Summary based on Chapter 12 of Baltagi: Panel Unit Root Tests

at the Department of Economics

at University of Vienna

PhD-Course: Panel Data

Lecturer: Prof. Dr. Robert Kunst

Christopher Nell

Stefan Zimmermann

Vienna, June 25th, 2011

1 Introduction

Panel unit root testing emerged from time series unit root testing. The major difference to time series testing of unit roots is that we have to consider asymptotic behavior of the time-series dimension T *and* the cross-sectional dimension N. The way in which N and T converge to infinity is critical if one wants to determine the asymptotic behavior of estimators and tests used for nonstationary panels. There are several possibilities to handle the asymptotics:

1. sequential limit theory (one dimension is fixed, say N, and the other dimension T is allowed to go to infinity and provides an intermediate limit; starting from this intermediate point, N is allowed to grow large)

2. diagonal path limits (N and T go to infinity along a diagonal path—e.g., there is a monotonic increasingly connection between N and T)

3. joint limits (N and T are allowed to go to infinity at the same time)[1]

2 Levin-Lin-Chu Test

Individual unit root tests have limited power. The power of a test is the probability of rejecting the null when it is false and the null hypothesis is unit root. It follows that we find too many unit roots. Levin-Lin-Chu Test (LLC) suggest the following hypotheses

$$H_0\text{: each time series contains a unit root}$$
$$H_1\text{: each time series is stationary}$$

where the lag order p is permitted to vary across individuals. The procedure works as follows:

[1] Joint limits are generally more robust than both sequential limit theory and diagonal path limits, but they are often harder to derive and need stronger conditions to be fulfilled (e.g., existence of higher moments).

First, we run augmented Dickey-Fuller (ADF) for each cross-section on the equation:

$$\triangle y_{it} = \rho_i y_{i,t-1} + \sum_{L=1}^{p_i} \theta_{iL} \triangle y_{it-L} + \alpha_{mi} d_{mt} + \varepsilon_{it} \tag{1}$$

In the second step, we run two auxiliary regressions:

1. $\triangle y_{it}$ on $\triangle y_{i,t-L}$ and d_{mt} to obtain the residuals $\hat{e}_{it}$ and

2. $y_{i,t-1}$ on $\triangle y_{i,t-L}$ and d_{mt} to get residuals $\hat{\nu}_{i,t-1}$.

The third step involves standardization of the residuals by performing

$$\tilde{e}_{it} = \hat{e}_{it}/\hat{\sigma}_{\varepsilon_i} \tag{2}$$

$$\tilde{\nu}_{i,t-1} = \hat{\nu}_{it}/\hat{\sigma}_{\varepsilon_i} \tag{3}$$

where σ_{ε_i} denotes the standard error from each ADF.

Finally, we run the pooled OLS regression

$$\tilde{e}_{it} = \rho \tilde{\nu}_{i,t-1} + \tilde{\varepsilon}_{it} \tag{4}$$

The null hypothesis is $\rho = 0$. Notice that the standard deviation for t-statistics has to be adjusted, as can be seen in Table 2 in the original paper of Levin et al. (2002). The necessary condition for the Levin-Lin-Chu test is $\sqrt{N_T}/T \rightarrow 0$, while sufficient conditions would be $N_T/T \rightarrow 0$ and $N_T/T \rightarrow \kappa$. (N_T means that the cross-sectional dimension N is a monotonic function of time dimension T.) According to the authors, the statistic performs well when N lies between 10 and 250 and when T lies between 5 and 250 (e.g., the test is suitable for most macro panels). If T is very small, the test is undersized and has low power. One disadvantage of the test statistic is that it relies critically on the assumption of cross-sectional independence (e.g., Austria's GDP does not depend on Germany's GDP) . Moreover, the null hypothesis that *all* cross-sections have a unit root is very restrictive. That is, it does not allow the intermediate case, where some individuals are subject to a unit root and some are not. If T is very large, then Levin et al. (2002) suggest individual unit root time-series tests. If N is very large (or T very small) usual panel data procedures can be applied.

