CHAPTER 29

OVERVIEW

Part IV describes several important econometric models and applies the general tools from Part III to these models. The models themselves grow out of particular empirical situations and latent-variable models that describe simply key features. The econometric analysis capitalizes upon the latent-variable models to identify, estimate, and test parameters of interest.

1. Panel data replicate observations in several ways, typically across individuals and time periods. To account for covariance among the observations for an individual across time periods, a basic regression function contains a latent individual-specific effect α_n:

$$\mathrm{E}[y_{nt} \mid \mathbf{X}, \boldsymbol{\alpha}] = \mathbf{x}_{nt}'\boldsymbol{\beta}_0 + \alpha_n, \quad n = 1, \ldots, N$$
$$t = 1, \ldots, T$$

   Depending upon the assumptions about the α_n, one may be able to estimate $\boldsymbol{\beta}_0$ with an IV or FGLS estimator.

2. Time series data produce a need for flexible models of autocovariance. As in panel-data models, autoregressive-moving-average (ARMA) models use shared latent variables to produce a large class of autocovariance functions: if we decompose the dependent variable y_t into its regression $\mathbf{x}_t'\boldsymbol{\beta}_0$ function and a latent disturance term ε_t,

$$y_t = \mathbf{x}_t'\boldsymbol{\beta}_0 + \varepsilon_t$$

   then

$$\phi(L)\varepsilon_t = \psi(L)u_t$$

   makes $\{\varepsilon_t\}$ an ARMA(p,q) sequence where $\phi(L)$ is a pth-order polynomial, $\psi(L)$ is a qth order polynomial, and $\{u_t\}$ is a white noise sequence. For $\{\varepsilon_t\}$ to be covariance stationary, the characteristic roots of $z^p\phi(z) = 0$ must lie strictly inside the complex unit circle. GLS estimation usually exploits a prediction-error decomposition constructed with the recursive structure of the ARMA specification.

3. Multivariate dependent data presents the difficulty of separating simultaneous structural dependence from other sources of covariance. In a simple market model, quantity transacted and price are codetermined by equilibrium in supply and demand; covariance between quantity and price results from equilibrium and shared, latent, determinants. This is captured by the simultaneous system of linear equations

$$\mathbf{y}_t'\boldsymbol{\gamma}_{0j} + \mathbf{x}_t'\boldsymbol{\beta}_{0j} = \varepsilon_{tj}, \quad j = 1, \ldots, J$$
$$t = 1, \ldots, T$$

where the ε_{tj} are latent, correlated variables. One builds relatively efficient IV estimators, when $\boldsymbol{\gamma}_{0j}$ and $\boldsymbol{\beta}_{0j}$ are identified, out of the $\mathbf{x}_t$.

4. Limited dependent variables generally possess nonlinear conditional expectations; their expected values are restricted by the limits of the supports of their distributions. If a limited dependent variable y_n is a many-to-one transformation of a latent dependent variable y_t^* with a linear conditional expectation, one can derive the implied, nonlinear, conditional expectation for y_n. For example, if $y_n \in \{0, 1\}$ then

$$y_n^* \mid \mathbf{x}_n \sim \mathfrak{N}(\mathbf{x}_n' \boldsymbol{\beta}_0, \sigma_0^2)$$

and $y_n = \mathbf{1}\{y_n^* \geq 0\}$ implies that

$$\mathrm{E}[y_n \mid \mathbf{x}_n] = \mathrm{E}[\mathbf{1}\{y_n^* \geq 0\} \mid \mathbf{x}_n] = \Phi(\mathbf{x}_n' \boldsymbol{\beta}_0 / \sigma_0)$$

Estimation proceeds with NLS or ML, because the conditional distribution of y_n follows from the conditional distribution of y_n^*.