The Econometrics of Stability and Convergence

Andrew Harvey
Faculty of Economics, University of Cambridge

March 29, 2001

Abstract

This paper develops tests for stability and convergence and critically examines the way in which unit root tests have been used to assess convergence in the literature. Unobserved components models are used to characterise convergence in trends and to separate out cyclical movements.

KEYWORDS: Common trends; stationarity tests; unit roots; unobserved components.

JEL classification: C22, C32, O47

1. Introduction

There is a confusion between the econometric treatment of stability and convergence in the literature in that most of the time series tests for convergence are actually tests of stability. In many situations these tests are worse than useless as tests of convergence. The aim of the paper is to clarify the issues involved in testing for convergence and stability and to devise test procedures for relevant hypotheses.

Section 2 reviews tests for stationarity, while section 3 shows how unit root tests can be set up in such a way that their distributions under the null hypothesis belong to the same family as those of the stationarity tests. This provides a unified approach to testing hypotheses involving nonstationarity. Section 4 then extends the tests to multivariate series. It is pointed out that panel data methods are often invalid because of cross-correlations across series.

Issues of stability are addressed in section 5. Stability tests are needed for testing, amongst other things, the validity of purchasing power parity and the constancy of wage distributions over time. It is argued that stationarity tests are more appropriate than unit root tests in these situations since the null hypothesis is usually that the series are stable. Section 6 examines convergence and argues that the appropriate tests depend very much on what exactly it is required to test. For example it matters whether the aim is to test if two series have converged or if they are in the process of converging. Direct tests of convergence are proposed and it is shown that unit root tests are relevant provided that they are adapted to take account of the influence of initial conditions. The unit root tests are devised in the context of a model for the dynamics of convergence of two series. Section 7 shows how the model of the difference can be derived from a bivariate model for the levels of two converging economies.

2. Stationarity and unit root tests

2.1. Testing against the presence of a random walk component

Consider a univariate unobserved components model consisting of a random walk plus noise for a set of observations, y_t :

$$y_t = \mu_t + \varepsilon_t, \qquad \mu_t = \mu_{t-1} + \eta_t, \quad t = 1, ..., T, \tag{2.1}$$

where the $\eta_t' s$ and $\varepsilon_t' s$ are mutually and serially uncorrelated Gaussian disturbances with variances σ_η^2 and σ_ε^2 respectively. When $\sigma_\eta^2 = 0$ the random walk becomes a constant level. Nyblom and Mäkeläinen (1983) showed that the locally best invariant test (LBI) test of the null hypothesis that $\sigma_\eta^2 = 0$, against the alternative that $\sigma_\eta^2 > 0$, can be formulated as

$$\eta = T^{-2} \sum_{i=1}^{T} \left[\sum_{t=1}^{i} e_t \right]^2 / s^2 > c, \tag{2.2}$$

where $e_t = y_t - \overline{y}$, $s^2 = T^{-1} \sum_{t=1}^{T} (y_t - \overline{y})^2$ and c is a critical value. The test can also be interpreted as a one-sided Lagrange multiplier (LM) test. The asymptotic distribution of the statistic under the null hypothesis is the Cramér-von Mises distribution, denoted as CvM. If a linear time trend is included in (2.1) so that

$$y_t = \mu_t + \beta t + \varepsilon_t, \qquad t = 1, ..., T, \tag{2.3}$$

the asymptotic distribution of the test statistic, η_2, is a second level Cramér-von Mises distribution, denoted as $CvM_{2.}$ With neither constant nor trend, the distribution is $CvM_{0.}$ In the case of any ambiguity the standard CvM distribution will be referred to as CvM_1.

If the model is extended so that ε_t is any indeterministic stationary process the asymptotic distribution of the η test statistic remains the same if s^2 is replaced by a consistent estimator of the long-run variance σ_L^2. Kwiatkowski et al (1992) - KPSS - construct such an estimator nonparametrically. Leybourne and McCabe (1994) attack the problem of serial correlation by introducing lagged dependent variables into the model. The test statistic obtained after removing the effect of the lagged dependent variables is then of the same form as (2.2) and is asymptotically CvM under the null hypothesis. Since we are testing for the presence of an unobserved component it will often be natural to work with structural time series models as in Harvey and Streibel (1997). As in the Leybourne-McCabe test, the nuisance parameters in the STM need to be estimated in practice and this is best done under the alternative hypothesis. This has the compensating advantage that since there will often be some doubt about a suitable model specification, estimation of the unrestricted model affords the opportunity to check its suitability by the usual diagnostics and goodness of fit tests. Once the nuisance parameters have been estimated, the test statistic is calculated from a set of T innovations obtained with σ_η^2 set to zero.

