Title

> **xtunitroot** — Panel-data unit-root tests

Syntax   Menu   Description   Options
Remarks and examples   Stored results   Methods and formulas   Acknowledgments
References   Also see

Syntax

Levin–Lin–Chu test

    xtunitroot llc *varname* [*if*] [*in*] [, *LLC_options*]

Harris–Tzavalis test

    xtunitroot ht *varname* [*if*] [*in*] [, *HT_options*]

Breitung test

    xtunitroot breitung *varname* [*if*] [*in*] [, *Breitung_options*]

Im–Pesaran–Shin test

    xtunitroot ips *varname* [*if*] [*in*] [, *IPS_options*]

Fisher-type tests (combining p-values)

    xtunitroot fisher *varname* [*if*] [*in*], {<u>df</u>uller|<u>pp</u>erron} <u>lags</u>(*#*)
        [*Fisher_options*]

Hadri Lagrange multiplier stationarity test

    xtunitroot hadri *varname* [*if*] [*in*] [, *Hadri_options*]

LLC_options	Description
<u>trend</u>	include a time trend
<u>noconst</u>ant	suppress panel-specific means
demean	subtract cross-sectional means
<u>lags</u>(*lag_spec*)	specify lag structure for augmented Dickey–Fuller (ADF) regressions
<u>kernel</u>(*kernel_spec*)	specify method to estimate long-run variance

lag_spec is either a nonnegative integer or one of aic, bic, or hqic followed by a positive integer.
kernel_spec takes the form *kernel maxlags*, where *kernel* is one of <u>bartlett</u>, <u>parzen</u>, or <u>quadraticspectral</u>
  and *maxlags* is either a positive number or one of nwest or llc.

HT_options	Description
<u>trend</u>	include a time trend
<u>noconst</u>ant	suppress panel-specific means
<u>demean</u>	subtract cross-sectional means
altt	make small-sample adjustment to T

Breitung_options	Description
<u>trend</u>	include a time trend
<u>noconst</u>ant	suppress panel-specific means
<u>demean</u>	subtract cross-sectional means
<u>robust</u>	allow for cross-sectional dependence
<u>lags</u>(#)	specify lag structure for prewhitening

IPS_options	Description
<u>trend</u>	include a time trend
<u>demean</u>	subtract cross-sectional means
<u>lags</u>(*lag_spec*)	specify lag structure for ADF regressions

lag_spec is either a nonnegative integer or one of aic, bic, or hqic followed by a positive integer.

Fisher_options	Description
*<u>dfuller</u>	use ADF unit-root tests
*<u>pperron</u>	use Phillips–Perron unit-root tests
*<u>lags</u>(#)	specify lag structure for prewhitening
<u>demean</u>	subtract cross-sectional means
[dfuller_opts](#)	any options allowed by the dfuller command
[pperron_opts](#)	any options allowed by the pperron command

*Either dfuller or pperron is required.
*lags(#) is required.

Hadri_options	Description
<u>trend</u>	include a time trend
<u>demean</u>	subtract cross-sectional means
<u>robust</u>	allow for cross-sectional dependence
<u>kernel</u>(*kernel_spec*)	specify method to estimate long-run variance

kernel_spec takes the form *kernel* $\lceil # \rceil$, where *kernel* is one of <u>bart</u>lett, <u>parzen</u>, or <u>quadraticspectral</u> and # is a positive number.

varname may contain time-series operators; see [U] 11.4.4 Time-series varlists.

Menu

Statistics > Longitudinal/panel data > Unit-root tests

Description

xtunitroot performs a variety of tests for unit roots (or stationarity) in panel datasets. The Levin–Lin–Chu (2002), Harris–Tzavalis (1999), Breitung (2000; Breitung and Das 2005), Im–Pesaran–Shin (2003), and Fisher-type (Choi 2001) tests have as the null hypothesis that all the panels contain a unit root. The Hadri (2000) Lagrange multiplier (LM) test has as the null hypothesis that all the panels are (trend) stationary. The top of the output for each test makes explicit the null and alternative hypotheses. Options allow you to include panel-specific means (fixed effects) and time trends in the model of the data-generating process.

Options

LLC_options

trend includes a linear time trend in the model that describes the process by which the series is generated.

noconstant suppresses the panel-specific mean term in the model that describes the process by which the series is generated. Specifying noconstant imposes the assumption that the series has a mean of zero for all panels.

lags(*lag_spec*) specifies the lag structure to use for the ADF regressions performed in computing the test statistic.

Specifying lags(#) requests that # lags of the series be used in the ADF regressions. The default is lags(1).

Specifying lags(aic #) requests that the number of lags of the series be chosen such that the Akaike information criterion (AIC) for the regression is minimized. xtunitroot llc will fit ADF regressions with 1 to # lags and choose the regression for which the AIC is minimized. This process is done for each panel so that different panels may use ADF regressions with different numbers of lags.

Specifying lags(bic #) is just like specifying lags(aic #), except that the Bayesian information criterion (BIC) is used instead of the AIC.

Specifying lags(hqic #) is just like specifying lags(aic #), except that the Hannan–Quinn information criterion is used instead of the AIC.

kernel(*kernel_spec*) specifies the method used to estimate the long-run variance of each panel's series. *kernel_spec* takes the form *kernel maxlags*. *kernel* is one of bartlett, parzen, or quadraticspectral. *maxlags* is a number, nwest to request the Newey and West (1994) bandwidth selection algorithm, or llc to request the lag truncation algorithm in Levin, Lin, and Chu (2002).

Specifying, for example, kernel(bartlett 3) requests the Bartlett kernel with 3 lags.

Specifying kernel(bartlett nwest) requests the Bartlett kernel with the maximum number of lags determined by the Newey and West bandwidth selection algorithm.

Specifying kernel(bartlett llc) requests the Bartlett kernel with a maximum lag determined by the method proposed in Levin, Lin, and Chu's (2002) article:

$$maxlags = \text{int}\left(3.21T^{1/3}\right)$$

where T is the number of observations per panel. This is the default.

demean requests that xtunitroot first subtract the cross-sectional averages from the series. When specified, for each time period xtunitroot computes the mean of the series across panels and subtracts this mean from the series. Levin, Lin, and Chu suggest this procedure to mitigate the impact of cross-sectional dependence.

HT_options

trend includes a linear time trend in the model that describes the process by which the series is generated.

noconstant suppresses the panel-specific mean term in the model that describes the process by which the series is generated. Specifying noconstant imposes the assumption that the series has a mean of zero for all panels.

altt requests that xtunitroot use $T - 1$ instead of T in the formulas for the mean and variance of the test statistic under the null hypothesis. When the number of time periods, T, is small (less than 10 or 15), the test suffers from severe size distortions when fixed effects or time trends are included; in these cases, using altt results in much improved size properties at the expense of significantly less power.

demean requests that xtunitroot first subtract the cross-sectional averages from the series. When specified, for each time period xtunitroot computes the mean of the series across panels and subtracts this mean from the series. Levin, Lin, and Chu suggest this procedure to mitigate the impact of cross-sectional dependence.

Breitung_options

trend includes a linear time trend in the model that describes the process by which the series is generated.

noconstant suppresses the panel-specific mean term in the model that describes the process by which the series is generated. Specifying noconstant imposes the assumption that the series has a mean of zero for all panels.

lags(#) specifies the number of lags used to remove higher-order autoregressive components of the series. The Breitung test assumes the data are generated by an AR(1) process; for higher-order processes, the first-differenced and lagged-level data are replaced by the residuals from regressions of those two series on the first # lags of the first-differenced data. The default is to not perform this prewhitening step.

robust requests a variant of the test that is robust to cross-sectional dependence.

demean requests that xtunitroot first subtract the cross-sectional averages from the series. When specified, for each time period xtunitroot computes the mean of the series across panels and subtracts this mean from the series. Levin, Lin, and Chu suggest this procedure to mitigate the impact of cross-sectional dependence.

IPS_options

trend includes a linear time trend in the model that describes the process by which the series is generated.

lags(*lag_spec*) specifies the lag structure to use for the ADF regressions performed in computing the test statistic. With this option, xtunitroot reports Im, Pesaran, and Shin's (2003) $W_{t\text{-}bar}$ statistic that is predicated on T going to infinity first, followed by N going to infinity. By default, no lags are included, and xtunitroot instead reports Im, Pesaran, and Shin's $\widetilde{t}\text{-}bar$ and $Z_{\widetilde{t}\text{-}bar}$ statistics that assume T is fixed while N goes to infinity, as well as the $t\text{-}bar$ statistic and exact critical values that assume both N and T are fixed.

Specifying lags(*#*) requests that *#* lags of the series be used in the ADF regressions. By default, no lags are included.

Specifying lags(aic *#*) requests that the number of lags of the series be chosen such that the AIC for the regression is minimized. xtunitroot llc will fit ADF regressions with 1 to *#* lags and choose the regression for which the AIC is minimized. This process is done for each panel so that different panels may use ADF regressions with different numbers of lags.

Specifying lags(bic *#*) is just like specifying lags(aic *#*), except that the BIC is used instead of the AIC.

Specifying lags(hqic *#*) is just like specifying lags(aic *#*), except that the Hannan–Quinn information criterion is used instead of the AIC.

If you specify lags(0), then xtunitroot reports the $W_{t\text{-}bar}$ statistic instead of the $Z_{t\text{-}bar}$, $Z_{\widetilde{t}\text{-}bar}$, and $t\text{-}bar$ statistics.

demean requests that xtunitroot first subtract the cross-sectional averages from the series. When specified, for each time period xtunitroot computes the mean of the series across panels and subtracts this mean from the series. Levin, Lin, and Chu suggest this procedure to mitigate the impact of cross-sectional dependence.

