14.383

Fall 2003.

Handout 1 ½

(- OLS on SF and RF.
- The Identification Problem
- The identifying restrictions, with IV interpretation).

T.A.: Alexis León
E52 - 351
aleon@mit.edu
(Office hours: Tuesdays at 4pm in E52 - 398).

A. Brief review of SF vs. RF:

Consider a system of M-equations in M- jointly endog. variables, and K- pre-determined variables:

$$\underset{(T\times M)}{Y}\ \underset{(M\times M)}{B} + \underset{(T\times K)}{Z}\cdot\underset{(K\times M)}{\Gamma} = \underset{(T\times M)}{U}$$

(eg: $\underbrace{[y_{1t}\ y_{2t}]}_{y_t'}\begin{bmatrix}\beta_{11} & \beta_{21}\\ \beta_{12} & \beta_{22}\end{bmatrix} + \underbrace{[z_{1t}\ z_{2t}\ z_{3t}]}_{z_t'}\begin{bmatrix}\gamma_{11} & \gamma_{21}\\ \gamma_{12} & \gamma_{22}\\ \gamma_{13} & \gamma_{23}\end{bmatrix} = \underbrace{[u_{1t}\ u_{2t}]}_{u_t'}$ for $t=1,\dots,T$).

[→ Note: Equation "i" in this system will be:
$Y B_{(i)} + Z\Gamma_{(i)} = U_{(i)}$]

From the SF above, we can always derive the RF: (just solve for the Y's in terms of the Z's):

$$\Rightarrow\quad Y = -Z\Gamma B^{-1} + UB^{-1} = \underset{(T\times K)}{Z}\ \underset{(K\times M)}{\Pi} + \underset{(T\times M)}{V},$$

where: $\begin{cases} -\Gamma B^{-1} \equiv \Pi & (\Leftrightarrow \underline{\Pi B + \Gamma = 0}) \\ UB^{-1} \equiv V & (\Leftrightarrow \underset{(1\times M)}{u_t'}\underset{(M\times M)}{B^{-1}} = \underset{(1\times M)}{v_t'}). \end{cases}$

- If we assume: $\{u_t\}$ iid, with:

$$\underset{(1\times M)}{E[u_t']} = E[u_t'\,|\,z_t] = 0$$

$$\underset{(M\times M)}{Var[u_t']} = E[u_t u_t'] = E[u_t u_t'\,|\,z_t] = \begin{bmatrix} E[u_{1t}^2|z_t] & E[u_{1t}u_{2t}|z_t] & \dots & E[u_{1t}u_{Mt}|z_t] \\ \text{≶} & E[u_{2t}^2|z_t] & \dots & E[u_{2t}u_{Mt}|z_t] \\ & & \ddots & \vdots \\ & & & E[u_{Mt}^2|z_t] \end{bmatrix}$$

$$= \begin{bmatrix} \sigma_{11} & \sigma_{12} & \dots & \sigma_{1M} \\ & \sigma_{22} & \dots & \sigma_{2M} \\ & & \ddots & \vdots \\ & & & \sigma_{MM} \end{bmatrix} \equiv \underset{(M\times M)}{\Sigma}$$

← matrix of var./covar. b/w. the disturbances across diff't equations! (not across t !!!).

so:

$$\text{"}\underset{(T\times M)}{E[U]}\text{"} = \{\underset{(T\times 1)}{E[U_{(i)}]}\}_{i=1,\dots,M} =$$

$$= [\underset{(T\times 1)}{E[U_{(1)}]}\ \ \underset{(T\times 1)}{E[U_{(2)}]}\ \dots\ \underset{(T\times 1)}{E[U_{(M)}]}] = [\underset{(T\times 1)}{\vec{0}} : \underset{(T\times 1)}{\vec{0}} : \dots : \underset{(T\times 1)}{\vec{0}}] = \underset{(T\times M)}{0}$$