3. Unit root tests

3.1. Unobserved components formulation

The Dickey-Fuller test is based on the model

$$y_t = \alpha + \beta t + \phi y_{t-1} + \xi_t, \qquad t = 1, ..., T \tag{3.1}$$

with variations in which the trend and both the constant and the trend are omitted. The null is that ϕ is unity, so the model is nonstationary, while the alternative is that it is less than unity, so the model is (trend) stationary. The test statistic is based on the regression coefficient of the lagged dependent variable. If the model is reformulated with Δy_t as the dependent variable, the parameter associated with y_{t-1}, and denoted here as ρ, is equal to $\phi - 1$ and hence is zero under the null hypothesis. Lagged differences can be added to the right hand side without affecting the asymptotic distribution of the estimator of ρ.

Formulating the unit root test in an autoregressive framework is computationally convenient. However, as Schmidt and Phillips (1992, p 258) observe, the parameterizations of (3.1) are "...not convenient..." because "...they handle level and trend in a clumsy and potentially confusing way." Specifically the meanings of α and β differ under the null and alternative hypotheses. These difficulties can be avoided by following Bhargava (1986), Nabeya and Tanaka (1990) and Schmidt and Phillips (1992) and setting up the unit root test of $H_0 : \phi = 1$ against $H_0 : \phi < 1$ within the components framework

$$y_t = \alpha + \beta t + \mu_t, \qquad \mu_t = \phi\mu_{t-1} + \eta_t, \quad t = 1, ..., T, \tag{3.2}$$

The interpretation of α and β is now the same under both the null and alternative hypotheses.

Schmidt and Phillips (1992) show that LM tests of the unit root hypothesis are based on the residuals obtained by estimating α and β under the null hypothesis. Since

$$\Delta y_t = \beta + \eta_t, \quad t = 2, ..., T \tag{3.3}$$

under the null hypothesis, these residuals are defined by

$$\widetilde{\mu}_t = y_t - \widetilde{\alpha}_0 - \widetilde{\beta}t, \qquad t = 1, ..., T$$

where $\widetilde{\beta} = \overline{\Delta y} = \sum \Delta y_t/(T-1) = (y_T - y_1)/(T-1)$ and $\widetilde{\alpha}_0 = y_1 - \widetilde{\beta}$, where $\alpha_0 = \alpha + \mu_0$. Note that $\widetilde{\mu}_1 = 0$, while $\widetilde{\mu}_T = 0$ provided a slope, β, is estimated. Schmidt and Phillips (1992) formulate their test in terms of a regression analogous to the one used in the Dickey-Fuller test, with y_{t-1} replaced by $\widetilde{\mu}_{t-1}$ and a constant but no time trend included. The tests are based on the regression coefficient of $\widetilde{\mu}_{t-1}$ or its 't-statistic'. A variant of the test, studied further in Schmidt and Lee (1991), excludes the constant.

Now consider the test with critical region

$$T^{-1}\sum_{t=1}^{T}\widetilde{\mu}_t^2 / \sum_{t=1}^{T}(\widetilde{\mu}_t - \widetilde{\mu}_{t-1})^2 = \zeta < c \tag{3.4}$$

where $\widetilde{\mu}_0$ is taken to be zero. This corresponds to the N_2 test suggested by Bhargava (1986) and to the variant of the Schmidt-Phillips test statistic obtained by regressing $\Delta\widetilde{\mu}_t$ on $\widetilde{\mu}_{t-1}$ without a constant term. The test statistic is the same as R_4 in Nabeya and Tanaka (1990), who show that the test is locally best invariant and unbiased (LBIU). If it is written in first differences it becomes

$$\zeta = T^{-1} \sum_{i=1}^{T} \left[\sum_{t=1}^{i} \Delta\widetilde{\mu}_t \right]^2 / \sum_{t=1}^{T} (\Delta\widetilde{\mu}_t)^2 \tag{3.5}$$

This is of the same form as the η test statistic, (2.2), except that it applies to observations in first differences. Provided the slope is estimated so that $\Delta\widetilde{\mu}_t = \Delta y_t - \overline{\Delta y}$ for $t = 2, .., T$, it is immediately apparent that the statistic has a CvM_1 distribution under the null hypothesis. However, while the value of the stationarity statistic, η, increases under the alternative, the value of ζ decreases as it is $T\zeta$ which has a limiting distribution under the alternative. Thus the appropriate critical values are those in the lower (left-hand) tail of the CvM distribution. At the 5% level these are 0.056 if no time trend is included and 0.037 if one is included. The right hand tail of ζ can be used to test against explosive processes, that is $\phi > 1$.

