Time Series Analysis

Dynamic Heterogeneous TSCS Models (Pesaran, Chapter 28, 29, 31)

Jude C. Hays

March 31, 2021

Outline

1. Pooled Estimators, Mean-Group Estimator, and Het Tests
2. TSCS Common Factor Models
3. Unit Roots and Cointegration in TSCS Data

Dynamic Heterogeneous TSCS Models

- The dynamic heterogeneous TSCS model is an $ARDL(p, \underbrace{q, q, q....., q}_{k-times})$ of the form

$$y_{it} = \alpha_i + \sum_{j=1}^{p} \lambda_{ij} y_{i,t-j} + \sum_{j=1}^{q} \boldsymbol{\delta}'_{ij} \mathbf{x}_{i,t-j} + u_{it}, \text{ for } i = 1, 2, ..., N$$

- If this model is stationary, there is a long-run relationship between y_{it} and $\mathbf{x}_{it}$ such that

$$y_{it} = \boldsymbol{\theta}_i \mathbf{x}_{it} + \eta_{it},$$

where $\boldsymbol{\theta}_i = -\boldsymbol{\beta}_i / \phi_i, \phi_i = -(1 - \sum_{j=1}^{p} \lambda_{ij}), \boldsymbol{\beta}_i = \sum_{j=1}^{p} \boldsymbol{\delta}_{ij}.$

The Bias of Pooled Estimators

- Consider the $ARDL(1,0)$,

$$y_{it} = \alpha_i + \lambda_i y_{i,t-1} + \beta_i x_{it} + u_{it}$$

  with $\lambda_i = \lambda + \eta_{i1}$ and $\beta_i = \beta + \eta_{i2}$.
- After substituting, we have

$$\begin{aligned} y_{it} &= \alpha_i + \lambda y_{i,t-1} + \beta x_{it} + \upsilon_{it} \\ \upsilon_{it} &= u_{it} + \eta_{i1} y_{i,t-1} + \eta_{i2} x_{it} \end{aligned}$$

- It is clear that if we assume fixed λ and β, the estimates from either the FE or RE model will be biased.

The Mean-Group Estimator

- Consider again the $ARDL(1, 0...0)$

$$y_{it} = \lambda_i y_{i,t-1} + \mathbf{x'}_{it}\beta_i + u_{it}, \text{ for } i = 1, 2, ..., N; t = 1, 2, ..., T$$

- Let $\psi_i = (\lambda_i, \boldsymbol{\beta'}_i)'$, and assume the ψ_i are iid with

$$E(\psi_i) = \psi$$
$$E[(\psi_i - \psi)(\psi_i - \psi)'] = \boldsymbol{\Delta}$$

- The pooled least squares regression of y_{it} on $y_{i,t-1}$ and $\mathbf{x_{it}}$ will produce inconsistent estimates of ψ.

The Mean-Group Estimator

- However, unit by unit regressions of y_{it} on $y_{i,t-1}$ and $\mathbf{x_{it}}$ will produce consistent estimates ($T \to \infty$) of ψ_i.
- The mean-group estimator is

$$\hat{\psi}_{MG} = \frac{1}{N}\sum_{i=1}^{N}\hat{\psi}_i, \text{ with}$$

$$\widehat{var}(\hat{\psi}_{MG}) = \frac{1}{N(N-1)}\sum_{i=1}^{N}(\hat{\psi}_i - \hat{\psi}_{MG})(\hat{\psi}_i - \hat{\psi}_{MG})'$$

- The MG estimator is asymptotically normal for large N and T if $\sqrt{N}/T \to 0$ as both N and $T \to \infty$.

The Mean-Group Estimator

- The estimates for $\hat{\psi}_i$ suffer from a small-sample (Hurwicz) bias on the order of $1/T$.
- The MG estimator is unlikely to perform well when either N or T is small.

Pesaran and Yamagata Δ-test

- Pesaran and Yamagata have proposed a standardized dispersion statistic

$$\tilde{\Delta}_{adj} = \sqrt{\frac{N(T+1)}{(T-k-1)}}\left(\frac{N^{-1}\tilde{S}-k}{\sqrt{2k}}\right)$$

where $\tilde{S}$ is a modified Swamy statistic calculated by

$$\tilde{S} = \sum_{i=1}^{N}\left(\hat{\beta}_i - \tilde{\beta}_{WFE}\right)' \frac{\mathbf{X}'_i \mathbf{M}_\tau \mathbf{X}_i}{\tilde{\sigma}_i^2}\left(\hat{\beta}_i - \tilde{\beta}_{WFE}\right)$$

Pesaran and Yamagata Δ-test

$$\tilde{S} = \sum_{i=1}^{N} \left(\hat{\beta}_i - \tilde{\beta}_{WFE}\right)' \frac{\mathbf{X'}_i \mathbf{M}_\tau \mathbf{X}_i}{\tilde{\sigma}_i^2} \left(\hat{\beta}_i - \tilde{\beta}_{WFE}\right)$$

where $\mathbf{M}_\tau$ is the mean deviation matrix $\mathbf{I}_T - \frac{1}{T}\boldsymbol{\tau}_T\boldsymbol{\tau'}_T$ with $\boldsymbol{\tau}_T$ defined as a $T \times 1$ vector of ones.