$$\text{"}\underset{(T\times M)}{Var[U]}\text{"} = \{\underset{(T\times 1)}{Var[U_{(i)}]}\}_{i=1,\dots,M} =$$

$$= \begin{bmatrix} E[\underset{(T\times 1)}{U_{(1)}}\underset{(1\times T)}{U_{(1)}'}] & E[\underset{(T\times 1)}{U_{(1)}}\underset{(1\times T)}{U_{(2)}'}] & \dots & E[\underset{(T\times 1)}{U_{(1)}}\underset{(1\times T)}{U_{(M)}'}] \\ & E[\underset{(T\times 1)}{U_{(2)}}\underset{(1\times T)}{U_{(2)}'}] & \dots & E[\underset{(T\times 1)}{U_{(2)}}\underset{(1\times T)}{U_{(M)}'}] \\ & & \ddots & \vdots \\ & & & E[U_{(M)}U_{(M)}'] \end{bmatrix} = \begin{bmatrix} \sigma_{11}I_T & \sigma_{12}I_T & \dots & \sigma_{1M}I_T \\ & \sigma_{22}I_T & \dots & \sigma_{2M}I_T \\ & & \ddots & \vdots \\ & & & \sigma_{MM}I_T \end{bmatrix}$$

$$= \underbrace{\Sigma\otimes I_T}_{(MT\times MT)}$$

then we'll have:

$$\underset{(1\times M)}{E[v_t']} = E[v_t'/z_t] = E[\underset{(1\times M)(M\times M)}{u_t' B^{-1}}/z_t] = E[u_t'/z_t]\cdot B^{-1} = \underline{\underline{0_{(1\times M)}}}$$

$$Var[v_t'] = E[v_t v_t'] = E[v_t v_t'/z_t] = E[(B^{-1})' u_t u_t' B^{-1} / z_t] =$$

$$= B^{-1\prime}\cdot E[u_t u_t'/z_t]\cdot B^{-1} =$$

$$= \underset{(M\times M)}{B^{-1\prime}}\cdot \underset{(M\times M)}{\Sigma}\cdot \underset{(M\times M)}{B^{-1}} \equiv \underline{\underset{(M\times M)}{\Omega}} \left(= \begin{bmatrix} \omega_{11} & \cdots & \omega_{1M} \\ & \ddots & \vdots \\ & & \omega_{MM} \end{bmatrix} \right).$$

$$(\Leftrightarrow \underline{\Omega B = B^{-1\prime}\Sigma}).$$

$$\left(\begin{array}{l} \text{so:} \\ \text{"}\underset{(T\times M)}{E[V]}\text{"} = \left[E[V_{(1)}] \dots E[V_{(M)}] \right] = [\vec{0} \vdots \dots \vdots \vec{0}] = \underset{(T\times M)}{0} \\ \text{"}\underset{(T\times M)}{Var[V]}\text{"} = \begin{bmatrix} E[V_{(1)}V_{(1)}'] & \dots & E[V_{(1)}V_{(M)}'] \\ & \ddots & \vdots \\ & & E[V_{(M)}V_{(M)}'] \end{bmatrix} = \begin{bmatrix} \omega_{11} I_T & \dots & \omega_{1M} I_T \\ & \ddots & \vdots \\ & & \omega_{MM} I_T \end{bmatrix} = \Omega \otimes I_T \\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad (= B^{-1\prime}\Sigma B^{-1} \otimes I_T). \end{array} \right)$$

. Last week we saw that we can <u>always</u> estimate the <u>RF</u> by <u>OLS</u>.
Why? Because we can check for the orthogonality condition in each RF equation "i": $\underset{(T\times 1)}{y_i} = \underset{(T\times K)(K\times 1)}{Z\Pi_{(i)}} + \underset{(T\times 1)}{V_{(i)}}$:

$$\text{plim } \frac{1}{T} \underset{(K\times T)(T\times 1)}{Z' V_{(i)}} = \text{plim } \frac{1}{T}\sum_{t=1}^{T} \underset{(K\times 1)\ (1\times 1)}{z_t \cdot v_{(i)t}} = E[z_t \cdot v_{(i)t}] = E[z_t \cdot \underbrace{E[v_{(i)t}/z_t]}_{\vec{0}}] = 0$$