The statistical software package Stata allows an easy implementation of this panel unit root test with the following command:

```
xtunitroot llc varname [if] [in] [,LLC options].
```

3 Im, Pesaran and Shin Test

The Im-Pesaran-Shin (IPS) test is not as restrictive as the Levin-Lin-Chu test, since it allows for heterogeneous coefficients. The null hypothesis is that all individuals follow a unit root process:

$$H_0 : \rho_i = 0 \, \forall i$$

The alternative hypothesis allows some (but not all) of the individuals to have unit roots:

$$H_1 : \begin{cases} \rho_i < 0 \quad for \; i = 1, 2, \ldots, N_1 \\ \rho_i = 0 \quad for \; i = N_1 + 1, \ldots, N \end{cases}$$

When t_{ρ_i} is the individual t-statistic for testing the null hypothesis: $\rho_i = 0$ for all i, then the test is based on averaging individual unit root tests $\bar{t} = \frac{1}{N}\sum_{i=1}^{N} t_{\rho_i}$. If this statistic is properly standardized, it is asymptotically $N(0,1)$ distributed. Monto Carlo simulations reveal that the small sample performance of the Im-Pesaran-Shin test is better than Levin-Lin-Chu test. Im-Pesaran-Shin requires $N/T \to 0$ for $N \to \infty$. If either N is small or if N is large relative to T, then both Im-Pesaran-Shin and Levin-Lin-Chu show size distortions.[2] Additionally, the tests have little power if deterministic terms are included in the analysis.[3] The Stata command looks as follows:

```
xtunitroot ips varname [if] [in] [,IPS options]
```

4 Breitung's Test

The procedure of the Breitung's test can be described as follows. The first step is same as in the Levin-Lin-Chu test, except that we do not include deterministic terms. We regress $\triangle y_{it}$ on $\triangle y_{i,t-L}$ and obtain the residuals $\hat{e}_{it}$. Also, we run $y_{i,t-1}$ on $\triangle y_{i,t-L}$ and obtain the residuals $\hat{\nu}_{i,t-1}$. Afterwards, forward orthogonalization transformation is applied to the residuals $\hat{e}_{it}$ such that we obtain e^*_{it}.[4] Finally, we run the pooled regression $e^*_{it} = \rho\nu^*_{i,t-1} + \varepsilon^*_{it}$, which is asymptotically $N(0,1)$ distributed. The Stata command is:

```
xtunitroot breitung varname [if] [in] [,Breitung options]
```

5 Combining p-Value Tests

5.1 Fisher-type Test

The Fisher-type test uses p-values from unit root tests for each cross-section i. The formula of the test looks as follows:

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

[2]Size is the probability of rejecting the null when it is true. Thus, a size distortion implies that the null is rejected too often.

[3]The reason is the bias correction that removes the mean under the sequence of local alternatives.

[4]Forward orthogonalization: to each of the first $(T-1)$ observations the mean of the remaining future observations available in the sample is subtracted.

The test is asymptotically chi-square distributed with $2N$ degrees of freedom ($T_i \to \infty$ for finite N). A big benefit is that the test can handle unbalanced panels. Furthermore, the lag lengths of the individual augmented Dickey-Fuller tests are allowed to differ.[5]

A drawback of the test is that the p-values have to be obtained by Monte Carlo simulations. However, the test is easily implemented in Stata:

```
xtunitroot fisher varname [if] [in] {dfuller or:pperron} lags( )
```

5.2 Further Tests and Properties

There are several other combining p-value tests, in particular, the inverse normal test Z (standard normal distribution), the logit test L (logistic distribution), and the modified Fisher-type test (when N is large). All of these tests share certain advantages:

- The number of cross-section observations N can be finite or infinite (except for the Fisher-type test).
- Each individual allows for different types of non-stochastic and stochastic components.
- The time dimension T can vary for each individual.
- The power is superior to both the Levin-Lin-Chu test and Im-Pesaran-Shin test (IPS has higher power than LLC).

