Econometric Methods for Panel Data

Based on the books by Baltagi: *Econometric Analysis of Panel Data* and by Hsiao: *Analysis of Panel Data*

Robert M. Kunst
robert.kunst@univie.ac.at

University of Vienna
and
Institute for Advanced Studies Vienna

May 17, 2016

Outline

Unit roots and panels: the general idea

Unit-root tests for time series are known to have low power. It appears attractive to check whether a variable 'has a unit root'—i.e., becomes stationary after taking first differences only—using parallel data for several 'individuals' (countries).

This, of course, only makes sense in the presence of some homogeneity across individuals.

An extension is the multivariate generalization: cointegration in panels (for example, does uncovered interest parity hold in all countries?).

Panel asymptotics

Particularly unit-root panel analysis uses several concepts of letting samples grow beyond all bounds:

- *Sequential limits* investigates what happens if, first, $T \to \infty$ for fixed N, and then $N \to \infty$;
- *Joint limits* lets both T and N go to infinity simultaneously, typically restricting T to be a function of N or imposing conditions such as $T/N \to 0$.

A surprising feature is that unit-root test statistics often have standard N(0,1) limit distributions, in contrast to the time-series case $N = 1$ with its non-standard limits.

Generations of panel unit-root tests

The *first generation* of panel unit-root tests assumes that individuals are independent, while some heterogeneity across i is admitted. Most popular panel unit-root tests (LLC, IPS) are first-generation tests.

The *second generation* of panel unit-root tests admit some dependence among individuals. Dynamic cross-sample dependence, however, is often ruled out (some researchers see the development of a 'third generation' that allows for dynamic interaction).

The Levin-Lin-Chu (LLC) test: The idea

For several years, LEVIN & LIN were circulating a working paper that finally got published as LEVIN, LIN, & CHU in 2002. The LLC test builds on the Dickey-Fuller test for $H_0 : \rho = 0$ in the model

$$\Delta y_{it} = \rho y_{i,t-1} + \sum_{k=1}^{p_i} \theta_{i,k} \Delta y_{i,t-k} + \alpha_{mi} d_{mt} + \varepsilon_{it},$$

with d_{mt} the deterministic part (none, constant, or linear trend). In this model, the parameter of main concern ρ is homogeneous across i (*homogeneous alternative*), while other parameters, including lag orders p_i and the variances $\mathrm{E}\varepsilon_{it}^2$, are heterogeneous.

Three steps in the construction of LLC

1. Like in a first step of a Frisch-Waugh approach for estimating ρ by regression, Δy_{it} and $y_{i,t-1}$ are regressed separately on all remaining regressors $\Delta y_{i,t-k}, k = 1, \ldots, p_i$ and the deterministic terms. This yields residuals $\hat{e}$ and $\hat{v}$, which are re-scaled to $\tilde{e}$ and $\tilde{v}$ to mitigate heteroskedasticity across i;
2. Determine variance ratios of long-run variance of Δy_i to the variances of ε_i, and average square roots (standard errors) across i to yield $\hat{S}_N$;
3. Regress $\tilde{e}$ on $\tilde{v}$ to yield the coefficient estimate $\hat{\rho}$ as in the second Frisch-Waugh step. Calculate the test statistic from $\hat{\rho}$, $\hat{S}_N$, and some tabulated(!) correction factors.

If done correctly, the evolving statistic LLC is asymptotically N(0,1) distributed under its null $\rho = 0$, for $\sqrt{N}/T \to 0$.

LLC: the first step

In the first step, regressions

$$\Delta y_{it} = \sum_{k=1}^{p_i} \theta_{i,k}^{(1)} \Delta y_{i,t-k} + \alpha_{mi}^{(1)} d_{mt} + e_{it},$$
$$y_{i,t-1} = \sum_{k=1}^{p_i} \theta_{i,k}^{(2)} \Delta y_{i,t-k} + \alpha_{mi}^{(2)} d_{mt} + v_{i,t-1},$$

are run, with p_i determined via sequential elimination of the highest lag in the 'levels' model. These yield residuals $\hat{e}$ and $\hat{v}$, which are divided by empirical standard errors from full DF regressions $\hat{\sigma}_{\varepsilon,i}$, which yields $\tilde{e}$ and $\tilde{v}$ that have been adjusted for heteroskedasticity across i.