Fisher_options

dfuller requests that xtunitroot conduct ADF unit-root tests on each panel by using the dfuller command. You must specify either the dfuller or the pperron option.

pperron requests that xtunitroot conduct Phillips–Perron unit-root tests on each panel by using the pperron command. You must specify either the pperron or the dfuller option.

lags(*#*) specifies the number of lags used to remove higher-order autoregressive components of the series. The Fisher test assumes the data are generated by an AR(1) process; for higher-order processes, the first-differenced and lagged-level data are replaced by the residuals from regressions of those two series on the first *#* lags of the first-differenced data. lags(*#*) is required.

dfuller_opts are any options accepted by the dfuller command, including noconstant, trend, drift, and lags(). Because xtunitroot calls dfuller quietly, the dfuller option regress has no effect. See [TS] **dfuller**.

pperron_opts are any options accepted by the pperron command, including noconstant, trend, and lags(). Because xtunitroot calls pperron quietly, the pperron option regress has no effect. See [TS] **pperron**.

demean requests that xtunitroot first subtract the cross-sectional averages from the series. When specified, for each time period xtunitroot computes the mean of the series across panels and subtracts this mean from the series. Levin, Lin, and Chu suggest this procedure to mitigate the impact of cross-sectional dependence.

Hadri_options

trend includes a linear time trend in the model that describes the process by which the series is generated.

robust requests a variant of the test statistic that is robust to heteroskedasticity across panels.

kernel(*kernel_spec*) requests a variant of the test statistic that is robust to serially correlated errors. *kernel_spec* specifies the method used to estimate the long-run variance of each panel's series. *kernel_spec* takes the form *kernel* $\lceil$*#*$\rceil$. Three kernels are supported: bartlett, parzen, and quadraticspectral.

Specifying, for example, kernel(bartlett 3) requests the Bartlett kernel with 3 lags.

If # is not specified, then 1 lag is used.

demean requests that xtunitroot first subtract the cross-sectional averages from the series. When specified, for each time period xtunitroot computes the mean of the series across panels and subtracts this mean from the series. Levin, Lin, and Chu suggest this procedure to mitigate the impact of cross-sectional dependence.

Remarks and examples

stata.com

Remarks are presented under the following headings:

> *Overview*
> *Levin–Lin–Chu test*
> *Harris–Tsavalis test*
> *Breitung test*
> *Im–Pesaran–Shin test*
> *Fisher-type tests*
> *Hadri LM test*

Overview

We consider a simple panel-data model with a first-order autoregressive component:

$$y_{it} = \rho_i y_{i,t-1} + \mathbf{z}'_{it}\boldsymbol{\gamma}_i + \epsilon_{it} \tag{1}$$

where $i = 1, \ldots, N$ indexes panels; $t = 1, \ldots, T_i$ indexes time; y_{it} is the variable being tested; and ϵ_{it} is a stationary error term. The $\mathbf{z}_{it}$ term can represent panel-specific means, panel-specific means and a time trend, or nothing, depending on the options specified to xtunitroot. By default, $\mathbf{z}_{it} = 1$, so that the term $\mathbf{z}'_{it}\boldsymbol{\gamma}_i$ represents panel-specific means (fixed effects). If trend is specified, $\mathbf{z}'_{it} = (1, t)$ so that $\mathbf{z}'_{it}\boldsymbol{\gamma}_i$ represents panel-specific means and linear time trends. For tests that allow it, specifying noconstant omits the $\mathbf{z}'_{it}\boldsymbol{\gamma}_i$ term. The Im–Pesaran–Shin (xtunitroot ips), Fisher-type (xtunitroot fisher), and Hadri LM (xtunitroot hadri) tests allow unbalanced panels, while the remaining tests require balanced panels so that $T_i = T$ for all i.

Panel unit-root tests are used to test the null hypothesis $H_0 : \rho_i = 1$ for all i versus the alternative $H_a : \rho_i < 1$. Depending on the test, H_a may hold, for one i, a fraction of all i or all i; the output of the respective test precisely states the alternative hypothesis. Equation (1) is often written as

$$\Delta y_{it} = \phi_i y_{i,t-1} + \mathbf{z}'_{it}\boldsymbol{\gamma}_i + \epsilon_{it} \tag{1'}$$

so that the null hypothesis is then $H_0 : \phi_i = 0$ for all i versus the alternative $H_a : \phi_i < 0$.

The Hadri LM test for panel stationarity instead assumes the null hypothesis that all panels are stationary versus the alternative that at least some of the panels contain unit roots. We discuss the Hadri LM test in detail later, though for now our remarks focus on tests whose null hypothesis is that the panels contain unit roots.

The various panel unit-root tests implemented by xtunitroot differ in several key aspects. First, the Levin–Lin–Chu (xtunitroot llc), Harris–Tsavalis (xtunitroot ht), and Breitung (xtunitroot breitung) tests make the simplifying assumption that all panels share the same autoregressive parameter so that $\rho_i = \rho$ for all i. The other tests implemented by xtunitroot, however, allow the autoregressive parameter to be panel specific. Maddala and Wu (1999) provide an example of testing whether countries' economic growth rates converge to a long-run value. Imposing the restriction that $\rho_i = \rho$ for all i implies that the rate of convergence would be the same for all countries, an implication that is too restrictive in practice.

Second, the various tests make differing assumptions about the rates at which the number of panels, N, and the number of time periods, T, tend to infinity or whether N or T is fixed. For microeconomic panels of firms, for example, increasing the sample size would involve gathering data on more firms while holding the number of time periods fixed; here N tends to infinity whereas T is fixed. In a macroeconomic analysis of OECD countries, one would typically assume that N is fixed whereas T tends to infinity.

Related to the previous point, the size of one's sample will in large part determine which test is most appropriate in a given situation. If a dataset has a small number of panels and a large number of time periods, then a panel unit-root test that assumes that N is fixed or that N tends to infinity at a slower rate than T will likely perform better than one that is designed for cases where N is large.

Hlouskova and Wagner (2006) provide a good overview of the types of panel unit-root tests available with xtunitroot, and they present exhaustive Monte Carlo simulations examining the tests' performance. Baltagi (2013, chap. 12) also concisely discusses the tests implemented by xtunitroot.

The following table summarizes some of the key differences among the various tests:

Test	Options	Asymptotics	ρ under H_a	Panels
LLC	`noconstant`	$\sqrt{N}/T \to 0$	common	balanced
LLC		$N/T \to 0$	common	balanced
LLC	`trend`	$N/T \to 0$	common	balanced
HT	`noconstant`	$N \to \infty,\ T$ fixed	common	balanced
HT		$N \to \infty,\ T$ fixed	common	balanced
HT	`trend`	$N \to \infty,\ T$ fixed	common	balanced
Breitung	`noconstant`	$(T, N) \to_{\text{seq}} \infty$	common	balanced
Breitung		$(T, N) \to_{\text{seq}} \infty$	common	balanced
Breitung	`trend`	$(T, N) \to_{\text{seq}} \infty$	common	balanced
IPS		$N \to \infty,\ T$ fixed or N and T fixed	panel-specific	unbalanced
IPS	`trend`	$N \to \infty,\ T$ fixed or N and T fixed	panel-specific	unbalanced
IPS	`lags()`	$(T, N) \to_{\text{seq}} \infty$	panel-specific	unbalanced
IPS	`trend lags()`	$(T, N) \to_{\text{seq}} \infty$	panel-specific	unbalanced
Fisher-type		$T \to \infty,\ N$ finite or infinite	panel-specific	unbalanced
Hadri LM		$(T, N) \to_{\text{seq}} \infty$	(not applicable)	balanced
Hadri LM	`trend`	$(T, N) \to_{\text{seq}} \infty$	(not applicable)	balanced

The first column identifies the test procedure, where we use LLC to denote the Levin–Lin–Chu test, HT to denote the Harris–Tsavalis test, and IPS to denote the Im–Pesaran–Shin test. The second column indicates the deterministic components included in (1) or (1′). The column labeled "Asymptotics" indicates the behavior of the number of panels, N, and time periods, T, required for the test statistic to have a well-defined asymptotic distribution. For example, the LLC test without the noconstant

option requires that T grow at a faster rate than N so that N/T approaches zero; with the `noconstant` option, we need only for T to grow faster than the square root of N (so T could grow more slowly than N).

The HT tests and the IPS tests without accommodations for serial correlation assume that the number of time periods, T, is fixed, whereas N tends to infinity; `xtunitroot` also reports critical values for the IPS tests that are valid in finite samples (where N and T are fixed).

Many of the tests are justified using sequential limit theory, which we denote as $(T, N) \rightarrow_{\text{seq}} \infty$. First, the time dimension goes to infinity, and then the number of panels goes to infinity. As a practical matter, these tests work best with "large" T and at least "moderate" N. See Phillips and Moon (2000) for an introduction to asymptotics that depend on both N and T and their relation to nonstationary panels. Phillips and Moon (1999) contains a more technical discussion of "multi-indexed" asymptotics.

The fourth column refers to the parameter ρ_i in (1) and ϕ_i in (1′). As we mentioned previously, some tests assume that all panels have the same autoregressive parameter under the alternative hypothesis of stationarity (denoted "common" in the table), while others allow for panel-specific autoregressive parameters (denoted "panel-specific" in the table). The Hadri LM tests are not framed in terms of an equation like (1) or (1′), so the distinction based on ρ is not applicable.

The final column indicates whether the panel dataset must be strongly balanced, meaning each panel has the same number of observations covering the same time span. Except for the Fisher tests, all the tests require that there be no gaps in any panel's series.

We now discuss each test in turn.

Levin–Lin–Chu test

The starting point for the Levin–Lin–Chu (LLC) test is (1′) with the restriction that all panels share a common autoregressive parameter. In a regression model like (1), ϵ_{it} is likely to be plagued by serial correlation, so to mitigate this problem, LLC augment the model with additional lags of the dependent variable:

$$\Delta y_{it} = \phi y_{i,t-1} + \mathbf{z}'_{it}\gamma_i + \sum_{j=1}^{p} \theta_{ij}\Delta y_{i,t-j} + u_{it} \tag{2}$$

The number of lags, p, can be specified using the `lags()` option, or you can have `xtunitroot llc` select the number of lags that minimizes one of several information criteria. The LLC test assumes that ϵ_{it} is independently distributed across panels and follows a stationary invertible autoregressive moving-average process for each panel. By including sufficient lags of $\Delta y_{i,t}$ in (2), u_{it} will be white noise; the test does not require u_{it} to have the same variance across panels.