If there is no time trend in the model, $\overline{\Delta y}$ is omitted and the asymptotic distribution of the statistic is CvM_0. In this case it is useful to label the statistic ζ_1, and to denote the time trend statistic as ζ_2 when there is any ambiguity. If there is neither constant nor time trend, so that the statistic, ζ_0, is constructed by setting $\widetilde{\mu}_t = y_t$ for all $t = 0, 1, ..., T$ the asymptotic distribution is again CvM_0 provided.

The modified statistic

$$T^{-1} \sum (\widetilde{\mu}_t - \overline{\widetilde{\mu}})^2 / \sum (\widetilde{\mu}_t - \widetilde{\mu}_{t-1})^2, \tag{3.6}$$

where $\overline{\widetilde{\mu}}$ is the mean of the $\widetilde{\mu}_t' s$, is a transformation of the test statistic favoured by Schmidt and Phillips (1992) and the R_2 statistic in Bhargava (1986). It corresponds directly to R_3 in Nabeya and Tanaka (1990). Schmidt and Lee (1991) compare the tests based on (3.5) and (3.6) using Monte Carlo simulations and seem to come down in favour of (3.6) though the evidence is by no means clear-cut. Nabeya and Tanaka (1990), using an analysis based on limiting powers find that there is no dominance of one test over the other for the time trend model considered by Schmidt and Lee (1991). Furthermore, if a time trend is not present, then ζ is better. Further discussion can be found in Tanaka (1996, p348), where ζ is labelled R_2.

3.2. Serial correlation and additional unobserved components

Nabeya and Tanaka(1990) consider methods of adjusting the statistics (3.5) and (3.6) so that the same asymptotic distribution is obtained under the null hypothesis when η_t is serially correlated. They suggest using a nonparametric estimator of the long-run variance, constructed in a similar way as is done in the KPSS test. This corresponds to the KPSS statistic computed from first differences. However, under the alternative the spectrum of first differences is zero at the origin. Schmidt and Phillips (1992, p267) make a similar proposal but argue that a consistent test requires the use of residuals obtained (under the alternative hypothesis) from a Dickey-Fuller regression based on (3.1). Another option would be to use the coefficient of $\widetilde{\mu}_{t-1}$ from an augmented Dickey-Fuller regression.

If a fully parameterized UC model is set up, a pure LM test may be carried out. Evidence suggests that parametric test may have a more reliable size and higher power. ll that needs to be done is to estimate the model under the null hypothesis and then form a test statistic from the standardized innovations, $\widetilde{\nu}_t$, calculated starting with the smoothed estimator of μ_0 so they run from $t = 1$ to T. Assuming the innovations have been standardized so as to have unit variance, the unobserved components unit root test statistic is simply

$$\zeta = T^{-2} \sum_{i=1}^{T} \left[\sum_{t=1}^{i} \widetilde{\nu}_t \right]^2 \qquad (3.7)$$

Alternatively smoothing may be avoided by reversing the order of the observations and calculating (a different set of) innovations starting from the filtered estimator of μ_T.

The case for a parametric UC approach can be illustrated by the first-order AR with added white noise, that is

$$y_t = \alpha + \beta t + \mu_t + \varepsilon_t, \qquad t = 1, ..., T, \qquad (3.8)$$

The model in (3.8) is easily estimated when $\phi = 1$, for example by using the STAMP package of Koopman et al (2000). Hence forming the test statistic from the innovations as in (3.7) is easily done. (Note that if σ_ε^2 is zero, so that the model reduces to (3.2), then $\widetilde{\nu}_1 = 0$). Applying the Dickey-Fuller test when the data are best approximated by (3.8) is likely to result in too many rejections under the null hypothesis if the ratio of σ_η^2 to σ_ε^2 is low. The reduced form is an

ARIMA(0,1,1) model with MA parameter close to minus one and the poor performance of the augmented Dickey-Fuller test is well documented in this situation; see, for example, Pantula (1991).