We also saw that estimation of the <u>SF</u> by <u>OLS</u> is <u>generally</u> inconsistent
Why? Because we can also check ~~for~~ that the orthog. cond. is not met:
Take equation 1, for example: $Y B_{(1)} + Z\Gamma_{(1)} = U_{(1)} \Leftrightarrow$

$$\Leftrightarrow \underset{(T\times 1)}{y_1} = \underset{(T\times(M-1))}{Y_1} \cdot \underset{(M-1)\times 1}{\tilde{B}_{(1)}} + \underset{(T\times K)(K\times 1)}{Z\Gamma_{(1)}} + \underset{(T\times 1)}{U_{(1)}}.$$

where: we imposed $\beta_{11} = 1$ (normalization restr.)
and $\tilde{B}_{(1)}$ is $B_{(1)}$ without the own element β_{11}
and Y_1 is Y without the own column y_1.

$$\text{plim } \frac{1}{T} \underset{(K\times T)(T\times 1)}{Z' U_{(1)}} = \text{plim } \frac{1}{T}\sum_{t=1}^{T} \underset{(K\times 1)(1\times 1)}{z_t \cdot u_{(1)t}} = E[z_t \cdot u_{(1)t}] = E[z_t\, E[u_{(1)t}/z_t]] = 0.$$

but:

$\text{plim } \frac{1}{T} Y_1' U_{(1)} \neq 0$ in general. Let us analyze this one ~~[illegible]~~ separately by regressor (y_2 through y_M):

for each $j = \{2, \dots, M\}$ (= ie: $j \neq 1$), we have:

$$\text{plim } \frac{1}{T} \underset{(1\times T)(T\times 1)}{y_j' U_{(1)}} \left(= \text{plim } \frac{1}{T}\sum_{t=1}^{T} y_{j_t} u_{(1)_t}\right) = \text{plim } \frac{1}{T}\big(Z\Pi_{(j)} + \underset{(1\times T)}{V_{(j)}}\big)' \underset{(T\times 1)}{U_{(1)}} =$$

$$= \text{plim } \frac{1}{T} \Pi_{(j)}' Z' \cancel{U_{(1)}} + \text{plim } \frac{1}{T} V_{(j)}' U_{(1)} =$$

$$= \text{plim } \frac{1}{T} \underbrace{\big(\underset{(T\times M)(M\times 1)}{U B^{-1}_{(j)}}\big)'}_{(1\times T)} \underset{(T\times 1)}{U_{(1)}} =$$

$$= (B^{-1}_{(j)})' \cdot \text{plim } \frac{1}{T} \underset{(M\times T)(T\times 1)}{U' U_{(1)}} =$$

$$= B^{-1\prime}_{(j)} \cdot E[U' U_{(i)}] =$$

$$= \underline{B^{-1\prime}_{(j)} \cdot \Sigma_{(1)}} \quad \left(= [B^{-1\prime}\Sigma]_{j1}\right).$$

$$= [B^{-1\prime}]_{j1} \cdot \sigma_{11} + \sum_{k\neq 1} [B^{-1\prime}]_{jk} \cdot \sigma_{1k} \quad \underline{\neq 0}, \text{ in } \underline{\text{general}}.$$

Note, however, that if the system is <u>recursive</u> (B triangular, Σ diagonal), then we'll have: $\left\{ \begin{array}{l} \sigma_{1k} = 0 \quad \forall k \neq 1 \\ [B^{-1\prime}]_{j1} = 0 \end{array} \right\}$ (*), so $\underline{= 0}$ for all "j", and hence ~~the residual will~~ all regressors in the SF will satisfy the orthog. condition, so OLS is consistent.