Under the null hypothesis of a unit root, y_{it} is nonstationary, so a standard OLS regression t statistic for ϕ will have a nonstandard distribution that depends in part on the specification of the $\mathbf{z}_{it}$ term. Moreover, the inclusion of a fixed-effect term in a dynamic model like (2) causes the OLS estimate of ϕ to be biased toward zero; see Nickell (1981). The LLC method produces a bias-adjusted t statistic, which the authors denote as t_δ^*, that has an asymptotically normal distribution.

The LLC test without panel-specific intercepts or time trends, requested by specifying the `noconstant` option with `xtunitroot llc`, is justified asymptotically if $\sqrt{N}/T \rightarrow 0$, allowing the time dimension T to grow more slowly than the cross-sectional dimension N; LLC (2002) mention that this assumption is particularly relevant for panel datasets typically encountered in microeconomic applications.

If model (2) includes panel-specific means (the default for `xtunitroot llc`) or time trends (requested with the `trend` option), then you must assume that $N/T \rightarrow 0$ for the t_δ^* statistic to have

an asymptotically standard normal distribution. This implies that the time dimension, T, must grow faster than the cross-sectional dimension, N, a situation more plausible with macroeconomic datasets.

LLC (2002) recommend using their test with panels of "moderate" size, which they describe as having between 10 and 250 panels and 25 to 250 observations per panel. Baltagi (2013, sec. 12.2.3) mentions that the requirement $N/T \rightarrow 0$ implies that N should be small relative to T.

❑ Technical note

Panel unit-root tests have frequently been used to test the purchasing power parity (PPP) hypothesis. We use a PPP dataset to illustrate the xtunitroot command, but understanding PPP is not required to understand how these tests are applied. Here we outline PPP and explain how to test it using panel unit-root tests; uninterested readers can skip the remainder of this technical note. Our discussion and examples are motivated by those in Oh (1996) and Patterson (2000, chap. 13). Also see Rogoff (1996) for a broader introduction to PPP.

The PPP hypothesis is based on the Law of One Price, which stipulates that the price of a tradeable good will be the same everywhere. Absolute PPP stipulates that the nominal exchange rate, E, is

$$E = \frac{P}{P^*}$$

where P is the price of a basket of goods in the home country and P^* is the price of the same basket in the foreign country. The exchange rate, E, indicates the price of a foreign currency in terms of our "home" currency or, equivalently, how many units of the home currency are needed to buy one unit of the foreign currency.

Now consider the real exchange rate, λ, which tells us the prices of goods and services—things we actually consume—in a foreign country relative to their prices at home. We have

$$\lambda = \frac{EP^*}{P} \qquad (3)$$

λ in general does not equal unity for many reasons, including the fact that not all goods are tradeable across countries (haircuts being the textbook example), trade barriers such as tariffs and quotas, differences among countries in how price indices are constructed, and the Harrod–Balassa–Samuelson effect, which links productivity and price levels; see Obstfeld and Rogoff (1996, 210–216).

Taking logs of both sides of (3), we have

$$y \equiv \ln\lambda = \ln E + \ln P^* - \ln P$$

PPP holds only if the real exchange rate reverts to its equilibrium value over time. Thus, to test for PPP, we test whether y contains a unit root. If y does contain a unit root, we reject PPP.

The dataset pennxrate.dta contains real exchange-rate data based on the Penn World Table version 6.2 (Heston, Summers, and Aten 2006). The data are a balanced panel consisting of 151 countries observed over 34 years, from 1970 through 2003. The United States was treated as the domestic country and is therefore not included. The variable lnrxrate contains the log of the real exchange rate and is the variable on which we conduct panel unit-root tests in the examples.

Two indicator variables are included in the dataset as well. The variable oecd flags 27 countries aside from the United States that are members of the Organization for Economic Cooperation and Development (OECD). (The Czech Republic and the Slovak Republic are excluded because they did not become independent countries until 1993.) The variable g7 flags the six countries aside from the United States that are members of the Group of Seven (G7) nations.

❑

▷ Example 1

The dataset `pennxrate.dta` contains real exchange-rate data for a panel of countries observed over 34 years. Here we use the LLC test to determine whether the series `lnrxrate`, the log of real exchange rates, contains a unit root for six nations that are currently in the G7 group of advanced economies. We do not have any reason to believe `lnrxrate` should exhibit a global trend, so we do not include the `trend` option.

Looking at (2), we have no a priori knowledge of the number of lags, p, needed to ensure that u_{it} is white noise, so we let `xtunitroot` choose the number of lags for each panel by minimizing the AIC, subject to a maximum of 10 lags.

We type

```
. use http://www.stata-press.com/data/r13/pennxrate
. xtunitroot llc lnrxrate if g7, lags(aic 10)

Levin-Lin-Chu unit-root test for lnrxrate
─────────────────────────────────────────────────────────────────
Ho: Panels contain unit roots            Number of panels  =      6
Ha: Panels are stationary                Number of periods =     34

AR parameter: Common                     Asymptotics: N/T -> 0
Panel means:  Included
Time trend:   Not included

ADF regressions: 1.00 lags average (chosen by AIC)
LR variance:    Bartlett kernel, 10.00 lags average (chosen by LLC)
─────────────────────────────────────────────────────────────────
                       Statistic       p-value
─────────────────────────────────────────────────────────────────
  Unadjusted t          -6.7538
  Adjusted t*           -4.0277         0.0000
─────────────────────────────────────────────────────────────────
```

The header of the output summarizes the exact specification of the test and dataset. Because we did not specify the `noconstant` option, the test allowed for panel-specific means. On average, $p = 1$ lag of the dependent variable of (2) were included as regressors in the ADF regressions. By default, `xtunitroot` estimated the long-run variance of $\Delta\text{lnrxrate}_{it}$ by using a Bartlett kernel with an average of 10 lags.

The LLC bias-adjusted test statistic $t_{\delta}^* = -4.0277$ is significantly less than zero ($p < 0.00005$), so we reject the null hypothesis of a unit-root [that is, that $\phi = 0$ in (2)] in favor of the alternative that `lnrxrate` is stationary (that is, that $\phi < 0$). This conclusion supports the PPP hypothesis.

Labeled "Unadjusted t" in the output is a conventional t statistic for testing $H_0 : \phi = 0$. When the model does not include panel-specific means or trends, this test statistic has a standard normal limiting distribution and its p-value is shown in the output; the unadjusted statistic, t_{δ}, diverges to negative infinity if trends or panel-specific constants are included, so a p-value is not displayed in those cases.

Because the G7 economies have many similarities, our results could be affected by cross-sectional correlation in real exchange rates; O'Connell's (1998) results showed that the LLC test exhibits severe size distortions in the presence of cross-sectional correlation. LLC (2002) suggested removing cross-sectional averages from the data to help control for this correlation. We can do this by specifying the `demean` option to `xtunitroot`:

```
. xtunitroot llc lnrxrate if g7, lags(aic 10) demean

Levin-Lin-Chu unit-root test for lnrxrate
───────────────────────────────────────────────────────────────────────
Ho: Panels contain unit roots            Number of panels  =       6
Ha: Panels are stationary                Number of periods =      34

AR parameter: Common                     Asymptotics: N/T -> 0
Panel means:  Included
Time trend:   Not included               Cross-sectional means removed

ADF regressions: 1.50 lags average (chosen by AIC)
LR variance:     Bartlett kernel, 10.00 lags average (chosen by LLC)
───────────────────────────────────────────────────────────────────────
                      Statistic      p-value
───────────────────────────────────────────────────────────────────────
 Unadjusted t         -5.5473
 Adjusted t*          -2.0813         0.0187
───────────────────────────────────────────────────────────────────────
```

Once we control for cross-sectional correlation by removing cross-sectional means, we can no longer reject the null hypothesis of a unit root at the 1% significance level, though we can reject at the 5% level.

◁

Here we chose the number of lags based on the AIC criterion in an admission that we do not know the true number of lags to include in (2). However, the test statistics are derived under the assumption that the lag order, p, is known. If we happen to choose the wrong number of lags, then the distribution of the test statistic will depart from its expected distribution that assumes p is known.

Harris–Tsavalis test

In many datasets, particularly in microeconomics, the time dimension, T, is small, so tests whose asymptotic properties are established by assuming that T tends to infinity can lead to incorrect inference. HT (1999) derived a unit-root test that assumes that the time dimension, T, is fixed. Their simulation results suggest that the test has favorable size and power properties for N greater than 25, and they report (p. 213) that power improves faster as T increases for a given N than when N increases for a given T.

The HT test statistic is based on the OLS estimator, ρ, in the regression model

$$y_{it} = \rho y_{i,t-1} + \mathbf{z}'_{it}\boldsymbol{\gamma}_i + \epsilon_{it} \tag{4}$$

where the term $\mathbf{z}'_{it}\boldsymbol{\gamma}_i$ allows for panel-specific means and trends and was discussed in *Overview*. Harris and Tsavalis assume that ϵ_{it} is independent and identically distributed (i.i.d.) normal with constant variance across panels. Because of the bias induced by the inclusion of the panel means and time trends in this model, the expected value of the OLS estimator is not equal to unity under the null hypothesis. Harris and Tsavalis derived the mean and standard error of $\widehat{\rho}$ for (4) under the null hypothesis $H_0 : \rho = 1$ when neither panel-specific means nor time trends are included (requested with the noconstant option), when only panel-specific means are included (the default), and when both panel-specific means and time trends are included (requested with the trend option). The asymptotic distribution of the test statistic is justified as $N \to \infty$, so you should have a relatively large number of panels when using this test. Notice that, like the LLC test, the HT test assumes that all panels share the same autoregressive parameter.