The model in (3.8) may be generalised by including other components such as seasonals and cycles. Such models are easily estimated with ϕ set to one. The η statistic is computed from the innovations obtained from the Kalman filter by setting σ_η^2 to zero. Its aim is to determine whether a restriction should be placed on the model, while the ζ test is to find out if it should be more general.

4. Multivariate tests

4.1. Testing against a multivariate random walk

If $\mathbf{y}_t$ is a vector containing N time series the Gaussian multivariate local level model is

$$\begin{aligned} \mathbf{y}_t &= \boldsymbol{\mu}_t+\boldsymbol{\varepsilon}_t, \qquad \boldsymbol{\varepsilon}_t \sim NID(\mathbf{0}, \boldsymbol{\Sigma}_\epsilon), \\ \boldsymbol{\mu}_t &= \boldsymbol{\mu}_{t-1}+\boldsymbol{\eta}_t, \qquad \boldsymbol{\eta}_t \sim NID(\mathbf{0}, \boldsymbol{\Sigma}_\eta), \quad t=1,...,T, \end{aligned} \tag{4.1}$$

where $\boldsymbol{\Sigma}_\varepsilon$ is an $N \times N$ positive definite (p.d.) matrix. Nyblom and Harvey (2000) show that an LBI test of the null hypothesis that $\boldsymbol{\Sigma}_\eta = \mathbf{0}$ can be constructed against the homogeneous alternative $\boldsymbol{\Sigma}_\eta = q\boldsymbol{\Sigma}_\varepsilon$. The test has the rejection region

$$\eta(N) = \text{ tr } \left[\mathbf{S}^{-1}\mathbf{C}\right] > c, \tag{4.2}$$

where

$$\mathbf{C} = T^{-2}\sum_{i=1}^{T}\left[\sum_{t=1}^{i}\mathbf{e}_t\right]\left[\sum_{t=1}^{i}\mathbf{e}_t\right]' \quad \text{and} \quad \mathbf{S} = T^{-1}\sum_{t=1}^{T}\mathbf{e}_t\mathbf{e}_t'. \tag{4.3}$$

where $\mathbf{e}_t = \mathbf{y}_t - \overline{\mathbf{y}}$. Under the null hypothesis, the limiting distribution of (4.2) is Cramér-von Mises with N degrees of freedom, $CvM(N)$; see Harvey (2001, p13-4). The distribution is $CvM_2(N)$ if the model contains a vector of time trends.

The $\eta(N)$ test can be generalized along the lines of the KPSS test quite straightforwardly as in Nyblom and Harvey (2000). Parametric adjustments can also be made by the procedure outlined for univariate models. This requires estimation under the alternative hypothesis, but is likely to lead to an increase in power. If there are no constraints across parameters, it may be more convenient to construct the test statistic, (4.2), using the innovations from fitted univariate models. Alternatively, the lagged dependent variable method of Leybourne and McCabe (1994) may be used. This is the approach taken by Kuo and Mikkola (2001) in their study of purchasing power parity. They conclude that dealing with serial correlation in this way leads to tests with higher power than those formed using the nonparametric correction.

In the above tests no restrictions are put on the matrices $\mathbf{S}$ and $\mathbf{C}$. Similarly no restrictions are put on the covariance matrices of stationary components if more general UC models are estimated. If N is large this may be a problem and it may be necessary to assume some structure on the covariance matrices to reduce the number of parameters to be estimated. One possibility is to impose some spatial pattern. Panel methods may appear to offer a way out of this problem since if the units are mutually independent, the individual η statistics may be summed to give an overall test statistic which, by the central limit theorem is asymptotically normal under the null hypothesis. However, such statistics will be different from the above multivariate statistic unless $\mathbf{S}$ is diagonal and will be invalid with correlated units.