6 Residual-Based LM Test

Hadri (2000) proposes a test that builds on the Kwiatkowski-Phillips-Schmidt-Shin test (KPSS) from time series testing. KPSS tests are used for testing the

[5] The same principle can be used for every other unit root test.

null hypothesis that an observable time series is stationary around a deterministic trend. The Hadri test is based on OLS residuals obtained from regressing y_{it} on a constant (or a constant plus trend). The null hypothesis is that there is *no* unit root in any series (stationarity). The alternative hypothesis is that the panel has a unit root.

$$y_{it} = r_{it} + \varepsilon_{it} \tag{5}$$

$$r_{it} = r_{i,t-1} + u_{it} \tag{6}$$

$$H_0 : \sigma_u^2 = 0 \tag{7}$$

If variance u_{it} is zero, then r_{it} becomes a constant and thus y_{it} is stationary. The Hadri test allows for heteroskedasticity adjustments. Its empirical size is close to its nominal size if N and T are large. The Stata command for the Hadri test is:

```
xtunitroot hadri varname [if] [in] [,Hadri options]
```

7 Summary

Both the Im-Pesaran-Shin and Fisher-type test relax the restrictive assumption of Levin-Lin-Chu that ρ_i must be the same for all series under the alternative hypothesis. Also, when N is small, the empirical size of both tests is close to their nominal size of 5 percent. (Fisher shows some distortions at $N = 100$.) With respect to the size-adjusted power, the Fisher-type test outperforms the Im-Pesaran-Shin test. It should be mentioned that in the presence of a linear time trend, the power of all tests decreases considerably.

The Levin-Lin-Chu test has high power if the time dimension T is large. This can be problematic, because one might infer stationarity for the whole panel, even if it is only true for a few individuals. On the other hand, it has low power for a small time dimension. In this case, one can conclude non-stationarity when in fact most series display stationary behavior. Thus, it is advisable to analyze the outcome of both the Levin-Lin-Chu and the Im-Pesaran-Shin test.

All in all, there is no dominant performance of one particular test. That is, the econometrician must actively think about the benefits and drawbacks of the different tests (and might compare the outcomes of different tests). A caveat is that all tests discussed here assume cross-sectional independence. Therefore, they cannot be applied to panels of the second and third generation. There are, however, tests that can deal with cross-sectional dependence—for instance, Choi (2002), Chang (2002, 2004), and Pesaran (2007).

References

[1] Breitung, J. 2000. The local power of some unit root tests for panel data. Advances in Econometrics, Volume 15: Nonstationary Panels, Panel Cointegration, and Dynamic Panels, ed., B. H. Baltagi, 161–178. Amsterdam: JAY Press.

[2] Breitung, J., and S. Das. 2005. Panel unit root tests under cross-sectional dependence. Statistica Neerlandica 59: 414–433.

[3] Chang, Y. 2002. Nonlinear IV unit root tests in panels with cross sectional dependency. Journal of Econometrics 110: 261–292.

[4] Chang, Y. 2004. Bootstrap unit root tests in panels with cross sectional dependency. Journal of Econometrics 120: 263–293.

[5] Choi, I. 2001. Unit root tests for panel data. Journal of International Money and Finance 20: 249–272.

[6] Choi, I. 2002. Instrumental variables estimation of a nearly nonstationary, heterogeneous error component model. Journal of Econometrics 109: 1–32.

[7] Hadri, K. 2000. Testing for stationarity in heterogeneous panel data. Econometrics Journal 3: 148–161.

[8] Harris, R.D.F., and E. Tzavalis. 1999. Inference for unit roots in dynamic panels where the time dimension is fixed. Journal of Econometrics 91: 201–226.

[9] Im, K.S., M.H. Pesaran, and Y. Shin. 2003. Testing for unit roots in heterogeneous panels. Journal of Econometrics 115: 53–74.

[10] Levin, A., C.F. Lin, and C.S.J. Chu. 2002. Unit root tests in panel data: Asymptotic and finite-sample properties. Journal of Econometrics 108: 1–24.

[11] Pesaran, M.H. 2007. A simple panel unit root test in the presence of cross section dependence. Journal of Applied Econometrics 27: 265–312.