▷ Example 2

Because the HT test is designed for cases where N is relatively large, here we test whether the series lnrxrate contains a unit root using all 151 countries in our dataset. We will again remove cross-sectional means to help control for contemporaneous correlation. We type

```
. xtunitroot ht lnrxrate, demean

Harris-Tzavalis unit-root test for lnrxrate
```

	Number of panels =	151

Ho: Panels contain unit roots   Number of panels =   151
Ha: Panels are stationary   Number of periods =   34

AR parameter: Common   Asymptotics: N -> Infinity
Panel means: Included   T Fixed
Time trend: Not included   Cross-sectional means removed

	Statistic	z	p-value
rho	0.8184	-13.1239	0.0000

Here we strongly reject the null hypothesis of a unit root, again finding support for PPP. The point estimate of ρ in (4) is 0.8184, and the z statistic is -13.12.

◁

Can we directly compare the results from the LLC and HT tests? We used a subset of the data for the LLC test but used all the data for the HT test. That leads to the obvious answer that no, our results are not entirely comparable. However, a more subtle issue regarding the asymptotic properties of the tests also warrants caution when comparing results.

The LLC test assumes that $N/T \rightarrow 0$, so N should be small relative to T. Moreover, with our exchange-rate dataset, we are much more likely to be able to add more years of data rather than add more countries, because the number of countries in the world is for the most part fixed. Hence, assuming T grows faster than N is certainly plausible.

On the other hand, the HT test assumes that T is fixed whereas N goes to infinity. Is that assumption plausible for our dataset? As we just mentioned, T likely grows faster than N here, so using a test that assumes T is fixed whereas N grows is hard to justify with our dataset.

In short, when selecting a panel unit-root test, you must consider the relative sizes of N and T and the relative speeds at which they tend to infinity or whether either N or T is fixed.

Breitung test

Both the LLC and HT tests take the approach of first fitting a regression model and subsequently adjusting the autoregressive parameter or its t statistic to compensate for the bias induced by having a dynamic regressor and fixed effects in the model. The Breitung (2000; Breitung and Das 2005) test takes a different tact, adjusting the data before fitting a regression model so that bias adjustments are not needed.

In the LLC test, additional lags of the dependent variable could be included in (2) to control for serial correlation. The Breitung procedure instead allows for a prewhitening of the series before computing the test. If the trend option is not specified, we regress Δy_{it} and $y_{i,t-1}$ on $\Delta y_{i,t-1}, \ldots, \Delta y_{i,t-p}$ and use the residuals from those regressions in place of $\Delta y_{i,t}$ and $y_{i,t-1}$ in computing the test. You specify the number of lags, p, to use by specifying lags(#). If the trend option is specified, then the Breitung method uses a different prewhitening procedure that involves fitting only one (instead of two) preliminary regressions; see *Methods and formulas* for details.

Monte Carlo simulations by Breitung (2000) show that bias-corrected statistics such as LLC's t_δ^* suffer from low power, particularly against alternative hypotheses with autoregressive parameters near one and when panel-specific effects are included. In contrast, the Breitung (2000) test statistic exhibits much higher power in these cases. Moreover, the Breitung test has good power even with small datasets ($N = 25$, $T = 25$), though the power of the test appears to deteriorate when T is fixed and N is increased.

The Breitung test assumes that the error term ϵ_{it} is uncorrelated across both i and t. `xtunitroot` `breitung` optionally also reports a version of the statistic based on Breitung and Das (2005) that is robust to cross-sectional correlation.

▷ Example 3

Here we test whether `lnrxrate` contains a unit root for the subset of 27 OECD countries in our dataset. We will use the `robust` option to obtain a test statistic that is robust to cross-sectional correlation, so we will not subtract the cross-sectional means via the `demean` option. We type

```
. xtunitroot breitung lnrxrate if oecd, robust

Breitung unit-root test for lnrxrate
────────────────────────────────────────

Ho: Panels contain unit roots          Number of panels  =    27
Ha: Panels are stationary              Number of periods =    34

AR parameter: Common                   Asymptotics:  T,N -> Infinity
Panel means:  Included                               sequentially
Time trend:   Not included             Prewhitening: Not performed
──────────────────────────────────────────────────────────────────
                   Statistic       p-value
──────────────────────────────────────────────────────────────────
   lambda*          -1.6794        0.0465
──────────────────────────────────────────────────────────────────

 * Lambda robust to cross-sectional correlation
```

We can reject the null of a unit root at the 5% level but not at the 1% level.

◁

Im–Pesaran–Shin test

All the tests we have discussed thus far assume that all panels share a common autoregressive parameter, ρ. Cultural, institutional, and other factors make such an assumption tenuous for both macro- and microeconometric panel datasets. IPS (2003) developed a set of tests that relax the assumption of a common autoregressive parameter. Moreover, the IPS test does not require balanced datasets, though there cannot be gaps within a panel. The starting point for the IPS test is a set of Dickey–Fuller regressions of the form

$$\Delta y_{it} = \phi_i y_{i,t-1} + \mathbf{z}_{it}'\boldsymbol{\gamma}_i + \epsilon_{it} \tag{5}$$

Notice that here ϕ is panel-specific, indexed by i, whereas in (2), ϕ is constant. Im, Pesaran, and Shin assume that ϵ_{it} is independently distributed normal for all i and t, and they allow ϵ_{it} to have heterogeneous variances σ_i^2 across panels.

As described by Maddala and Wu (1999), one way to view the key difference between the IPS and LLC tests is that here we fit (5) to each panel separately and average the resulting t statistics, whereas in the LLC test we pool the data before fitting an equation such as (2) (thus we impose a common autoregressive parameter) and compute a test statistic based on the pooled regression results.

Under the null hypothesis that all panels contain a unit root, we have $\phi_i = 0$ for all i. The alternative is that the fraction of panels that follow stationary processes is nonzero; that is, as N tends to infinity, the fraction N_1/N converges to a nonzero value, where N_1 is the number of panels that are stationary.

Whether you allow for serially correlated errors determines the test statistics produced, and because there are substantive differences in the output, we consider the serially uncorrelated and serially correlated cases separately. First, we consider the serially uncorrelated case, which `xtunitroot` assumes when you do not specify the `lags()` option.

The IPS test allowing for heterogeneous panels with serially uncorrelated errors assumes that the number of time periods, T, is fixed; `xtunitroot ips` produces statistics both for the case where N is fixed and for the case where $N \to \infty$. Under the null hypothesis of a unit root, the usual t statistic, t_i, for testing $H_0: \phi_i = 0$ in (5) does not have a mean of zero. For the case where N is fixed, IPS used simulation to tabulate "exact" critical values for the average of the t_i statistics when the dataset is balanced; these critical values are not available with unbalanced datasets. The critical values are "exact" only when the error term is normally distributed and when T corresponds to one of the sample sizes used in their simulation studies. For other values of T, `xtunitroot ips` linearly interpolates the values in IPS (2003, table 2).

For the case where $N \to \infty$, they used simulation to tabulate the mean and variance of t_i for various values of T under the null hypothesis and showed that a bias-adjusted average of the t_i's has a standard normal limiting distribution. We illustrate the test with an example.

▷ Example 4

Here we test whether `lnrxrate` contains a unit root for the subset of OECD countries. We type

```
. xtunitroot ips lnrxrate if oecd, demean
Im-Pesaran-Shin unit-root test for lnrxrate
```

```
Ho: All panels contain unit roots          Number of panels  =       27
Ha: Some panels are stationary             Number of periods =       34

AR parameter: Panel-specific               Asymptotics: T,N -> Infinity
Panel means:  Included                                      sequentially
Time trend:   Not included                 Cross-sectional means removed

ADF regressions: No lags included
```

	Statistic	p-value	Fixed-N exact critical values		
			1%	5%	10%
t-bar	-3.1327		-1.810	-1.730	-1.680
t-tilde-bar	-2.5771				
Z-t-tilde-bar	-7.3911	0.0000			

As with the other unit-root tests available with `xtunitroot`, the header of the output contains a summary of the dataset's dimensions and the null and alternative hypotheses. First, consider the statistic labeled `t-bar`, which IPS denote as $t\text{-}bar_{NT}$. This statistic is appropriate when you assume that both N and T fixed; exact critical values reported in IPS (2003) are reported immediately to its right. Here, because $t\text{-}bar_{NT}$ is less than even its 1% critical value, we strongly reject the null hypothesis that all series contain a unit root in favor of the alternative that a nonzero fraction of the panels represent stationary processes.

The statistic labeled `t-tilde-bar` is IPS's $\widetilde{t}\text{-}bar_{NT}$ statistic and is similar to the $t\text{-}bar_{NT}$ statistic, except that a different estimator of the Dickey–Fuller regression error variance is used. A standardized version of this statistic, $Z_{\widetilde{t}\text{-}bar}$, is labeled `Z-t-tilde-bar` in the output and has an asymptotic standard

normal distribution. Here the p-value corresponding to Z-t-tilde-bar is essentially zero, so we strongly reject the null that all series contain a unit root.

◁

❑ Technical note

Just as the $Z_{\tilde{t}\text{-}bar}$ statistic corresponds to $\tilde{t}\text{-}bar_{NT}$, IPS present a $Z_{t\text{-}bar}$ statistic corresponding to $t\text{-}bar_{NT}$. However, the $Z_{t\text{-}bar}$ statistic does not have an asymptotic normal distribution, and so it is not presented in the output. $Z_{t\text{-}bar}$ is available in the stored results as r(zt).

❑

When serial correlation is present, we augment the Dickey–Fuller regression with further lags of the dependent variable:

$$\Delta y_{it} = \phi_i y_{i,t-1} + \mathbf{z}_{it}'\boldsymbol{\gamma}_i + \sum_{j=1}^{p} \Delta y_{i,t-j} + \epsilon_{it} \tag{6}$$

where the number of lags, p, is specified using the lags() option, and if the trend option is specified, we also include a time trend with panel-specific slope. You can either specify a number or have xtunitroot choose the number of lags for each panel by minimizing an information criterion. Here xtunitroot produces the IPS $W_{t\text{-}bar}$ statistic, which has an asymptotically standard normal distribution as $T \to \infty$ followed by $N \to \infty$. As a practical matter, this means you should have a reasonably large number of both time periods and panels to use this test.