If, however, the above is true only for some j but not for others, then we say that we have <u>relative recursivity</u> (: equations j and 1 are <u>rel. rec.</u> <u>iff</u> $[B^{-1\prime}\Sigma]_{j1} = 0$, which implies that $\underline{\text{plim } \frac{1}{T} y_j' U_{(1)} = 0}$ and hence $\underline{y_j}$ <u>can be treated as pre-det. in equ. 1</u>).

But unless we have that all the included endog vars. (all the possible y_j's) in equation 1 can be treated as pre-det., OLS will yield inconsistent estimates of the SF parameters in equation 1.

(*) It turns out we get this result ($[B^{-1\prime}]_{j1} = 0$) when the equations $(j, 1)$ are <u>relatively triangular</u>, this is, if there is no way how a shock in equation 1 can get to affect equation j through the coefficients.

B. The Liu + Sims critique of Structural Estimation:

- OLS on RF is always possible, why not always use it?

In fact, Liu + Sims critique of the focus on SF argues that we should stick to RF and stay away from SF, since the only way to estimate SF parameters is to impose – as we'll see in a sec. – restrictions that are arbitrary, and RF will do just fine for prediction purposes.

But: In deciding what pre-det. variables to include in the model, and thus in the RF regression, we are already imposing arbitrary exclusion restrictions! If we were to really allow for any variable (incl. lag and lead) to enter the eqn., we'd run out of dof !!! We've got to draw the line somewhere, and in doing so, we're effectively incurring in the same problem as in estimating the SF: we're imposing restrictions!

· Note: It is always possible to use a linear transformation of the SF model such that we "triangularize" B and "diagonalize" Σ.

Why don't we do that and thus allow ourselves to always have a recursive system which we can estimate by OLS?

Because: the coeff. we'd be estimating are not the SF (or even the RF) coeffs. anymore, but a non-linear f^n of all of them, — and that's really very hard to interpret!

ie: $YB + Z\Gamma = U$, post-mult. by some P:

$$YBP + Z\Gamma P = UP$$

such that / BP: triangular

/ $\underbrace{P'\Sigma P}_{= Var(u_t P)}$: diagonal.

C. The identification problem:

We have seen it is always possible to derive the RF from the SF ~~form~~ — so we could always get the RF coeff. from the SF coeffs.
But is it possible to get the SF coeffs. from the RF coeffs.? (if we just know RF coeffs.).

The id. problem consists in deriving a unique solution for the (unknown) elements in the matrices of SF parameters (B, Γ, Σ) from our knowledge of: (given)

- the equations that relate the RF to the SF coeffs.:

$$\Pi = -\Gamma B^{-1} \iff \Pi B + \Gamma = 0$$
$$\iff [\Pi \;\; I]\begin{bmatrix} B \\ \Gamma \end{bmatrix} = 0$$

$$\Omega = B^{-1\prime} \Sigma \cdot B^{-1} \iff \Omega B = B^{-1\prime}\Sigma$$
$$\iff [\Omega \;\; 0]\begin{bmatrix} B \\ \Gamma \end{bmatrix} = B^{-1\prime}\Sigma$$

(← not really useful if we have no prior knowledge of Σ whatsoever! — In that case, # unknowns in Σ and # equations in Ω increase by the same # !).

