

Estimation follows standard 2SLS.

System IV Estimator

- As before, the SUR and panel data cases can both be written as

$$Y_i = X_i\beta + u_i \tag{10}$$

where Y_i is $G \times 1$ (or $T \times 1$) and X_i is $G \times K$ (or $T \times K$), and β is the $K \times 1$ vector of parameters to be estimated.

- How do we choose the matrix of instruments, Z_i?
  - SUR: Assume that for each equation g, the moment conditions are

$$E(z'_{ig}u_{ig}) = 0, \tag{11}$$

  for a $1 \times L_g$ vector z_{ig}, written for a random draw.

  - Then the $G \times L$ matrix of instruments is

$$Z_i = \begin{pmatrix} z_{i1} & 0 & \cdots & 0 \\ 0 & z_{i2} & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & z_{iG} \end{pmatrix}, \tag{12}$$

  where $L = L_1 + L_2 + \cdots + L_G$.

System IV Estimator

- Write the system as in the system OLS/GLS case:

$$Y_i = X_i\beta + u_i.$$

- Assumption SIV.1 (Moment Conditions): For a $G \times L$ matrix Z_i,

$$E(Z_i'u_i) = 0.$$

- The assumption is weaker, often in important ways, than the assumption that all elements of Z_i are uncorrelated with all elements in u_i: $E(Z_i \otimes u_i) = 0$. We return to this distinction later.

- Assumption SIV.2 (Rank Condition):

$$\text{rank } E(Z_i'X_i) = K.$$

- Later it is useful to define $C = E(Z_i'X_i)$.

System IV Estimator

- In the SUR case

$$E(z_i'x_i) = \begin{pmatrix} E(z_{i1}'x_{i1}) & 0 & \cdots & 0 \\ 0 & E(z_{i2}'x_{i2}) & \cdots & 0 \\ 0 & 0 & \ddots & \vdots \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \cdots & E(z_{iG}'x_{iG}) \end{pmatrix},$$

  and so SIV.2 holds if and only if

$$\text{rank } E(z_{ig}'x_{ig}) = K_g, \quad g = 1, \ldots, G.$$

- Suppose $L = K$. Then

$$\beta = [E(Z_i'X_i)]^{-1} E(Z_i'y_i).$$

- Now, replace population averages with sample averages to get the system instrumental variables (SIV) estimator:

$$\hat{\beta}_{SIV} = \left(N^{-1} \sum_{i=1}^{N} Z_i'X_i \right)^{-1} \left(N^{-1} \sum_{i=1}^{N} Z_i'y_i \right).$$

Empirical Example: Epple and McCallum (2005)

- Edible meat of young chicken (broiler)

$$\text{(supply)} \qquad Q_t = \alpha_0 + \alpha_1 P_t + \alpha_2 W_t + u_t \tag{13}$$

$$\text{(demand)} \qquad Q_t = \beta_0 + \beta_1 P_t + \beta_2 Y_t + v_t. \tag{14}$$

As suggested above, we presume that $\alpha_1 > 0$, $\alpha_2 < 0$, $\beta_1 < 0$, and $\beta_2 > 0$. In 13 and 14, u_t and v_t are stochastic disturbances representing measurement error, a multitude of individually-unimportant omitted variables, and purely random influences. We assume that $E[u_t] = 0$, $E[v_t] = 0$, $E[u_t^2] = \sigma_u^2$, and $E[v_t^2] = \sigma_v^2$ for all $t = 1, 2, \ldots, T$. We also assume that W_t and Y_t can legitimately be treated as exogenous to the particular market under consideration, so that W_t and Y_t will be uncorrelated with values of u_t and v_t for all current and past periods.

- Time series data from 1960 – 1999
  - Fast technological progress
  - Large change of relative price of chicken to other types of meat

Alternative Estimates of Demand (per capita)

- OLS

$$
\begin{gathered}
q = -4.860 + 0.871y - 0.277p \\
(0.669) \quad (0.068) \quad (0.070) \\
R^2 = 0.980 \quad SE = 0.0572 \quad DW = 0.343 \quad T = 52
\end{gathered}
$$

- Add price of substitute good (beef)

$$
\begin{gathered}
q = -4.679 + 0.852y - 0.264p - 0.118p_b \\
(0.675) \quad (0.069) \quad (0.070) \quad (0.084) \\
R^2 = 0.981 \quad SE = 0.0566 \quad DW = 0.443 \quad T = 52
\end{gathered}
$$

- Benchmark model: first differencing, no constant

$$
\begin{gathered}
\Delta q = 0.711\Delta y - 0.374\Delta p + 0.251\Delta p_b \\
(0.150) \quad (0.058) \quad (0.068) \\
R^2 = 0.331 \quad SE = 0.0294 \quad DW = 2.38 \quad T = 51
\end{gathered}
$$

Alternative Estimates of Supply (aggregate)

- Key exclusion restriction: feed price/corn price
- Simple specification does not work well

$$q^A = -9.185 - 1.203p - 0.338p_{cor}$$
$$(0.029) \quad (0.110) \quad (0.075)$$
$$R^2 = 0.942 \quad SE = 0.1412 \quad DW = 0.591 \quad T = 52$$

- Tech progress – time trend, control for autocorrelation, more promising

$$q^A = -2.478 - 0.041p - 0.083p_f + 0.0102\text{time} + 0.647q^A(-1)$$
$$(0.698) \quad (0.052) \quad (0.032) \quad (0.0038) \quad (0.108)$$
$$R^2 = 0.997 \quad SE = 0.0252 \quad DW = 1.883 \quad T = 39.$$

Identification

- Demand: excluded variables on supply side are feed price, lagged quantity.
- Supply: excluded variables on demand side are changes in beef price and income per capita.
- Also include population change , lagged outputs price, since demand equation in first difference.

Results

$$\Delta q = 0.843\Delta y - 0.404\Delta p + 0.279\Delta p_b$$
$$(0.143) \quad (0.086) \quad (0.093)$$
$$R^2 = 0.291 \quad SE = 0.0253 \quad DW = 1.929 \quad T = 40$$

$$q^A = -2.371 + 0.105p - 0.113p_f + 0.0123\text{time} + 0.640q^A(-1)$$
$$(0.773) \quad (0.077) \quad (0.037) \quad (0.0043) \quad (0.119)$$
$$R^2 = 0.996 \quad SE = 0.0279 \quad DW = 1.869 \quad T = 40$$

Results

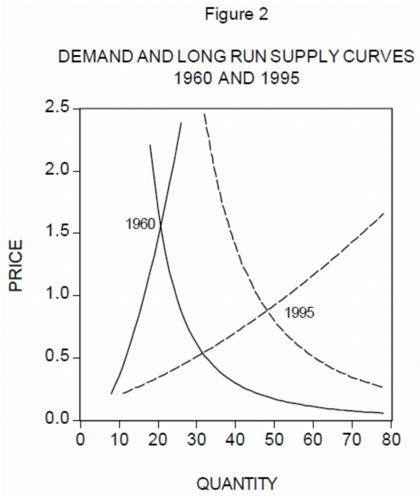

Figure: Epple and McCallum (2005) Figure 2