4.2. Testing for common trends

If the rank of $\mathbf{\Sigma}_\eta$ is K, the model has K common trends. Suppose we wish to test the null hypothesis that model (4.1) has a specific number of common trends against the alternative that it has more. Formally the test is of

$$H_0 : rank(\mathbf{\Sigma}_\eta) = K \qquad \text{against} \quad H_1 : rank(\mathbf{\Sigma}_\eta) > K, \quad K = 1, .., N-1. \quad (4.4)$$

Nyblom and Harvey (2000) show that a test of this hypothesis can be based on the sum of the $N-K$ smallest eigenvalues of $\mathbf{S}^{-1}\mathbf{C}$, denoted (K, N), and they tabulate significance points. Of course if we allow K to be zero, then $\eta(0, N) = \eta(N)$.

Parametric and nonparametric adjustment for serial correlation may be made in the common trends test in much the same way as was suggested for the $\eta(N)$ test.

4.3. Multivariate unit root tests

The UC model in (3.2) generalizes to

$$\mathbf{y}_t = \boldsymbol{\alpha} + \boldsymbol{\beta} t + \boldsymbol{\mu}_t, \qquad \boldsymbol{\mu}_t = \boldsymbol{\phi}\boldsymbol{\mu}_{t-1} + \boldsymbol{\eta}_t, \quad t = 1, ..., T, \qquad (4.5)$$

with $Var(\boldsymbol{\eta}_t) = \boldsymbol{\Sigma}_\eta$. As in the univariate case, residuals are formed by estimating the level and slope coefficients under the null hypothesis. Generalizing the test statistic (3.4) based on detrended observations yields

$$\zeta(N) = tr\left\{\frac{1}{T}\left[\sum_{t=1}^{T}\Delta\widetilde{\boldsymbol{\mu}}_t\Delta\widetilde{\boldsymbol{\mu}}_t'\right]^{-1}\sum_{t=1}^{T}\widetilde{\boldsymbol{\mu}}_t\widetilde{\boldsymbol{\mu}}_t'\right\} \tag{4.6}$$

where $\widetilde{\boldsymbol{\mu}}_t = \mathbf{y}_t - \widetilde{\boldsymbol{\alpha}}_0 - \widetilde{\boldsymbol{\beta}}t$ for $t = 1, .., T$ and $\widetilde{\boldsymbol{\mu}}_0 = \mathbf{0}$ with $\widetilde{\boldsymbol{\beta}} = (\mathbf{y}_T - \mathbf{y}_1)/(T-1)$ and $\widetilde{\boldsymbol{\alpha}}_0 = \mathbf{y}_1 - \widetilde{\boldsymbol{\beta}}$. Writing $\zeta(N)$ in a form analogous to (4.2) makes it apparent that its asymptotic distribution under the null hypothesis is $CvM_1(N)$ with the lower tail defining the critical region. If there is no time trend the critical values are taken from the $CvM_0(N)$ distribution.

Provided the slope is included, the above test corresponds to an LM test in a model in which $\boldsymbol{\phi} = \phi\mathbf{I}_N$ where ϕ is a scalar. The null hypothesis is $\phi = 1$ and the alternative is $\phi < 1$; see Harvey and Nyblom (2001).

If the model is generalised to include more components a parametric test statistic can be constructed from the vector of standardized innovations. Corresponding to (3.7) this statistic is

$$\zeta(N) = tr\left\{T^{-2}\sum_{i=1}^{T}\left[\sum_{t=1}^{i}\widetilde{\boldsymbol{\nu}}_t\right]\left[\sum_{t=1}^{i}\widetilde{\boldsymbol{\nu}}_t'\right]\right\} = T^{-2}\sum_{i=1}^{T}\left[\sum_{t=1}^{i}\widetilde{\boldsymbol{\nu}}_t'\right]\left[\sum_{t=1}^{i}\widetilde{\boldsymbol{\nu}}_t\right]. \tag{4.7}$$

If the innovations from fitted univariate models are used, the test statistic is of the form (4.2) so as to allow for cross-correlation.

4.4. Common trend unit root tests

If there are K common trends, the asymptotic distribution of $\zeta(N)$ will be $CvM(K)$. This suggests that a test of the null hypothesis that there are K common trends against the alternative that there are none could be carried out by comparing the K largest eigenvalues of $\zeta(N)$ with the lower tail of the $CvM(K)$ distribution. More generally, a test of the null hypothesis that there are K common trends against the alternative that there are $0 \leq M < K$ common trends could be based on the $K - M$ smallest of the K largest eigenvalues. Critical values for different combinations of M and K are tabulated in Harvey and Nyblom (2001). This unit root common trends test can also be interpreted as a test of the null hypothesis that there are $N-K$ co-integrating vectors against the alternative that there are $N-M$, that is $K-M$ more.