Part of the computation of the $W_{t\text{-}bar}$ statistic involves retrieving expected values and variances of the t statistic for β_i in (6) in table 3 of IPS (2003). Because expected values have not been computed beyond $p = 8$ lags in (6), you cannot request more than 8 lags in the lags() option.

▷ Example 5

We again test whether lnrxrate contains a unit root for the subset of OECD countries, except we allow for serially correlated errors. We will choose the number of lags for the ADF regressions by minimizing the AIC criterion, subject to a maximum of 8 lags. We type

```
. xtunitroot ips lnrxrate if oecd, lags(aic 8) demean

Im-Pesaran-Shin unit-root test for lnrxrate
─────────────────────────────────────────────────────────────────────────
Ho: All panels contain unit roots            Number of panels  =     27
Ha: Some panels are stationary               Number of periods =     34

AR parameter: Panel-specific                 Asymptotics: T,N -> Infinity
Panel means:  Included                                        sequentially
Time trend:   Not included                   Cross-sectional means removed

ADF regressions: 1.48 lags average (chosen by AIC)
─────────────────────────────────────────────────────────────────────────
                        Statistic        p-value
─────────────────────────────────────────────────────────────────────────
    W-t-bar             -7.3075          0.0000
─────────────────────────────────────────────────────────────────────────
```

◁

Fisher-type tests

In our discussion of the IPS test, we intimated that the test statistics could be viewed as averages of bias-adjusted t statistics for each panel. As Maddala and Wu (1999, 635) describe the IPS test, "... the IPS test is a way of combining the evidence on the unit-root hypothesis from the N unit-root tests performed on the N cross-section units." Fisher-type panel unit-root tests make this approach explicit.

Meta-analysis, frequently used in biostatistics and medical sciences, is the combination of results from multiple studies designed to test a similar hypothesis in order to yield a more decisive conclusion. One type of meta-analysis, first proposed by R. A. Fisher, combines the p-values from independent tests to obtain an overall test statistic and is frequently called a Fisher-type test. See Whitehead (2002, sec. 9.8) for an introduction. In the context of panel data unit-root tests, we perform a unit-root test on each panel's series separately, then combine the p-values to obtain an overall test of whether the panel series contains a unit root.

`xtunitroot fisher` performs either ADF or Phillips–Perron unit-root tests on each panel depending on whether you specify the `dfuller` or `pperron` option. The actual tests are conducted by the `dfuller` and `pperron` commands, and you can specify to `xtunitroot fisher` any options those commands take; see [TS] **dfuller** and [TS] **pperron**.

`xtunitroot fisher` combines the p-values from the panel-specific unit-root tests using the four methods proposed by Choi (2001). Three of the methods differ in whether they use the inverse χ^2, inverse-normal, or inverse-logit transformation of p-values, and the fourth is a modification of the inverse χ^2 transformation that is suitable for when N tends to infinity. The inverse-normal and inverse-logit transformations can be used whether N is finite or infinite.

The null hypothesis being tested by `xtunitroot fisher` is that all panels contain a unit root. For a finite number of panels, the alternative is that at least one panel is stationary. As N tends to infinity, the number of panels that do not have a unit root should grow at the same rate as N under the alternative hypothesis.

▷ Example 6

Here we test for a unit root in `lnrxrate` using all 151 countries in our sample. We will use the ADF test. As before, we do not include a trend in real exchange rates and will therefore not specify the `trend` option. However, because the mean real exchange rate for any country is nonzero, we will specify the `drift` option. We will use two lags in the ADF regressions, and we will remove cross-sectional means by using `demean`. We type

```
. xtunitroot fisher lnrxrate, dfuller drift lags(2) demean

Fisher-type unit-root test for lnrxrate
Based on augmented Dickey-Fuller tests
```

| Ho: All panels contain unit roots | Number of panels | = | 151 |
| Ha: At least one panel is stationary | Number of periods | = | 34 |

AR parameter: Panel-specific	Asymptotics: T -> Infinity
Panel means: Included	
Time trend: Not included	Cross-sectional means removed
Drift term: Included	ADF regressions: 2 lags

		Statistic	p-value
Inverse chi-squared(302)	P	975.9130	0.0000
Inverse normal	Z	-19.6183	0.0000
Inverse logit t(759)	L*	-20.9768	0.0000
Modified inv. chi-squared	Pm	27.4211	0.0000

```
P statistic requires number of panels to be finite.
Other statistics are suitable for finite or infinite number of panels.
```

All four of the tests strongly reject the null hypothesis that all the panels contain unit roots. Choi's (2001) simulation results suggest that the inverse normal Z statistic offers the best trade-off between size and power, and he recommends using it in applications. We have observed that the inverse logit L^* test typically agrees with the Z test. Under the null hypothesis, Z has a standard normal distribution and L^* has a t distribution with $5N + 4$ degrees of freedom. Low values of Z and L^* cast doubt on the null hypothesis.

When the number of panels is finite, the inverse χ^2 P test is applicable; this statistic has a χ^2 distribution with $2N$ degrees of freedom, and large values are cause to reject the null hypothesis. Under the null hypothesis, as $T \to \infty$ followed by $N \to \infty$, P tends to infinity so that P has a degenerate limiting distribution. For large panels, Choi (2001) therefore proposes the modified inverse χ^2 P_m test which converges to a standard normal distribution; a large value of P_m casts doubt on the null hypothesis. Choi's simulation results do not reveal a specific value of N over which P_m should be preferred to P, though he mentions that $N = 100$ is still too small for P_m to have an approximately normal distribution.

◁

Hadri LM test

All the tests we have discussed so far take as the null hypothesis that the series contains a unit root. Classical statistical methods are designed to reject the null hypothesis only when the evidence against the null is sufficiently overwhelming. However, because unit-root tests typically are not very powerful against alternative hypotheses of somewhat persistent but stationary processes, reversing roles and testing the null hypothesis of stationarity against the alternative of a unit root is appealing. For pure time series, the KPSS test of Kwiatkowski et al. (1992) is one such test.

The Hadri (2000) LM test uses panel data to test the null hypothesis that the data are stationary versus the alternative that at least one panel contains a unit root. The test is designed for cases with large T and moderate N. The motivation for the test is straightforward. Suppose we include a panel-specific time trend (using the trend option with xtunitroot hadri) and write our series, y_{it}, as

$$y_{it} = r_{it} + \beta_i t + \epsilon_{it}$$

where r_{it} is a random walk,

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

and ϵ_{it} and u_{it} are zero-mean i.i.d. normal errors. If the variance of u_{it} were zero, then r_{it} would collapse to a constant; y_{it} would therefore be trend stationary. Using this logic, the Hadri LM test tests the hypothesis

$$H_0 : \lambda = \frac{\sigma_u^2}{\sigma_\epsilon^2} = 0 \qquad \text{versus} \qquad H_a : \lambda > 0$$

Two options to `xtunitroot hadri` allow you to relax the assumption that ϵ_{it} is i.i.d., though normality is still required. You can specify the `robust` option to obtain a variant of the test that is robust to heteroskedasticity across panels, or you can specify `kernel()` to obtain a variant that is robust to serial correlation and heteroskedasticity. Asymptotically, the Hadri LM test is justified as $T \to \infty$ followed by $N \to \infty$. As a practical matter, Hadri (2000) recommends this test for "large" T and "moderate" N.

▷ Example 7

We now test the null hypothesis that `lnrxrate` is stationary for the subset of OECD countries. To control for serial correlation, we will use a Bartlett kernel with 5 lags. We type

```
. xtunitroot hadri lnrxrate if oecd, kernel(bartlett 5) demean

Hadri LM test for lnrxrate
─────────────────────────────
Ho: All panels are stationary           Number of panels =      27
Ha: Some panels contain unit roots       Number of periods =     34

Time trend:           Not included       Asymptotics: T, N -> Infinity
Heteroskedasticity: Robust                                  sequentially
LR variance:          Bartlett kernel, 5 lags Cross-sectional means removed
─────────────────────────────────────────────────────────────────────
                   Statistic      p-value
─────────────────────────────────────────────────────────────────────
 z                   9.6473        0.0000
─────────────────────────────────────────────────────────────────────
```

We strongly reject the null hypothesis that all panels' series are stationary in favor of the alternative that at least one of them contains a unit root. In contrast, the previous examples generally rejected the null hypothesis that all series contain unit roots in favor of the alternative that at least some are stationary. For cautionary remarks on the use of panel unit-root tests in the examination of PPP, see, for example, Banerjee, Marcellino, and Osbat (2005). In short, our results are qualitatively quite similar to those reported in the literature, though Banerjee, Marcellino, and Osbat argue that because of cross-unit cointegration and long-run relationships among countries, panel unit-root tests quite often reject the null hypothesis even when true.