LLC: the second step

The concept of a long-run variance (LRV, spectral density at frequency 0) is relevant in unit-root testing. The LRV is defined by

$$\sigma_{LRV}^2 = \sum_{j=-\infty}^{\infty} \gamma_j = \sigma^2 + 2\sum_{j=1}^{\infty} \gamma_j,$$

where γ_j denotes autocovariances. A LRV of Δy_i is estimated via

$$\hat{\sigma}_{yi}^2 = \frac{1}{T-1}\sum_{t=2}^{T}(\Delta y_{it})^2 + 2\sum_{j=1}^{K} w_{K,j}\left(\frac{1}{T-1}\sum_{t=2+j}^{T}\Delta y_{it}\Delta y_{i,t-j}\right),$$

with $w_{K,j}$ denoting a *kernel function* that assigns lesser weights to $\hat{\gamma}_j$ at larger j. From this step, keep the average standard deviation ratio $\hat{S}_N = \frac{1}{N}\sum_{i=1}^{N}\hat{\sigma}_{yi}/\hat{\sigma}_{\varepsilon,i}$.

LLC: the third step

A simple regression is now run on the tilde residuals

$$\tilde{e}_{it} = \rho \tilde{v}_{i,t-1} + \mathrm{error},$$

which yields a t–statistic t_ρ and residuals $\tilde{\varepsilon}_{it}$. The LLC statistic is calculated from this t_ρ via

$$t_\rho^* = \frac{t_\rho - N\tilde{T}\hat{S}_N \hat{\sigma}_{\tilde{\varepsilon}}^{-2} \hat{\sigma}(\hat{\rho})\mu^*}{\sigma^*},$$

where μ^*, σ^* are tabulated correction factors and $\tilde{T}$ the actually available sample size in the T dimension. Note that the asymptotic null distribution is N(0,1). The statistic is hard to calculate but then it is easy to use.

The Im, Pesaran, and Shin test

To some, a drawback of the LLC test is that it imposes a common ρ. Im, Pesaran, & Shin (2003, IPS) consider the null $H_0 : \rho = 0 \quad \forall i$, and the alternative (*heterogeneous alternative*)

$$H_A : \begin{cases} \rho_i < 0, & i = 1, 2, \ldots, N_1, \\ \rho_i = 0, & i = N_1 + 1, \ldots, N. \end{cases}$$

Why not calculate DF statistics for each i and average them to yield $\tilde{t}$? Again, correction constants come in:

$$t_{IPS} = \frac{\sqrt{N}(\tilde{t} - f_E)}{f_{sd}}$$

converges to N(0,1) for sequential limits or for $N/T \to 0$. The correction factors f_E, f_{sd} are tabulated or bootstrapped.

The Breitung test

Breitung (2000) found that the handling of deterministic terms in the LLC and IPS tests is not optimal and can lead to size distortion. His reformed test uses several steps:

1. Regressions identical to first step of LLC, though omitting the d_{mt} regressors, yield residuals $\hat{e}, \hat{v}$. These are then corrected for individual-specific heteroskedasticity, so they become $\tilde{e}, \tilde{v}$.
2. Residuals $\tilde{e}$ are transformed by forward orthogonalization. Residuals $\tilde{v}$ receive transformations specific to the assumed deterministic terms, such as differencing for intercept, but no trend. New residuals are called e^* and v^*.
3. e^*_{it} is regressed on $v^*_{i,t-1}$. The thus obtained t–statistic is distributed as N(0,1) in large samples. No further correction terms are needed.

The second step of Breitung's test

The technique of forward orthogonalization follows a dynamic panel estimator developed by Arellano and Bover:

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

The final residual $v_{i,t-1}^*$ equals $\tilde{v}_{i,t-1}$ in the unusual case without intercept and trend; it becomes the difference to the start

$$v_{i,t-1}^* = \tilde{v}_{i,t-1} - \tilde{v}_{i,1}$$

for intercept, though no trend; for a linear trend, it will be

$$v_{i,t-1}^* = \tilde{v}_{i,t-1} - \tilde{v}_{i,1} - \frac{t-1}{T} \tilde{v}_{i,T}.$$

The Fisher-type tests

If p–values p_i for individual unit-root tests are available by bootstrap procedures, the combined joint statistic

$$P = -2 \sum_{i=1}^{N} \log p_i$$

has an asymptotic $\chi^2(2N)$ under the unit-root null. A variant by Choi (2001) uses the normal cumulative distribution function $\Phi(.)$:

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

with a better approximation to its asymptotic $N(0, 1)$ distribution.

Which first-generation panel unit-root test to use

Hlouskova and Wagner (2006) compare many tests in extensive Monte Carlo simulations. None of the tests dominates its rivals for all designs. Breitung's test has the advantage that it needs no tabulated correction terms. A panel version of the KPSS stationarity tests by Hadri performs poorly and should not be used. Among the Fisher-type tests, the version by Choi has the best performance.

Testing for cross-sectional dependence

First-generation panel unit-root tests can perform poorly with cross-sectional dependence (CD). The most common test for cross-sectional dependence is Pesaran's CD test. The CD statistic

$$CD = \sqrt{\frac{2T}{N(N-1)}} \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} \hat{\rho}_{ij}$$

summarizes the empirical correlation coefficients among residuals from N separate OLS regressions per individual i. Under the null of cross-section independence, CD is distributed N(0,1).

