Nonstationary Panels

Based on chapters 12.1, 12.2, and 12.3 of Baltagi, B. (2005): Econometric Analysis of Panel Data, 3rd edition. Chichester, John Wiley&Sons.

Jürgen Holl, University of Vienna

Econometric Methods for Panel Data

June 3rd, 2009

Levin, Lin and Chu (LLC) test I

- Levin, Lin and Chu (2002)
- Idea:
  individual unit root tests with limited power against near unit root processes, especially in small samples
  $\rightarrow$ thus use increased sample size of panel data
- Null hypothesis:
  each individual time series contains a unit root
- Alternative hypothesis:
  each individual time series is stationary
- Maintained hypothesis, analogous to Dickey-Fuller, but double-indexed:
  1. without constant or trend:

$$\Delta y_{it} = \rho y_{i,t-1} + \sum_{L=1}^{p_i} \theta_{iL} \Delta y_{i,t-L} + \varepsilon_{it} \tag{1}$$

Levin, Lin and Chu (LLC) test II

2. with constant:

$$\Delta y_{it} = \rho y_{i,t-1} + \sum_{L=1}^{p_i} \theta_{iL} \Delta y_{i,t-L} + \alpha_i + \varepsilon_{it} \tag{2}$$

3. with constant and trend:

$$\Delta y_{it} = \rho y_{i,t-1} + \sum_{L=1}^{p_i} \theta_{iL} \Delta y_{i,t-L} + \alpha_i' \begin{pmatrix} 1 \\ t \end{pmatrix} + \varepsilon_{it} \tag{3}$$

   where α_i is the vector of coefficients corresponding to a constant and a trend component

- Three-step procedure:

Levin, Lin and Chu (LLC) test III

1. perform separate augmented Dickey-Fuller (ADF) regressions for each cross-section i:

$$\Delta y_{it} = \rho_i y_{i,t-1} + \sum_{L=1}^{p_i} \theta_{iL} \Delta y_{i,t-L} + \alpha'_{mi} d_{mt} + \varepsilon_{it} \qquad m = 1, 2, 3 \quad (4)$$

where m corresponds to the model without constant or trend ($d_{1t} = 0$), with constant ($d_{2t} = 1$), and with constant and trend ($d'_{3t} = (1, t)$), respectively.

   1.1 determine lag order to completely specify the ADF regression: choose p_{max} and use t-statistic of $\hat{\theta}_{iL}$ to decide whether smaller order is to be preferred (t-statistic of $\hat{\theta}_{iL}$ is distributed N(0,1) under $H_0 : \theta_{iL} = 0$)

   1.2 get orthogonalized residuals to use Frisch-Waugh theorem:
   regress Δy_{it} on $\Delta y_{i,t-L}$ for $L = 1, \ldots, p_i$ and d_{mt}
   $\rightarrow$ residuals $\hat{e}_{it}$
   regress $y_{i,t-1}$ on $\Delta y_{i,t-L}$ for $L = 1, \ldots, p_i$ and d_{mt}
   $\rightarrow$ residuals $\hat{\nu}_{i,t-1}$

Levin, Lin and Chu (LLC) test IV

1.3 standardize residuals to control for different variances across i:

$$\tilde{e}_{it} = \frac{\hat{e}_{it}}{\hat{\sigma}_{\varepsilon i}} \tag{5}$$

and

$$\tilde{\nu}_{i,t-1} = \frac{\hat{\nu}_{i,t-1}}{\hat{\sigma}_{\varepsilon i}} \tag{6}$$

where $\hat{\sigma}_{\varepsilon i}$ is the standard error from each ADF regression for $i = 1, \ldots, N$

2. estimate the ratio of long-run to short-run standard deviations:

2.1 long-run variance under H_0:

$$\hat{\sigma}^2_{yi} = \frac{1}{T-1}\sum_{t=2}^{T} \Delta y^2_{it} + 2\sum_{L=1}^{\bar{K}} w_{\bar{K}L}\left[\frac{1}{T-1}\sum_{t=2+L}^{T} \Delta y_{it}\Delta y_{i,t-L}\right] \tag{7}$$

where $w_{\bar{K}L} = L/(\bar{K}+1)$ for a Bartlett kernel

2.2 short-run variance:

$$\hat{\sigma}^2_{\varepsilon i} \tag{8}$$

Levin, Lin and Chu (LLC) test V

2.3 ratio:

$$\hat{s}_i = \frac{\hat{\sigma}_{yi}}{\hat{\sigma}_{\varepsilon i}} \tag{9}$$

which is a "standardized" standard deviation; the average standard deviation is then

$$\hat{S}_N = \frac{1}{N}\sum_{i=1}^{N}\hat{s}_i \tag{10}$$

3 calculate the panel test statistics:

3.1 run the pooled regression:

$$\tilde{e}_{it} = \rho\tilde{\nu}_{i,t-1} + \tilde{\varepsilon}_{it} \tag{11}$$

based on $N\tilde{T}$ observations where $\tilde{T} = T - \bar{p} - 1$ is the average number of observations per individual where $\bar{p}$ is the average lag order

Levin, Lin and Chu (LLC) test VI

3.2 calculate the conventional t-statistic under $H_0 : \rho = 0$:

$$t_\rho = \frac{\hat{\rho}}{\hat{\sigma}(\hat{\rho})} \tag{12}$$

where

$$\hat{\rho} = \frac{\sum_{i=1}^{N} \sum_{t=2+p_i}^{T} \tilde{\nu}_{i,t-1} \tilde{e}_{it}}{\sum_{i=1}^{N} \sum_{t=2+p_i}^{T} \tilde{\nu}_{i,t-1}^2} \tag{13}$$

and

$$\hat{\sigma}(\hat{\rho}) = \frac{\hat{\sigma}_{\tilde{\varepsilon}}}{\left[\sum_{i=1}^{N} \sum_{t=2+p_i}^{T} \tilde{\nu}_{i,t-1}^2\right]^{\frac{1}{2}}} \tag{14}$$

where

$$\hat{\sigma}_{\tilde{\varepsilon}}^2 = \frac{1}{N\tilde{T}} \sum_{i=1}^{N} \sum_{t=2+p_i}^{T} \left(\tilde{e}_{it} - \hat{\rho}\tilde{\nu}_{i,t-1}\right)^2 \tag{15}$$

is the estimated variance of $\tilde{\varepsilon}_{it}$

3.3 calculate the adjusted t-statistic:

$$t_\rho^* = \frac{t_\rho - N\tilde{T}\hat{S}_N \hat{\sigma}_{\tilde{\varepsilon}}^{-2} \hat{\sigma}(\hat{\rho}) \mu_{m\tilde{T}}^*}{\sigma_{m\tilde{T}}^*} \tag{16}$$

where $\mu_{m\tilde{T}}^*$ and $\sigma_{m\tilde{T}}^*$ are adjustments suggested and tabulated by LLC

Levin, Lin and Chu (LLC) test VII

- Asymptotic distribution of t_ρ^*:
  - N(0,1), which is different from univariate unit root tests converging to functionals of Brownian motions
  - requires $\sqrt{N_T}/T \to 0$ where N is an arbitrary monotonically increasing function of T
- Limitations:
  - assumes cross-sectional independence, problem in macro panels
  - does not assume $T \to \infty$ at a faster rate than $N \to \infty$ which would be sufficient but not necessary, then problem in micro panels
  - assumes ρ homogeneous across i, all cross-sections have or do not have a unit root, "every country converges at the same rate"
- Recommendations:
  - for panels of moderate size ($10 < N < 250$, $25 < T < 250$)
  - for very large T, individual unit root tests are sufficiently powerful
  - for very large N and very small T, use standard procedures

Im, Pesaran and Shin (IPS) test I

- Im, Pesaran and Shin (2003)
- Idea:
  LLC requires ρ to be homogeneous across i, implying convergence at the same rate for all i under the alternative
  $\longrightarrow$ allow ρ to be heterogeneous across i
- Null hypothesis:

$$H_0 : \rho_i = 0 \qquad \text{for all } i \tag{17}$$

- Alternative hypothesis:

$$H_1 : \begin{cases} \rho_i < 0 & \text{for } i = 1, 2, \dots, N_1 \\ \rho_i = 0 & \text{for } i = N_1 + 1, \dots, N \end{cases} \tag{18}$$

  where the fraction of stationary individual time series is assumed nonzero (necessary for the consistency of the test)