The $\zeta(M,K)$ test will be consistent against M common trends but will be inconsistent if there are more. Thus it may be safest to set $M = K-1$. In the special case when $K=1$, and the asymptotic distribution of the test statistic under the null is $CvM(1)$, the test can be regarded as a test of the null hypothesis that there is at least one common trend, that is fewer than N co-integrating vectors, against the alternative that all the series are stationary.

The tests extend to deal with more complex models either parametrically or nonparametrically. Parametric tests may be constructed using information on nuisance parameters obtained by estimating univariate models with nonstationary components. This is an enormous advantage since having to estimate a multivariate model may not be easy, particularly if N is not small.

5. Tests of Stability

Bivariate stability tests are testing whether two nonstationary series are evolving in such a way that their difference is stationary. Multivariate tests are testing whether all pairs have stationary differences. Formally these are co-integration tests where the co-integrating vector is known. An alternative view is that the series are exhibiting balanced growth with a common trend.

5.1. Stability test

A group of nonstationary series, $\mathbf{y}_t$, have a stable relationship over time if there is a full rank $(N-1)\times N$ matrix, $\mathbf{D}$, with no null columns and the property that $\mathbf{Di}=\mathbf{0}$, thereby rendering $\mathbf{Dy}_t$ jointly stationary. The implicit UC model for a set of stable series is

$$\mathbf{y}_t = \mathbf{i}\mu_t + \overline{\boldsymbol{\mu}} + \boldsymbol{\psi}_t, \qquad t = 1,...,T, \tag{5.1}$$

where μ_t is a univariate random walk (with drift), $\mathbf{i}$ is a vector of ones, $\overline{\boldsymbol{\mu}}$ is an $N\times 1$ vector with one zero entry and $\boldsymbol{\psi}_t$ is a stationary vector. This is sometimes called the balanced growth model and the rows of $\mathbf{D}$ may be termed balanced growth co-integrating vectors. Typically each row will contain a one, a minus one and zeroes elsewhere. For example, one country may be used as a benchmark or numeraire. In the context of the wage distribution data studied by Harvey and Bernstein (2000) the deciles are all compared with the median. Alternatively we might compare consecutive series. It is sometimes useful to work in terms of deviations from the cross-sectional mean with one of the series (in deviation form) dropped; in this case each row has one element equal to $(N-1)/N$, while the rest are $-1/N$.

The test of stability is carried out by applying the test of (4.2) to $\mathbf{Dy}_t$. Time trends would not normally be present under the null and so the test statistic is formed from deviations from the means and its limiting distribution under the null hypothesis is $CvM(N-1)$. Note that the choice of $\mathbf{D}$ is not crucial since pre-multiplication of $\mathbf{y}_t$ by a non-singular $(N-1)\times(N-1)$ matrix leaves the test statistic unchanged. Kuo and Mikkola (2001) argue that the $\eta(N)$ test is attractive for testing whether groups of countries exhibit purchasing power parity (PPP) because it accounts for cross-sectional correlation and is invariant to the benchmark currency.

Multivariate unit root tests are not particularly useful for testing the null of no stability - assuming this is felt to be a reasonable null hypothesis - since a test like $\zeta(N)$ can be shown to have a high probability of rejecting even when some of the series are nonstationary. Taylor and Sarno (1998) report similar findings for a multivariate generalisation of the ADF test in the context of a study of PPP. This suggests the use of $\zeta(0,1)$ to test the null hypothesis that there is at least one common trend against the alternative that all the series are stationary. The asymptotic distribution of $\zeta(0,1)$ under the null is $CvM(1)$ and, like $\eta(N)$, it is invariant to the choice of $\mathbf{D}$. Taylor and Sarno (1998) suggest a likelihood ratio test of this hypothesis in the context of PPP; the test requires estimation of a vector error correction model.

5.2. Identical trend

If there are no constant terms in (5.1), that is $\overline{\boldsymbol{\mu}}=\mathbf{0}$, the series contain an *identical* common trend. Stability tests can be carried out with this restriction imposed and the distribution of the $\eta(N)$ statistic is $CvM_0(N-1)$; see Hobijn and Franses (2000).