◁

Stored results

xtunitroot llc stores the following in r():

Scalars
r(N)	number of observations
r(N_g)	number of groups
r(N_t)	number of time periods
r(sig_adj)	standard deviation adjustment
r(mu_adj)	mean adjustment
r(delta)	pooled estimate of δ
r(se_delta)	pooled standard error of $\widehat{\delta}$
r(Var_ep)	variance of whitened differenced series
r(sbar)	mean of ratio of long-run to innovation standard deviations
r(ttilde)	observations per panel after lagging and differencing
r(td)	unadjusted t_δ statistic
r(p_td)	p-value for t_δ
r(tds)	adjusted t_δ^* statistic
r(p_tds)	p-value for t_δ^*
r(hac_lags)	lags used in HAC variance estimator
r(hac_lagm)	average lags used in HAC estimator
r(adf_lags)	lags used in ADF regressions
r(adf_lagm)	average lags used in ADF regressions

Macros
r(test)	llc
r(hac_kernel)	kernel used in HAC variance estimator
r(hac_method)	HAC lag-selection algorithm
r(adf_method)	ADF regression lag-selection criterion
r(demean)	demean, if the data were demeaned
r(deterministics)	noconstant, constant, or trend

xtunitroot ht stores the following in r():

Scalars
r(N)	number of observations
r(N_g)	number of groups
r(N_t)	number of time periods
r(rho)	estimated ρ
r(Var_rho)	variance of ρ under H_0
r(mean_rho)	mean of ρ under H_0
r(z)	z statistic
r(p)	p-value

Macros
r(test)	ht
r(demean)	demean, if the data were demeaned
r(deterministics)	noconstant, constant, or trend
r(altt)	altt, if altt was specified

`xtunitroot breitung` stores the following in `r()`:

Scalars

`r(N)`	number of observations
`r(N_g)`	number of groups
`r(N_t)`	number of time periods
`r(lambda)`	test statistic λ
`r(lrobust)`	robust test statistic λ_R
`r(p)`	p-value for λ
`r(p_lrobust)`	p-value for λ_R
`r(lags)`	lags used for prewhitening

Macros

`r(test)`	breitung
`r(demean)`	demean, if the data were demeaned
`r(robust)`	robust, if specified
`r(deterministics)`	noconstant, constant, or trend

`xtunitroot ips` stores the following in `r()`:

Scalars

`r(N)`	number of observations
`r(N_g)`	number of groups
`r(N_t)`	number of time periods
`r(tbar)`	test statistic $t\text{-}bar_{NT}$
`r(cv_10)`	exact 10% critical value for $t\text{-}bar_{NT}$
`r(cv_5)`	exact 5% critical value for $t\text{-}bar_{NT}$
`r(cv_1)`	exact 1% critical value for $t\text{-}bar_{NT}$
`r(zt)`	test statistic $Z_{t\text{-}bar}$
`r(ttildebar)`	test statistic $\widetilde{t}\text{-}bar_{NT}$
`r(zttildebar)`	test statistic $Z_{\underset{t\text{-}bar}{\sim}}$
`r(p_zttildebar)`	p-value for $Z_{\underset{t\text{-}bar}{\sim}}$
`r(wtbar)`	test statistic $W_{t\text{-}bar}$
`r(p_wtbar)`	p-value for $W_{t\text{-}bar}$
`r(lags)`	lags used in ADF regressions
`r(lagm)`	average lags used in ADF regressions

Macros

`r(test)`	ips
`r(demean)`	demean, if the data were demeaned
`r(adf_method)`	ADF regression lag-selection criterion
`r(deterministics)`	constant or trend

`xtunitroot fisher` stores the following in `r()`:

Scalars

r(N)	number of observations
r(N_g)	number of groups
r(N_t)	number of time periods
r(P)	inverse χ^2 P statistic
r(df_P)	P statistic degrees of freedom
r(p_P)	p-value for P statistic
r(L)	inverse logit L statistic
r(df_L)	L statistic degrees of freedom
r(p_L)	p-value for L statistic
r(Z)	inverse normal Z statistic
r(p_Z)	p-value for Z statistic
r(Pm)	modified inverse χ^2 P_m statistic
r(p_Pm)	p-value for P_m statistic

Macros

r(test)	fisher
r(urtest)	dfuller or pperron
r(options)	options passed to dfuller or pperron
r(demean)	demean, if the data were demeaned

`xtunitroot hadri` stores the following in `r()`:

Scalars

r(N)	number of observations
r(N_g)	number of groups
r(N_t)	number of time periods
r(var)	variance of z under H_0
r(mu)	mean of z under H_0
r(z)	test statistic z
r(p)	p-value for z
r(lags)	lags used for HAC variance

Macros

r(test)	hadri
r(demean)	demean, if the data were demeaned
r(robust)	robust, if specified
r(kernel)	kernel used for HAC variance
r(deterministics)	constant or trend

Methods and formulas

Methods and formulas are presented under the following headings:

> Levin–Lin–Chu test
> Harris–Tsavalis test
> Breitung test
>   Breitung test without trend
>   Breitung test with trend
> Im–Pesaran–Shin test
> Fisher-type tests
> Hadri LM test

We consider a simple panel-data model with a first-order autoregressive component:

$$y_{it} = \rho_i y_{i,t-1} + \mathbf{z}'_{it}\boldsymbol{\gamma}_i + \epsilon_{it}$$

where $i = 1, \ldots, N$ indexes panels and $t = 1, \ldots, T$ indexes time. For the IPS, Fisher-type, and Hadri LM tests, we instead have $t = 1, \ldots, T_i$, because they do not require balanced panels. ϵ_{it} is a zero-mean error term; we discuss the assumptions about ϵ_{it} for each test below. Here we use N to denote the number of panels, not the total number of observations. By default, $\mathbf{z}_{it} = 1$, so that the term $\mathbf{z}'_{it}\gamma_i$ represents panel-specific means (fixed effects). If noconstant is specified, $\mathbf{z}'_{it}\gamma_i$ vanishes. If trend is specified, $\mathbf{z}'_{it} = (1, t)$ so that $\mathbf{z}'_{it}\gamma_i$ represents panel-specific means and linear time trends.

Levin–Lin–Chu test

The starting point for the LLC test is the regression model

$$\Delta y_{it} = \phi y_{i,t-1} + \mathbf{z}'_{it}\gamma_i + \sum_{j=1}^{p_i} \theta_{ij}\Delta y_{i,t-j} + u_{it} \tag{7}$$

In $(1')$, LLC assume ϵ_{it} is independently distributed across panels and follows a stationary invertible process so that with sufficient lags of Δy_{it} included in (7), u_{it} will be white noise with potentially heterogeneous variance across panels. If lags(#) is specified with xtunitroot llc, then we set $p_i = \#$ for all panels $i = 1, \ldots, N$. Otherwise, we fit (7) for each panel individually for lags $1 \ldots p^{\max}$ and choose the lag length, p_i, that minimizes the information criterion requested by the user. During this step, we restrict estimation to the subset of observations that are valid when $p^{\max}$ lags are included. Information criteria are defined as follows:

$$\text{AIC} = (-2\ln L + 2k)/M$$
$$\text{BIC} = (-2\ln L + k\ln M)/M$$
$$\text{HQIC} = (-2\ln L + 2k\ln\ln M)/M$$

where $\ln L$ is the log likelihood assuming Gaussian errors, $M = T - p^{\max} - 2$, and k is the number of parameters in (7).

With the lag orders, p_i, in hand, the test proceeds in three main steps, the first of which is to use panel-by-panel OLS regressions to obtain the orthogonalized residuals

$$\widehat{e}_{it} = \Delta y_{it} - \sum_{j=1}^{p_i} \widehat{\theta}_{ij}\Delta y_{ij} - \mathbf{z}'_{it}\widehat{\gamma}_i \tag{8}$$

and

$$\widehat{v}_{i,t-1} = y_{i,t-1} - \sum_{j=1}^{p_i} \widetilde{\theta}_{ij}\Delta y_{ij} - \mathbf{z}_{it}\widetilde{\gamma}_i \tag{9}$$

To control for panel-level heterogeneity, compute

$$\widetilde{e}_{it} = \widehat{e}_{it}/\widehat{\sigma}_{\epsilon i} \qquad \text{and} \qquad \widetilde{v}_{i,t-1} = \widehat{v}_{i,t-1}/\widehat{\sigma}_{\epsilon i}$$

where

$$\widehat{\sigma}_{\epsilon i}^2 = \frac{1}{T - p_i - 1} \sum_{t=p_i}^{T} \left(\widehat{e}_{it} - \widehat{\delta}_i\widehat{v}_{i,t-1}\right)^2$$

and $\widehat{\delta}_i$ is the OLS coefficient from a regression of $\widehat{e}_{it}$ on $\widehat{v}_{i,t-1}$. If time trends are included (by specifying the trend option), then a linear time trend is included in regressions (7), (8), and (9).

In the second step, we estimate the ratio of long-run to short-run variances. Under the null hypothesis of a unit root, the long-run variance of the model without panel-specific intercepts or time trends ($z_{it} = \{\emptyset\}$) can be estimated as

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

where m is the maximum number of lags and $K(j,m)$ is the kernel weight function. Define $z = j/(m+1)$. If *kernel* is `bartlett`, then

$$K(j,m) = \begin{cases} 1-z & 0 \leq z \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

If *kernel* is `parzen`, then

$$K(j,m) = \begin{cases} 1 - 6z^2 + 6z^3 & 0 \leq z \leq 0.5 \\ 2(1-z)^3 & 0.5 < z \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

If *kernel* is `quadraticspectral`, then

$$K(j,m) = \begin{cases} 1 & z = 0 \\ 3\{\sin(\theta)/\theta - \cos(\theta)\}/\theta^2 & \text{otherwise} \end{cases}$$

where $\theta = 6\pi z/5$. If the user requests automatic bandwidth (lag) selection using the Newey–West algorithm, then we use the method documented in *Methods and formulas* of [R] **ivregress** with $\mathbf{z}_i = \mathbf{h} = 1$. If automatic lag selection with the LLC algorithm is chosen, then $m = \text{int}(3.21 T^{1/3})$.

If panel-specific intercepts are included (by not specifying `noconstant`), then in the formula for $\widehat{\sigma}_{yi}^2$ we replace Δy_{it} with $\Delta y_{it} - \overline{\Delta y_{it}}$, where $\overline{\Delta y_{it}}$ is the panel-level mean of Δy_{it} for panel i. Let $\widehat{s}_i = \widehat{\sigma}_{yi}/\widehat{\sigma}_{\epsilon i}$, and denote $\widehat{S}_N = N^{-1} \sum_i \widehat{s}_i$.

In the third step, we run the OLS regression

$$\widetilde{e}_{it} = \delta \widetilde{v}_{i,t-1} + \widetilde{\epsilon}_{it}$$

Called the "Basic test statistic" in the output of `xtunitroot llc` is the standard t statistic for δ computed as

$$t_\delta = \widehat{\delta}/\text{se}_{\widehat{(\delta)}}$$

where

$$\text{se}_{(\widehat{\delta})} = \widehat{\sigma}_{\widetilde{\epsilon}} \left(\sum_{i=1}^{N} \sum_{t=p_i+2}^{T} \widetilde{v}_{i,t-1}^2 \right)^{-1/2}$$

$$\widehat{\sigma}_{\widetilde{\epsilon}}^2 = \frac{1}{N\widetilde{T}} \sum_{i=1}^{N} \sum_{t=p_i+2} \left(\widetilde{e}_{it} - \delta \widetilde{v}_{i,t-1} \right)^2$$

and $\widetilde{T} = T - \overline{p} - 1$ with $\overline{p}$ the average of $p_1, \ldots, p_N$.