Im, Pesaran and Shin (IPS) test II

- Test statistic:

$$\bar{t} = \frac{1}{N} \sum_{i=1}^{N} t_{\rho i} \tag{19}$$

where $t_{\rho i}$ is the t-statistic from the individual ADF regressions for all i
$\rightarrow$ "average of the individual ADF statistics"

- Standard result:

$$t_{\rho i} \Rightarrow \frac{\int_0^1 W_i(r) dW_i(r)}{\left[\int_0^1 W_i(r)^2 dr\right]^{\frac{1}{2}}} = t_{iT} \tag{20}$$

for fixed N and $T \rightarrow \infty$ where t_{iT} is assumed i.i.d. with finite mean and variance

Im, Pesaran and Shin (IPS) test III

- Asymptotic distribution:

$$t_{IPS} = \frac{\sqrt{N}\left(\bar{t} - \frac{1}{N}\sum_{i=1}^{N} E[t_{iT}|\rho_i = 0]\right)}{\sqrt{\frac{1}{N}\sum_{i=1}^{N} var[t_{iT}|\rho_i = 0]}} \Rightarrow N(0,1) \qquad (21)$$

  as $T \to \infty$ and $N \to \infty$, by the Lindeberg-Lévy central limit theorem

- $E[t_{iT}|\rho_i = 0]$ and $var[t_{iT}|\rho_i = 0]$ obtained by simulations for different values of T and p_i's
- Recommendation:
  if lag order is large enough, then IPS outperforms LLC

Breitung's test I

- Breitung (2000)
- Idea:
  LLC and IPS have weak power performance with deterministic terms due to bias correction
- Three-step procedure:
  1. perform separate augmented Dickey-Fuller (ADF) regressions for each cross-section i:
     - 1.1 determine lag order to completely specify the ADF regression
     - 1.2 get orthogonalized residuals to use Frisch-Waugh theorem, but without using deterministic terms d_{mt}
     - 1.3 standardize residuals to control for different variances across i

Breitung's test II

2. transform $\tilde{e}_{it}$ using the forward orthogonalization transformation by Arellano and Bover (1995):

$$e^*_{it} = \sqrt{\frac{T-t}{T-t+1}}\left(\tilde{e}_{it} - \frac{\tilde{e}_{i,t+1} + \ldots + \tilde{e}_{i,T}}{T-t}\right) \tag{22}$$

$$\nu^*_{i,t-1} = \begin{cases} \tilde{\nu}_{i,t-1} & \text{without intercept or trend} \\ \tilde{\nu}_{i,t-1} - \tilde{\nu}_{i,1} & \text{with intercept} \\ \tilde{\nu}_{i,t-1} - \tilde{\nu}_{i,1} - \frac{t-1}{T}\tilde{\nu}_{i,T} & \text{with intercept and trend} \end{cases} \tag{23}$$

3. run the pooled regression:

$$e^*_{it} = \rho\nu^*_{i,t-1} + \varepsilon^*_{it} \tag{24}$$

$\rightarrow$ t-statistic for $H_0 : \rho = 0,\ t^*_\rho$

- Asymptotic distribution:

$$t^*_\rho \Rightarrow N(0,1) \tag{25}$$

Combined p-value tests I

- Maddala and Wu (1999) and Choi(2001)
- Let G_{iT_i} be a unit root test statistic for the ith individual
- Assume $G_{iT_i} \Rightarrow G_i$ as $T_i \rightarrow \infty$ where G_i is a nondegenerate random variable
- Let p_i be the corresponding asymptotic p-value
- Test statistic:

$$P = -2 \sum_{i=1}^{N} \ln p_i \tag{26}$$

  which uses the p-values from unit root tests for each cross-section i (e.g. ADF test) $\rightarrow$ "Fisher's inverse chi-square test"
- Asymptotic distribution:

$$P \Rightarrow \chi^2_{2N} \tag{27}$$

  as $T_i \rightarrow \infty$ for finite N
- Improved test statistic:

Combined p-value tests II

- "inverse normal test"

$$Z = \frac{1}{\sqrt{N}} \sum_{i=1}^{N} \Phi^{-1}(p_i) \qquad (28)$$

where Φ is the standard normal cumulative distribution function

- as $0 \leq p_i \leq 1$, $\phi^{-1}(p_i) \sim N(0,1)$

- Asymptotic distribution:

$$Z \Rightarrow N(0,1) \qquad (29)$$

as $T_i \rightarrow \infty$ for all i

- Advantages:
  - different lag orders may be used
  - other unit root tests may be applied
- Disadvantages:
  p-values have to be derived by Monte Carlo simulations
- Recommendations:
  - for ρ and/or T heterogeneous across i
  - combined p-value tests outperform the IPS test, the Z test performs best