When the CD test rejects, one may want to apply second-generation tests.

PESARAN's CADF test

Among the second-generation unit-root tests, PESARAN's CADF (cross-section augmented Dickey-Fuller) test is the simplest. In a first step, run CADF regressions

$$\Delta y_{it} = \alpha_i + \rho_i^* y_{i,t-1} + d_0 \bar{y}_{t-1} + d_1 \Delta \bar{y}_t + \varepsilon_{it},$$

with cross-section averages $\bar{y}_{t-1}, \Delta \bar{y}_t$ as additional regressors. Regressions can be augmented using two types of differences $\Delta y_{i,t-k}$ and $\Delta \bar{y}_{t-k}$. In a second step, averaging across all N t–statistics for ρ_i^* yields a statistic with a tabulated/bootstrapped non-standard null distribution.

Panel cointegration tests: general

Like in time-series cointegration, there are three types of cointegration tests:

- Residual-based tests with the null of no cointegration: residuals from (potentially) cointegrating regressions are subjected to unit-root tests. If these reject, cointegration is supported.
- System tests with the null of no or less cointegration: VAR models are estimated, the rank of the impact matrix determines the cointegrating rank. Such ideas only work if T is large and N remains small.
- Tests with the null of cointegration: extensions of the KPSS idea. These have been found to perform poorly in panels.

Residual-based tests for panel cointegration

Most panel cointegration tests assume a common cointegrating relationship across i. A fixed-effects (possibly 'cointegrating') regression

$$y_{it} = x_{it}'\beta + \mu_i + e_{it}$$

is estimated by OLS, and unit-root tests are applied to the residuals $\hat{e}$. Dickey-Fuller test statistics on these $\hat{e}$ can then be calculated by a pooled regression or averaged over i individual statistics (KAO and PEDRONI tests). The Pedroni DF test has been reported to have the best properties. It dominates tests based on Phillips-Perron tests or on variance bounds.

PEDRONI tests for panel cointegration

PEDRONI (1999) considers (among other ideas) two ideas for test statistics:

- Weighted averaging (because of heteroskedasticity) of N augmented Dickey-Fuller test statistics on residuals from the cointegrating regressions. After adjusting for tabulated correction factors, this statistic converges to $N(0,1)$ under its null of no cointegration (sequential limits).
- Pooled evaluation of a joint weighted augmented DF statistic. After adjusting using tabulated values, this statistic converges to $N(0,1)$ under its null of no contegration (sequential limits).

The first version has power against the alternative of *some* cointegrating individuals, the second version has more power against cointegration in all individuals.

Estimating the cointegrating vector

If the Pedroni cointegration tests reject, there is evidence on panel cointegration. The direct cointegrating regression estimated by FE/OLS yields poor coefficient estimates.

The literature recommends two estimation methods:

- D-OLS (dynamic OLS, Kao & Chiang, 2000) augments the cointegrating regression by lags and leads

$$\tilde{y}_{it} = \tilde{x}'_{it}\beta + \sum_{j=-p_i}^{p_i} \Delta \tilde{x}'_{i,t-j}\lambda_{ij} + u_{it}$$

- Breitung suggests a two-step estimator that extends the Johansen VAR procedure to the panel case. This method is preferable if there is more than one cointegrating vector.

Breitung's two-step panel cointegration estimator

The celebrated Johansen maximum-likelihood system cointegration procedure for time series relies on an algebraic transformation of the system using an estimate of the loading matrix α. Breitung assumes cointegrating spaces as identical for all i, whereas α_i may be heterogeneous. Thus, the two steps are:

1. Run Johansen-type procedures for each $i = 1, \ldots, N$. Transform (rotate) using the estimated α_i. Also determine the cointegrating rank.
2. Run pooled OLS on all transformed observations. This yields an estimate of the matrix β that contains the cointegrating vectors.

The estimator is asymptotically normally distributed. The statistic for the rank test is also distributed N(0,1) in large samples. It uses, however, correction factors for mean and variance bias.

Which panel cointegration procedure is best?

In analogy to their panel unit-root study, Wagner and Hlouskova (2007) compared panel cointegration methods by extensive simulation. Their main results are:

- Cointegration tests and estimators can be sensitive to cross-section dependence, particularly unmodeled *cross-section cointegration* (i and j connected by a cointegrating vector);
- Among the residual-based cointegration *tests*, Pedroni tests based on the Dickey-Fuller concept are the most reliable tests;
- The best *estimator* for a single cointegrating relation is the DOLS method;
- If the cointegrating rank r is larger than one, Breitung's two-step algorithm performs best, but it tends to overstate r.