- the (linear, or linearized, if necessary) restrictions on the SF parameters that come from a prior knowledge (econ. theory, hopefully...):

(ie: prior information on SF parameters — usu. exclusion or normalization restrictions:

$$GB + H\Gamma = \phi$$
$$(\iff \underbrace{[G \;\; H]}_{\Phi}\begin{bmatrix} B \\ \Gamma \end{bmatrix} = \phi$$

(if all restrictions are excl. restrictions, then ϕ will be full of zeroes!).

and (if we want to make assumptions on the relative recursiveness of some equs, ie, ~~on the~~ if we want to assume that some y's are pre-det. in some equs):

$$\psi \cdot B^{-1\prime}\Sigma = 0 \iff \psi\,\Omega B = 0 \iff [\psi\Omega \;\; 0]\begin{bmatrix} B \\ \Gamma \end{bmatrix} = 0.$$

(selection) matrix that 'selects' the 'j'-th row ~~(element)~~ and the 'i-th' column when we impose that y_j can be treated as pre-det. in equ. "i" (ie: that equations (j, i) are relatively recursive).

We can write all this info. together as a system in B, Γ:

$$\begin{array}{cc} & \begin{matrix} \quad M \quad & \quad K \quad \end{matrix} \\ \begin{matrix} K \\ g \\ M \\ h \end{matrix} & \left[\begin{array}{cc} \underset{(K\times M)}{\Pi} & \underset{(K\times K)}{I_K} \\ \hline \multicolumn{2}{c}{\underset{(g\times (M+K))}{\Phi}} \\ \hline \underset{(M\times M)}{\Omega} & \underset{(M\times K)}{0} \\ \hline \underset{(h\times M)}{\Psi}\underset{(M\times M)}{\Omega} & 0 \end{array}\right] \end{array} \cdot \begin{array}{c} M \\ \left[\begin{array}{c} \underset{(M\times M)}{B} \\ \underset{(K\times M)}{\Gamma} \end{array}\right] \end{array} \begin{matrix} M \\ K \end{matrix} = \begin{array}{c} M \\ \left[\begin{array}{c} \underset{(K\times M)}{0} \\ \hline \underset{(g\times M)}{\phi} \\ \hline \underset{(M\times M)}{B^{-1'}}\underset{(M\times M)}{\Sigma} \\ \hline \underset{(h\times M)}{0} \end{array}\right] \end{array} \begin{matrix} K \\ g \\ M \\ h \end{matrix}$$

(= # exclusion and normalization, or other linear restrictions)

These will include covariance restrictions, if we have any.

(= # Hausman-Taylor restrictions)

(= # relatively recursive pairs of equations)

If we find that this system of $(K+g+M+h)^{\times M}$ equations in $(M+K)^{\times}$ M unknowns has a unique solution, then we will have identified all the unknowns in $\begin{bmatrix} B \\ \Gamma \end{bmatrix} \equiv A$ (= all the SF parameters of interest).

We can separate this mega-system into M systems, one for each SF equation. For equation 1, this will be:

$$\begin{matrix} K \\ g_1 \\ M \\ h \end{matrix} \left[\begin{array}{cc} \Pi & I_K \\ \hline \multicolumn{2}{c}{\Phi_{(1)}} \\ \hline \Omega & 0 \\ \hline \Psi\Omega & 0 \end{array}\right] \cdot \begin{array}{c} 1 \\ \left[\begin{array}{c} B_{(1)} \\ \Gamma_{(1)} \end{array}\right] \end{array} \begin{matrix} M \\ K \end{matrix} = \begin{array}{c} 1 \\ \left[\begin{array}{c} 0 \\ \hline \phi_{(1)} \\ \hline B^{-1'}\Sigma_{(1)} \\ \hline 0 \end{array}\right] \end{array} \begin{matrix} K \\ g_1 \\ M \\ h \end{matrix}.$$

In the simple case, where we don't use any info. on elements in Σ, this is just:

$$\begin{array}{cc} & \begin{matrix} M & K \end{matrix} \\ \begin{matrix} K \\ g_1 \end{matrix} & \left[\begin{array}{cc} \Pi & I_K \\ \hline \multicolumn{2}{c}{\Phi_1} \end{array}\right] \\ & \underbrace{\qquad\qquad}_{(M+K)} \end{array} \cdot \begin{array}{c} 1 \\ \left[\begin{array}{c} B_{(1)} \\ \Gamma_{(1)} \end{array}\right] \end{array} \begin{matrix} M \\ K \end{matrix} = \begin{array}{c} 1 \\ \left[\begin{array}{c} 0 \\ \hline \phi_1 \end{array}\right] \end{array} \begin{matrix} K \\ g_1 \end{matrix}.$$

$\Phi_{(1)}$, $\phi_{(1)}$ contain all the exclusion restrictions!