Residual-based LM test I

- Hadri (2000)
- Null hypothesis: no unit root in any of the series
- Alternative hypothesis: unit root in the panel
- Model:

$$y_{it} = r_{it} + \beta_i t + \varepsilon_{it} \tag{30}$$

$$r_{it} = r_{i,t-1} + u_{it} \tag{31}$$

  where $\varepsilon_{it} \sim i.i.n.(0, \sigma_\varepsilon^2)$ and $u_{it} \sim i.i.n.(0, \sigma_u^2)$ are mutually independent across i and over t

- By back substitution:

$$y_{it} = r_{i0} + \beta_i t + \sum_{s=1}^{t} u_{is} + \varepsilon_{it} \tag{32}$$

Residual-based LM test II

- Null hypothesis revisited:

$$H_0 : \sigma_u^2 = 0 \tag{33}$$

associated with stationarity

- LM statistic:

$$LM_2 = \frac{1}{N}\left[\sum_{i=1}^{N}\left(\frac{1}{T^2}\sum_{t=1}^{T}\frac{S_{it}^2}{\hat{\sigma}_{\varepsilon i}^2}\right)\right] \tag{34}$$

where

$$S_{it} = \sum_{s=1}^{t}\hat{\varepsilon}_{is} \tag{35}$$

and

$$\hat{\sigma}_{\varepsilon i}^2 = \frac{1}{T}\sum_{t=1}^{T}\hat{\varepsilon}_{it}^2 \tag{36}$$

which allows for heteroskedasticity across i

Residual-based LM test III

- Test statistic:

$$Z = \sqrt{N}\frac{LM_2 - \xi}{\zeta} \tag{37}$$

where $\xi = 1/15$ and $\zeta = 11/6300$ for the model with constant and trend ($\xi = 1/6$ and $\zeta = 1/45$ with constant only)

- Asymptotic distribution:

$$Z \Rightarrow N(0,1) \tag{38}$$

Moon and Perron test I

- Moon and Perron (2004)
- Idea: tackle cross-sectional dependence, i.e. control for it
- Null hypothesis:

$$H_0 : \rho_i = 0 \qquad \text{for all } i \tag{39}$$

- Alternative hypothesis:

$$H_1 : \rho_i < 0 \qquad \text{for some } i \tag{40}$$

- Dynamic factor model:

$$y_{it} = \alpha_i + y_{it}^0 \tag{41}$$

$$y_{it}^0 = \rho_i y_{i,t-1}^0 + \varepsilon_{it} \tag{42}$$

where ε_{it} is generated by M unobservable random factors f_t, such that

$$\varepsilon_{it} = \Lambda_i' f_t + e_{it} \tag{43}$$

where Λ_i is a vector of nonrandom factor loading coefficients of unknown length M

Moon and Perron test II

- Let Q_Λ be the matrix projecting onto the space orthogonal to the factor loadings, purging from cross-sectional dependence; let $\sigma^2_{e,i}$ be the variance of $e_{i,t}$, $w^2_{e,i}$, the long-run variance of e_{it}, and $\lambda_{e,i}$, the one-sided long-run variance of e_{it}; let σ^2_e, w^2_e, and λ_e be their cross-sectional averages, and ϕ^4_e be the cross-sectional average of $w^4_{e,i}$
- Test statistic:

$$t_a = \frac{\sqrt{N}T(\hat{\rho}^+_{pool} - 1)}{\sqrt{\frac{2\phi^4_e}{w^4_e}}} \tag{44}$$

where the pooled bias-correlated estimate

$$\hat{\rho}^+_{pool} = \frac{tr(Y_{-1}Q_\Lambda Y') - NT\lambda^N_e}{tr(Y_{-1}Q_\Lambda Y'_{-1})} \tag{45}$$

where Y is a $T \times N$ matrix of the data, Y_{-1} contains lagged values

Moon and Perron test III

- Asymptotic distribution:

$$t_a \Rightarrow N(0,1) \tag{46}$$

where $N \to \infty$ and $T \to \infty$ such that $N/T \to 0$