The adjusted test statistic is then computed as

$$t_{\delta}^* = \frac{t_{\delta} - N\widetilde{T}\widehat{S}_N \text{se}_{(\widehat{\delta})} \mu_{\widetilde{T}}^*}{\sigma_{\widetilde{T}}^*}$$

where $\mu_{\widetilde{T}}^*$ and $\sigma_{\widetilde{T}}^*$ are obtained by linearly interpolating the values in LLC (2002, table 2). t_{δ}^* is asymptotically $N(0, 1)$, with very negative values casting doubt on H_0. If `noconstant` is specified, then the asymptotic properties hold as $\sqrt{N}/T \to \infty$. Otherwise, T must grow at a faster rate so that $N/T \to \infty$.

Harris–Tsavalis test

The starting point for the HT test is (4), where ϵ_{it} is assumed to be i.i.d. normal with constant variance across panels. Denote by $\widehat{\rho}$ the least-squares estimate of ρ.

HT show that $\sqrt{N}(\widehat{\rho} - \mu) \xrightarrow{D} N(0, \sigma^2)$ as $N \to \infty$ with T fixed, where μ and σ^2 depend on the specification of the deterministic component:

Option	μ	σ^2
`noconstant`	1	$\frac{2}{T(T-1)}$
none	$1 - \frac{3}{T+1}$	$\frac{3(17T^2 - 20T + 17)}{5(T-1)(T+1)^3}$
`trend`	$1 - \frac{15}{2(T+2)}$	$\frac{15(193T^2 - 728T + 1147)}{112(T+2)^3(T-2)}$

Breitung test

Suppose the data are generated by an AR(1) process so that we can express y_{it} as

$$y_{it} = \mathbf{z}_{it}' \boldsymbol{\gamma}_i + x_{it}$$

where

$$x_{it} = \alpha_1 x_{i,t-1} + \alpha_2 x_{i,t-2} + \epsilon_{it}$$

where ϵ_{it} is an error term. A prewhitening step is available to correct for serial correlation. The nonrobust version assumes that ϵ_{it} is uncorrelated across panels, whereas the robust version allows for the panels to be contemporaneously correlated with covariance matrix $\boldsymbol{\Omega}$.

Under the null hypothesis that y_{it} contains a unit root, that is, that y_{it} is difference stationary, $\alpha_1 + \alpha_2 = 1$. Under the alternative that y_{it} is stationary, $\alpha_1 + \alpha_2 < 1$. Some of the time indices and summation limits of the formulas below appear more complex than those in Breitung (2000) and Breitung and Das (2005) because our formulas make explicit the loss of observations because of the prewhitening step.

Breitung test without trend

Let $y_{i,t}^\ell = y_{i,t-1} - y_{i,p+1}$ unless `noconstant` is specified, in which case let $y_{i,t}^\ell = y_{i,t-1}$. If the `lags()` option is specified with `xtunitroot breitung`, then we replace Δy_{it} and $y_{i,t}^\ell$ in the following description with the residuals from running regressions of Δy_{it} and $y_{i,t}^\ell$ on $\Delta y_{i,t-1}, \ldots, \Delta y_{i,t-p}$, where p is the lag order specified in `lags()`.

Define

$$\sigma_i^2 = \frac{1}{T-p-2} \sum_{t=p+2}^{T} (\Delta y_{it})^2$$

Then

$$\lambda = \frac{\sum_{i=1}^{N} \sum_{t=p+2}^{T} y_{it}^\ell \cdot \Delta y_{it} / \sigma_i^2}{\sqrt{\sum_{i=1}^{N} \sum_{t=p+2}^{T} (y_{it}^\ell)^2 / \sigma_i^2}}$$

λ is asymptotically distributed $N(0,1)$ as $T \to \infty$ followed by $N \to \infty$; small values of λ cast doubt on H_0.

For the robust version of the test statistic, let

$$\phi = \frac{\sum_{i=1}^{N} \sum_{t=p+2}^{T} y_{it}^\ell \cdot \Delta y_{it} / \sigma_i^2}{\sum_{i=1}^{N} \sum_{t=p+2}^{T} (y_{it}^\ell)^2 / \sigma_i^2}$$

and define $u_{it} = \Delta y_{it} - \phi y_{it}^\ell$. Let $\mathbf{u}_i = (u_{i,p+2}, \ldots, u_{iT})'$ and let the $N \times N$ matrix $\boldsymbol{\Omega}$ have typical element $\mathbf{u}_i'\mathbf{u}_j / (T-p-2)$. Let $\Delta \mathbf{y}_t = (\Delta y_{1t}, \ldots, \Delta y_{Nt})'$ and $\mathbf{y}_t^\ell = (y_{1,t-1}, \ldots, y_{N,t-1})'$. Then

$$\lambda_{\text{robust}} = \frac{\sum_{t=p+2}^{T} (\Delta \mathbf{y}_t)' \mathbf{y}_t^\ell}{\sum_{t=p+2}^{T} (\mathbf{y}_t^\ell)' \boldsymbol{\Omega} \mathbf{y}_t^\ell}$$

For $\boldsymbol{\Omega}$ to be positive definite, we must have $T - p - 1 \geq N$. As a practical matter, for $\boldsymbol{\Omega}$ to have good finite-sample properties, we need $T \gg N$. λ_{robust} is asymptotically distributed $N(0,1)$ as $T \to \infty$ followed by $N \to \infty$; very negative values of λ_{robust} cast doubt on H_0.

Breitung test with trend

Let p denote the number of lags requested in the `lags()` option. We fit the regression

$$\Delta y_{it} = \alpha_{i0} + \sum_{j=1}^{p} \alpha_{ij} \Delta y_{i,t-j} + \nu_{it}$$

and compute the $1 \times (T - p - 1)$ vectors $\Delta \mathbf{u}_i$ and $\mathbf{u}_i^\ell$ with typical elements

$$\Delta u_{is} = \Delta y_{is} - \sum_{j=1}^{p} \widehat{\alpha}_{ij} \Delta y_{i,s-j}$$

and

$$u_{is}^{\ell} = y_{i,s-1} - \sum_{j=1}^{p} \widehat{\alpha}_{ij} y_{i,s-j-1}$$

for $s = 1, \ldots, T - p - 1$. Let

$$\sigma_i^2 = \frac{1}{T - p - 2} \sum_{s=1}^{T-p-1} \left(\Delta u_{is} - \overline{\Delta u_i} \right) \Delta u_{is}$$

where $\overline{\Delta u_i}$ is the mean of Δu_{is} over s. Let $\Delta \mathbf{v}_i$ and $\mathbf{v}_i^{\ell}$ denote $1 \times (T - p - 1)$ vectors with typical elements

$$\Delta v_{is} = \sqrt{\frac{T - p - s - 1}{T - p - s}} \left(\Delta u_{is} - \frac{1}{T - p - s - 1} \sum_{j=s+1}^{T-p-1} \Delta u_{ij} \right)$$

and

$$v_{is}^{\ell} = u_{is}^{\ell} - u_{i1}^{\ell} - (T - p - 1)\overline{\Delta u_i}$$

Now

$$\lambda = \frac{\sum_{i=1}^{N} \sum_{s=1}^{T-p-1} v_{is}^{\ell} \Delta v_{is} / \sigma_i^2}{\sqrt{\sum_{i=1}^{N} \sum_{s=1}^{T-p-1} (v_{is}^{\ell})^2 / \sigma_i^2}}$$

λ is asymptotically distributed $N(0,1)$ as $T \to \infty$ followed by $N \to \infty$; very negative values of λ cast doubt on H_0. The computation of the robust form of the statistic proceeds in a fashion entirely analogous to the case without trend.

Im–Pesaran–Shin test

Write the model as

$$\Delta y_{it} = \phi_i y_{i,t-1} + \mathbf{z}_{it}' \gamma_i + \epsilon_{it}$$

where ϵ_{it} is independently distributed normal for all i and t with panel-specific variance σ_i^2. Denote $\Delta \mathbf{y}_i = (\Delta y_{i2}, \ldots, \Delta y_{iT})'$ and $\mathbf{y}_{i,-1} = (y_{i1}, \ldots, y_{i,T-1})'$. Note that to be consistent with the notation used in the rest of this documentation, we start the time index at $t = 1$ instead of $t = 0$ as in IPS (2003). Also let τ_T be a conformable vector of ones, $\mathbf{M}_{\tau} = \mathbf{I} - \tau_T(\tau_T' \tau_T)^{-1} \tau_T'$, $\mathbf{X}_i = (\tau_T, \mathbf{y}_{i,-1})$, and $\mathbf{M}_{X_i} = \mathbf{I} - \mathbf{X}_i(\mathbf{X}_i' \mathbf{X}_i)^{-1} \mathbf{X}_i'$.