$$\tilde{\beta}_{WFE} = \left(\sum_{i=1}^{N} \frac{\mathbf{X'}_i \mathbf{M}_\tau \mathbf{X}_i}{\tilde{\sigma}_i^2}\right)^{-1} \sum_{i=1}^{N} \frac{\mathbf{X'}_i \mathbf{M}_\tau \mathbf{y}_i}{\tilde{\sigma}_i^2}, \text{ and}$$

$$\tilde{\sigma}_i^2 = \frac{1}{T-1}\left(\mathbf{y}_i - \mathbf{X}_i \hat{\beta}_{FE}\right)' \mathbf{M}_\tau \left(\mathbf{y}_i - \mathbf{X}_i \hat{\beta}_{FE}\right)$$

The $\tilde{\Delta}$-test has the correct size and good power in dynamic panels as long as the autoregressive coefficient is not too close to one and $T \geq N$.

Cross-sectional Dependence and Common Factor Models

- Cross-sectional dependence is a form of dependence driven by common shocks to the units in one's dataset.
- The modern approach to cross-sectional dependence estimates dynamic common factor models of the form

$$y_{it} = \boldsymbol{\alpha'}_i \mathbf{d}_t + \boldsymbol{\beta'}_i \mathbf{x}_{it} + u_{it},$$

where $\mathbf{d}_t$ is an $n \times 1$ vector of observed common effects, $\mathbf{x}_{it}$ is a $k \times 1$ vector of observed covariates for unit i at time t, and the disturbances have the following common factor structure

$$u_{it} = \boldsymbol{\gamma'}_i \mathbf{f}_t + e_{it}$$

where $\mathbf{f}_t$ is an m-dimensional vector of unobservable common factors, and γ' is the associated vector of factor loadings.

Principal Components and Pesaran's CCE Estimators

- Bai (2009) has proposed a two-stage estimation procedure.
  1. Extract the PCs from the OLS residuals as proxies for the unobservable factors.
  2. Estimate an augmented regression where the estimated factors are treated as observable.

$$y_{it} = \boldsymbol{\alpha}'_i \mathbf{d}_t + \boldsymbol{\beta}'_i \mathbf{x}_{it} + \boldsymbol{\gamma}'_i \hat{\boldsymbol{f}}_t + e_{it}$$

- One problem is that if the factors are correlated the the regressors, the two-stage estimator is inconsistent.
- Pesaran's (2009) Correlated Common Effects (CCE) estimator approximates the linear combinations of unobserved factors by cross-sectional averages of the dependent and explanatory variables, which are included in an augmented regression.

A Simple LM test for Cross-sectional Independence

- The Breusch and Pagan (1980) Lagrange multiplier (LM) test evaluates the null hypothesis that all the pairwise error correlations are zero.
- Each pairwise correlation is estimated by

$$\hat{\rho}_{ij} = \hat{\rho}_{ji} = \frac{\sum_{t=1}^{T} \hat{u}_{it}\hat{u}_{jt}}{\left(\sum_{t=1}^{T} \hat{u}_{it}^2\right)^{1/2}\left(\sum_{t=1}^{T} \hat{u}_{jt}^2\right)^{1/2}},$$

  where $\hat{u}_t$ are the OLS residuals.
- Under the null hypothesis, asymptotically, the sum of the squared correlations will follow a χ^2 distribution with $N(N-1)/2$ degrees of freedom.

Dickey-Fuller Tests for TSCS Data

- Diagnosing Unit Roots and Cointegration in TSCS is complicated by parameter heterogeneity, cross-sectional dependence and specification of the alternative hypothesis.
- Consider the following Dickey-Fuller regression with parameter heterogeneity

$$\Delta y_{it} = \mu_i + \phi_i y_{i,t-1} + \varepsilon_{it}$$

- Im, Pesaran and Shin (2003) developed the following alternative hypothesis:

$$H_A : \phi_i < 0, i = 1, 2..., N_1, \phi_i = 0, i = N_1 + 1, N_1 + 2, ...N,$$

such that $\lim_{N\to\infty} \frac{N_1}{N} = \delta, 0 < \delta < 1$, which allows us to state the null and alternative hypotheses as $H_0 : \delta = 0$ and $H_A : \delta > 0$ respectively.

- The test statistic is the mean of the unit specific t-statistics: $\bar{t} = \frac{1}{N} \sum_{i=1}^{N} t_i$.
- This test can be augmented to allow for cross-sectional dependence by including cross-sectional averages in the DF regressions.

Adapting the Engle-Granger Methodology for TSCS Data

- Identify the order of integration for each of the variables using the appropriate Cross-section augmented Dickey-Fuller test.
- Estimate the long-run equilibrium relationship between y_t and z_t using an estimator that allows for parameter heterogeneity and cross-sectional (or spatial) dependence (e.g., Pesaran's Correlated Common Effects (CCE) estimator).
- Identify the order of integration for the estimated disturbances $\{\hat{e}_t\}$.
- If $\{y_t\}$ and $\{z_t\}$ are determined to be $I(1)$ and the disturbances are $I(0)$, we can conclude that the variables are cointegrated.