If a stable model is estimated, an LR or Wald test of the identical trends restriction can be readily carried out. However, the test may be done without fitting a model since the variance of the mean of $\mathbf{Dy}_t$ is $2\pi\mathbf{F}(0)$, where $\mathbf{F}(0)$ is the (multivariate) spectrum of $\mathbf{Dy}_t$ at frequency zero. This can be estimated nonparametrically as is in the multivariate extension of the KPSS test. Such a test should logically come after the $CvM_1(N-1)$ test has indicated stability.

5.3. Detection of stable clusters

If there is a single common trend, as in (5.1), pre-multiplication by the matrix of balanced growth co-integrating vectors, $\mathbf{D}$, gives $N-1$ stationary series. If there is more than one common trend, so that there are several stable clusters, the multivariate stationarity test statistic, $\eta(N-1)$, will tend to reject since at least one of the transformed series will be non-stationary. How many are nonstationary will depend on $\mathbf{D}$ but, as already noted, the value of $\eta(N-1)$ is invariant to the choice of $\mathbf{D}$. Furthermore, the common trends test statistics are also invariant to $\mathbf{D}$; see appendix. Thus if the hypothesis of overall stability is rejected, one may go on to test for the number of common trends. This test will be less efficient than one in which the actual groupings are specified. If there are K stable balanced clusters, each with its own trend, then knowledge of the clustering will suggest an $(N-K)\times N$ matrix of balanced co-integrating vectors, $\mathbf{D}_K$, which can be used to transform the observations. Under the null hypothesis that the specification of these K stable groupings is correct, the stationarity test statistic constructed from the transformed observations will be asymptotically distributed as CvM_{N-K}.

The unit root common trends tests are also invariant to $\mathbf{D}$. Again prior knowledge of possible clusters will suggest tests specifically for these clusters.

The tests do not identify the composition of clusters. If it is feasible to estimate a multivariate model, the restriction that there are a certain number of common trends may be imposed and the composition of clusters determined from the estimate of $\mathbf{\Sigma}_\eta$. An alternative strategy might be to use pairwise comparisons as the basis for an algorithm for determining clusters as in Hobijn and Franses(2000).

6. Convergence

There is serious confusion in the convergence literature regarding what exactly is being tested. This needs to be clarified at the outset. Once this has been done we can start to look at what are the most appropriate tools to use. For some hypotheses sensible tests can be constructed without formulating a model. For others, models play an essential role in allowing us to assess and understand what is happening and so formulate appropriate tests. However, they may also raise issues concerning whether tests are addressing the most interesting questions about convergence.

The discussion will initially be confined to two economies and attention will be focused on the difference in the logarithms, denoted y_t.

6.1. Definitions of convergence and their implications for testing

Two countries *have converged* if the difference between them is stable. If initial conditions are unimportant, stability implies that the difference between the series, y_t, is stationary for virtually the whole period. If the mean of y_t is zero the countries are in a state of *absolute convergence.* If the mean, α, is not zero we have *conditional* or *relative convergence*; this is a possibility if we entertain the existence of increasing costs of convergence and possible barriers to perfect convergence in the sense of (6.3). The limiting growth paths for the the regions are then parallel, differing by α.

Although they sometimes purport to be testing whether economies are converging, most 'convergence tests' are actually testing stability. This is consistent with the view of Bernard and Durlauf (1996 p 170) who state that the time series approaches to convergence check for the compatibility of the difference in (log) output with an indeterministic stationary series. Furthermore Bernard and Durlauf (1996 p 171) point out 'In time series tests, one assumes that the data are generated by economies near their limiting distributions and convergence is interpreted to mean that initial conditions have no (statistically significant) effect on the expected value of output differences.' In most applied studies the convergence tests are unit root tests, usually (augmented) Dickey-Fuller. However, if it is stability which is being tested, stationarity tests are arguably more appropriate.

If all that can be managed is stability tests, it is a pity, because this is not addressing any questions on the *process* of convergence. Data on countries often shows that they are converging, have just converged or have converged some time ago but still have a large opart of the series dependent on initial conditions.