First, we consider the case of no serial correlation, where the user does not specify the lags() option. Then

$$\widetilde{t}\text{-}bar_{NT} = \frac{1}{N} \sum_{i=1}^{N} \widetilde{t}_{iT}$$

where

$$\widetilde{t}_{iT} = \frac{\Delta \mathbf{y}_i' \mathbf{M}_{\tau} \mathbf{y}_{i,-1}}{\widetilde{\sigma}_{iT} \left(\mathbf{y}_{i,-1}' \mathbf{M}_{\tau} \mathbf{y}_{i,-1} \right)^{1/2}}$$

and

$$\widetilde{\sigma}_{iT}^2 = \frac{\Delta \mathbf{y}_i' \mathbf{M}_{\tau} \Delta \mathbf{y}_i}{T - 1}$$

Also

$$t\text{-}bar_{NT} = \frac{1}{N} \sum_{i=1}^{N} t_{iT}$$

where

$$t_{iT} = \frac{\Delta\mathbf{y}_i'\mathbf{M}_\tau\mathbf{y}_{i,-1}}{\widehat{\sigma}_{iT}\left(\mathbf{y}_{i,-1}'\mathbf{M}_\tau\mathbf{y}_{i,-1}\right)^{1/2}}$$

and

$$\widetilde{\sigma}_{iT}^2 = \frac{\Delta\mathbf{y}_i'\mathbf{M}_{X_i}\Delta\mathbf{y}_i}{T-1}$$

Now

$$Z_{\widetilde{t}\text{-}bar} = \frac{\sqrt{N}\left\{\widetilde{t}\text{-}bar_{NT} - N^{-1}\sum_{i=1}^{N} E(\widetilde{t}_{T_i})\right\}}{\sqrt{N^{-1}\sum_i \mathrm{Var}(\widetilde{t}_{T_i})}}$$

where $E(\widetilde{t}_{T_i})$ and $\mathrm{Var}(\widetilde{t}_{T_i})$ are obtained by linearly interpolating the values shown in IPS (2003, table 1). $Z_{\widetilde{t}\text{-}bar}$ has a standard normal limiting distribution for fixed T and $N \to \infty$; very negative values cast doubt on H_0. Similarly,

$$Z_{t\text{-}bar} = \frac{\sqrt{N}\left\{t\text{-}bar_{NT} - N^{-1}\sum_i E(t_{T_i})\right\}}{\sqrt{N^{-1}\sum_i \mathrm{Var}(t_{T_i})}}$$

If the `lags()` option is specified, then we fit the ADF regressions

$$\Delta y_{it} = \phi_i y_{i,t-1} + \mathbf{z}_{it}'\gamma_i + \sum_{j=1}^{p_i} \rho_{ij}\Delta y_{i,t-j} + \epsilon_i t$$

In matrix form, we can write this more compactly as

$$\Delta\mathbf{y}_i = \phi_i \mathbf{y}_{i,-1} + \mathbf{Q}_i\boldsymbol{\theta}_i + \epsilon_i$$

where $\mathbf{Q}_i = (\tau_t, \Delta\mathbf{y}_{i,-1}, \ldots, \Delta\mathbf{y}_{i,-p_i})$ and $\boldsymbol{\theta}_i = (\alpha_i, \rho_{i1}, \ldots, \rho_{ip_i})'$. Then

$$t\text{-}bar_{NT} = \frac{1}{N} \sum_{i=1}^{N} t_{iT}(p_i)$$

where

$$t_{iT}(p_i) = \frac{\sqrt{T - p_i - 2}(\mathbf{y}_{i,-1}'\mathbf{M}_{Q_i}\Delta\mathbf{y}_i)}{(\mathbf{y}_{i,-1}'\mathbf{M}_{Q_i}\mathbf{y}_{i,-1})^{1/2}(\Delta\mathbf{y}_{i,-1}'\mathbf{M}_{Q_i}\Delta\mathbf{y}_{i,-1})^{1/2}}$$

where $\mathbf{M}_{Q_i} = \mathbf{I} - \mathbf{Q}_i(\mathbf{Q}_i'\mathbf{Q}_i)^{-1}\mathbf{Q}_i'$, $\mathbf{M}_{X_i} = \mathbf{I} - \mathbf{X}_i(\mathbf{X}_i'\mathbf{X}_i)^{-1}\mathbf{X}_i'$, and $\mathbf{X}_i = (\mathbf{y}_{i,-1}, \mathbf{Q}_i)$. Finally,

$$\mathbf{W}_{t\text{-}bar}(p) = \frac{\sqrt{N}\left[t\text{-}bar_{NT} - N^{-1}\sum_{i=1}^{N} E\left\{t_{iT}(p_i)\right\}\right]}{\sqrt{N^{-1}\sum_{i=1}^{N} \mathrm{Var}\left\{t_{iT}(p_i)\right\}}}$$

where $E\left\{t_{iT}(p_i)\right\}$ and $\mathrm{Var}\left\{t_{iT}(p_i)\right\}$ are obtained by linearly interpolating the values shown in IPS (2003, table 3). $\mathbf{W}_{t\text{-}bar}(p)$ has a standard normal limiting distribution as $T \to \infty$ followed by $N \to \infty$; very negative values cast doubt on H_0.

Fisher-type tests

We use `dfuller` or `pperron` to perform unit-root tests on each panel; denote the p-value for the respective test on the ith panel as p_i. All of these tests are predicated on $T \to \infty$ so that the unit-root test for each panel is consistent. The P test is for finite N; the other tests are valid whether N is finite or infinite. Then

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

$P \sim \chi^2(2N)$ and large values cast doubt on H_0.

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

where $\Phi^{-1}(\ )$ is the inverse of the standard normal cumulative distribution function. $Z \sim N(0,1)$; very negative values of Z cast doubt on H_0.

$$L = \sum_{i=1}^{N} \ln\left(\frac{p_i}{1 - p_i}\right)$$

$L^* = \sqrt{k}L \sim t(5N + 4)$ where

$$k = \frac{3(5N + 4)}{\pi^2 N(5N + 2)}$$

Very negative values of L^* cast doubt on H_0. Finally,

$$P_m = -\frac{1}{\sqrt{N}} \sum_{i=1}^{N} \{\ln(p_i) + 1\}$$

$P_m \sim N(0,1)$; very positive values of P_m cast doubt on H_0.

Hadri LM test

As discussed in the main text, the Hadri LM test can be viewed as a test of $H_0 : \sigma_u^2/\sigma_\epsilon^2 = 0$, where both u_{it} and ϵ_{it} are normally distributed random errors.

Let $\widehat{\epsilon}_{it}$ denote the residuals from a regression of y_{it} on a panel-specific intercept or a panel-specific intercept and time trend if `trend` is specified. Then

$$\widehat{\text{LM}} = \frac{\frac{1}{N} \sum_i \frac{1}{T^2} \sum_t S_{it}^2}{\widehat{\sigma}_\epsilon^2} \tag{10}$$

where

$$S_{it} = \sum_{j=1}^{t} \widehat{\epsilon}_{ij}$$

and

$$\widehat{\sigma}_\epsilon^2 = \frac{1}{NT'} \sum_{i=1}^{N} \sum_{t=1}^{T} \widehat{\epsilon}_{it}^2$$

where $T' = T - 2$ if `trend` is specified and $T' = T - 1$ otherwise. Then

$$Z = \frac{\sqrt{N}\left(\widehat{\text{LM}} - \mu\right)}{\sigma}$$

where $\mu = 1/15$ and $\sigma = 11/6300$ if `trend` is specified and $\mu = 1/6$ and $\sigma = 1/45$ otherwise. $Z \sim N(0,1)$ asymptotically as $T \to \infty$ followed by $N \to \infty$. Very positive values of Z cast doubt on H_0. If `robust` is specified, then we instead use

$$\widehat{\text{LM}} = \frac{1}{N} \sum_{i=1}^{N} \left(\frac{\sum_{t=1}^{T} S_{it}^2}{T^2 \widehat{\sigma}_{\epsilon,i}^2} \right)$$

where we calculate $\widehat{\sigma}_{\epsilon,i}^2$ individually for each panel:

$$\widehat{\sigma}_{\epsilon,i}^2 = \sum_{t=1}^{T} \widehat{\epsilon}_{it}^2$$

If `kernel()` is specified, then we use (10) with

$$\widehat{\sigma}_{\epsilon}^2 = \frac{1}{N} \sum_{i=1}^{N} \left\{ \frac{1}{T} \sum_{t=p+1}^{T} \widehat{\epsilon}_{it}^2 + \frac{2}{T} \sum_{j=1}^{m} K(j,m) \sum_{t=j+1}^{T} \widehat{\epsilon}_{it}\widehat{\epsilon}_{i,t-j} \right\}$$

where m is the maximum number of lags and $K(.,.)$ is the kernel function defined previously.

Acknowledgments

References

Baltagi, B. H. 2013. *Econometric Analysis of Panel Data*. 5th ed. Chichester, UK: Wiley.

Banerjee, A., M. Marcellino, and C. Osbat. 2005. Testing for PPP: Should we use panel methods? *Empirical Economics* 30: 77–91.

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

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

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

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

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.

Heston, A., R. Summers, and B. Aten. 2006. Penn World Table Version 6.2. Center for International Comparisons of Production, Income, and Prices, University of Pennsylvania. http://pwt.econ.upenn.edu/php_site/pwt_index.php.

Hlouskova, J., and M. Wagner. 2006. The performance of panel unit root and stationarity tests: Results from a large scale simulation study. *Econometric Reviews* 25: 85–116.

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

Kwiatkowski, D., P. C. B. Phillips, P. Schmidt, and Y. Shin. 1992. Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series have a unit root? *Journal of Econometrics* 54: 159–178.

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.

Maddala, G. S., and S. Wu. 1999. A comparative study of unit root tests with panel data and a new simple test. *Oxford Bulletin of Economics and Statistics* 61: 631–652.

Newey, W. K., and K. D. West. 1994. Automatic lag selection in covariance matrix estimation. *Review of Economic Studies* 61: 631–653.

Nickell, S. J. 1981. Biases in dynamic models with fixed effects. *Econometrica* 49: 1417–1426.

Obstfeld, M., and K. Rogoff. 1996. *Foundations of International Macroeconomics*. Cambridge, MA: MIT Press.

O'Connell, P. G. J. 1998. The overvaluation of purchasing power parity. *Journal of International Economics* 44: 1–19.

Oh, K.-Y. 1996. Purchasing power parity and unit root tests using panel data. *Journal of International Money and Finance* 15: 405–418.

Patterson, K. 2000. *An Introduction to Applied Econometrics: A Time Series Approach*. Basingstoke, UK: Palgrave.

Phillips, P. C. B., and H. R. Moon. 1999. Linear regression limit theory for nonstationary panel data. *Econometrica* 67: 1057–1111.

——. 2000. Nonstationary panel data analysis: An overview of some recent developments. *Econometric Reviews* 19: 263–286.

Rogoff, K. 1996. The purchasing power parity puzzle. *Journal of Economic Literature* 34: 647–668.

Whitehead, A. 2002. *Meta-Analysis of Controlled Clinical Trials*. Chichester, UK: Wiley.

Also see