Dynamic Panels with lagged dependent variables, VAR's

October 7, 2014

Recap of dynamic panel regression

With constants c_i varying across groups, short time series, model

$$y_{it} = c_i + y_{i,t-1}\rho + \varepsilon_{it} \ ,$$

we can write the likelihood for all the observables $\{y_{i0}, \ldots, y_{iT}\}$ as

$$\prod_{i=1}^{N} q(c_i, y_{i0}) \prod_{t=1}^{T} p(y_{it} \mid c_i, y_{i,t-1}) \ .$$

We use the assumption that data are independent across i and that dependence of y_{it} on the past is entirely through $y_{i,t-1}$.

Recap of dynamic panel regression

$$\prod_{i=1}^{N} q(c_i, y_{i0}) \prod_{t=1}^{T} p(y_{it} \mid c_i, y_{i,t-1}) \, .$$

We use the assumption that data are independent across i and that dependence of y_{it} on the past is entirely through $y_{i,t-1}$. And the usual assumption is

$$p(y_{i,t} \mid c_i, y_{i,t-1}) = \frac{1}{\sigma} \phi \left(\frac{y_{it} - \rho y_{i,t-1}}{\sigma} \right) \, .$$

Why an exogenous variable makes things more complicated

- With no x_t, we can use the model for the conditional distributions, so the only complication is specifying the marginal joint distribution of c_i, y_{i0}.

- This approach would also work for x's indexed only by i, though of course their effects are not identified if we also allow unconstrained group constants c_i.

Why an exogenous variable makes things more complicated

- But for x's indexed i, t, the model provides only a distribution for

$$y_{it} \mid c_i, \{y_{i,t-s-1}, x_{t-s}, s = 0, \ldots, \infty\} \ .$$

- Even if x_{it} is strictly exogenous (uncorrelated with ε_{it} at all leads and lags), to follow the strategy we used without x's would require forming a joint distribution for $y_{i0}, c_i, \{x_{is}, s = 1, \ldots, T\}$.

- Giving this an arbitrary (but probably joint normal) density q to form the likelihood conditional on $x_{i1}, \ldots, x_{iT}$ could work well only if T and the dimension of x_{it} are small relative to the number of groups.

Modeling the x's

- Any approach to estimating the model with x's involves making assumptions on their distribution and their joint distribution with c_i and y_{i0}.

- A straightforward approach is to extend the dynamic model to include x_{it}.

- New model:

$$\begin{bmatrix} y_{it} \\ x_{it} \end{bmatrix} = \begin{bmatrix} c_{iy} \\ c_{ix} \end{bmatrix} + \rho \begin{bmatrix} y_{i,t-1} \\ x_{i,t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{iyt} \\ \varepsilon_{ixt} \end{bmatrix} .$$

VAR's

- This is now, for each group, a vector autoregression, or VAR. (So the whole thing is a panel VAR.)

- Two notations for a VAR:

$$\underset{n\times 1}{y_t^*} = \underset{n\times n}{\rho}\ y_{t-1}^* + \varepsilon_t^*$$

$$y_t = B(L)y_t + \varepsilon_t \; B_s = 0 \text{ for } s < 0 \text{ or } s > k$$

$$\rho = \begin{bmatrix} B_1 & B_2 & \cdots & B_k \\ & I_{n(k-1)} & & 0 \end{bmatrix}, \qquad y_t^* = \begin{bmatrix} y_t \\ y_{t-1} \\ \vdots \\ y_{t-k+1} \end{bmatrix}, \qquad \varepsilon^* = \begin{bmatrix} \varepsilon_t \\ 0 \end{bmatrix}.$$

VAR's

- We interpret the equation as giving a conditional distribution for $y_t \mid \{y_{t-s}, s > 0\}$, via a specification for the distribution of ε_t, which is then by construction independent of past y's.

- The system is consistent with a stationary distribution for the $\{y_t\}$ sequence, i.e. one in which the distribution of any finite collection $\{y_{t_j - s}, j = 1, \ldots, J\}$ is the same for every s, if and only if all eigenvalues of ρ are less than one in absolute value.

- Any stationary, jointly normal stochastic process can be approximated arbitrarily well by a finite order autoregression, if k can be chosen as large as necessary.

- This is true both in the sense that the prediction error in using ρy_{t-1} to predict y_t comes arbitrarily close to the minimum possible, and in the sense that the joint distribution of an arbitrary $\left\{y_{t_j}, j = 1, \ldots, J\right\}$ is arbitrarily close to the correct distribution in a mean square sense.