- Equation (1) will then be identified iff these equations yield a unique solution for $[B_{(1)}, \Gamma_{(1)}]$.
- A necessary condition for this is that $\cancel{(K+)}g_1 \geq M\cancel{(+K)}$ (or $\geq M-1$ if we included a normalization restriction explicitly in Φ_1, ϕ_1).

This is the order condition for identification.

(cont.:) $\begin{bmatrix} \Pi & I_K \\ \multicolumn{2}{c}{\Phi_{(1)}} \end{bmatrix} \begin{bmatrix} B_{(1)} \\ \Gamma_{(1)} \end{bmatrix} = \begin{bmatrix} 0 \\ \phi_{(1)} \end{bmatrix}$

- A necessary and sufficient condition for this system of equations to yield a unique solution for $B_{(1)}, \Gamma_{(1)}$ is that: $\text{rank}\begin{bmatrix} \Pi & I_K \\ \multicolumn{2}{c}{\Phi_{(1)}} \end{bmatrix} \begin{matrix} \updownarrow M \\ \updownarrow g_{(1)} \end{matrix} = M + K$. (or $= M+K-1$, if we didn't include the norm restriction explicitly in $\Phi_{(1)}$).

This is the rank condition for identification.
Since this matrix has $M+K$ columns, its rank cannot possibly be more than $M+K$, but it can be less than that (and as little as M) if there aren't $g_1 = K$ linearly independent (= non-redundant) restrictions in $\Phi_{(1)}$.

$\Rightarrow$ An equivalent, easier-to-check, way to state the rank condition is:

$\text{rank}[\Phi_{(1)}A] = M$.
(= rank $[\Phi_{(1)}\left[\begin{smallmatrix} B \\ \Gamma \end{smallmatrix}\right]]$)

(We'll see this next week).
– along with examples to see how we do that! in practice!

D. The three major types of identifying restrictions, and IV interpretation of the restrictions

1) Exclusion restrictions:

If all restrictions in $\Phi_{(1)}A = 0$ are exclusion restrictions, then we can separate $g_{(1)}$ into g_1^Y exclusion restr. on endog. vars,
and: + g_1^Z exclusion restr. on exog. vars.

Therefore, the order condition can be rewritten as:

$$g \geq M \iff g_1^Y + g_1^Z \geq M \ (-1)$$

$$\iff \underbrace{g_1^Z}_{\substack{\#\text{ excluded} \\ \text{exog. vars.} \\ \text{in equ. (1)}}} \geq \underbrace{M - g_1^Y}_{\substack{\#\text{ included} \\ \text{endog. vars} \\ \text{in equ. (1)}}} \ (-1)$$

: we're id! (over-id if >).

~~ ~~ $\Rightarrow$ the excluded Z's from equ. (1) are used to IV for the included y's !!! in equ. (1).

2) Covariance restrictions:

If we assume $\sigma_{1j} = 0$,
~~and j is~~ and equation (j) can be estimated consistently (ie: it's id.) then we can use the consistent estimates of the residuals from equation (j), $\hat{u}_{(j)t}$ to instrument for y_j in equation (1).

3) Hausman-Taylor restrictions: (ie: (j,1) are relatively recursive)

If we assume $[B^{-1\prime}\Sigma]_{j1} = 0$, this is, if we assume that y_j can be ~~treated~~ as pre-determined in equation (1), then y_j can be used as an instrument for itself! (= we don't need to ~~another~~ exclude another Z from equation 1 to use it ~~to~~ instrument for y_j !).