Bernard and Durlauf (1996) offer a definition of 'convergence as catching up' over the period t to $t+\tau$. This is based on information, I_t, at time t and is

$$E(y_{t+\tau}|I_t) < y_t \tag{6.1}$$

Implementation of this definition requires a model, which can be univariate if I_t is just the series itself. However, if the model is thought of as being made up of a long-run underlying component, μ_t, and a short-run transient component, then the condition may be violated because of an unusually large short-run component at time t. Within a UC framework a better definition of catching up is

$$E(\mu_{t+\tau}|I_t) < E(\mu_t|I_t)$$

Catching up can be assessed after the event by defining convergence to have taken place over a particular time period if

$$E(\mu_{t+\tau}|I_T) < E(\mu_t|I_T), \quad t+\tau \leq T. \tag{6.2}$$

The next sub-section looks at how convergence in the sense of (6.2) can be assessed. The issue of whether two economies have converged absolutely can also be addressed by asking whether the estimate of $E(\mu_T|I_T)$ is significantly different from zero. The models used do not necessarily satisfy the long-run definition of convergence given in (6.3) below, though as will be seen in sub-section 6.3 there is a connection.

If a model is fitted, it will imply convergence to a fixed level, α, if

$$\lim_{j\to\infty} E(y_{t+j}|I_t) = \alpha \tag{6.3}$$

Convergence is absolute if $\alpha = 0$ as in Bernard and Durlauf (1996). The error correction models set up in sub-section 6.3 are able to satisfy this condition and they become stationary for economies which have converged. Within this modelling framework it can be seen that particular forms of unit root tests are appropriate for testing hypothesis concerning the process of convergence. Sub-sections 6.4 and 6.5 investigate the properties of these tests and suggest a completely new set of procedures.

6.2. Direct tests based on unobserved components models

The UC framework allows us to look at stylised facts without positing a particular mechanism for convergence. The difference, y_t, is assumed to be made up of a stochastic trend or level, μ_t, together with other components such as cycle, ψ_t, and irregular. Thus

$$y_t = \mu_t + \psi_t + \varepsilon_t, \qquad t = 1, ..., T, \tag{6.4}$$

with

$$\begin{aligned} \mu_t &= \mu_{t-1} + \beta_{t-1} + \eta_t, \quad t = 1, ..., T, \\ \beta_t &= \beta_{t-1} + \zeta_t, \end{aligned} \tag{6.5}$$

The smoothed estimates of the trend describe the time path reflecting the long-run difference between the two economies. Simply plotting this time path may be very informative. For example, figure 1 shows the difference in the trend of per capita GDP between the USA and Japan obtained by fitting a smooth trend, that is with σ_η^2 set to zero, plus cycle model using the STAMP package of Koopman et al (2000). We can go further and carry out tests of whether the gap between the two economies has narrowed significantly and/or whether the gap is zero, that is $\mu_T = 0$, indicating that absolute convergence has taken place. The result can be seen from the graph where a confidence interval of two RMSE's is shown. More precisely, the level in the trend at the end of the sample is -0.139 with a RMSE of 0.080 giving a '$t - value$' of -1.73. In fact it appears that Japan overtook the USA around 1990 and that convergence has taken place in the other direction since 1995.

The above test of absolute convergence was carried out simply by comparing the estimate of μ_T with its RMSE. This information is readily available from the output of the Kalman filter. More generally, if there are N series we model them all or we can model a set of $N-1$ contrasts, $\mathbf{Dy}_t$, where the matrix $\mathbf{D}$ is defined as in the sub-section on stability tests. A multivariate tests of the absolute convergence hypothesis, $\mathbf{D}\boldsymbol{\mu}_T = \mathbf{0}$, can be easily constructed and treated as having an asymptotic distribution under the null which is χ^2_{N-1}. For given values of the parameters in the model of all N series, the test will be invariant to the choice of $\mathbf{D}$.Relative convergence can be assessed by testing the null of no change in the difference between two economies after a certain point. A test of this kind requires that we assess the significance of the underlying change in the series between time $t = \tau$ and time $t = T$. This requires estimating the *temporal (level) contrast,* $\mu_\tau - \mu_T$, and its root mean square error (RMSE). The RMSE is most easily obtained by setting up a fixed-point smoother to calculate the estimator of the contrast, $\tilde{\mu}_{\tau|T} - \tilde{\mu}_{T|T}$; see Harvey and Bernstein (2001). A test of significance, possibly one-sided, can be carried out by comparing the *temporal contrast test statistic*

$$\frac{\tilde{\mu}_{\tau|T} - \tilde{\mu}_{T|T}}{RMSE(\tilde{\mu}_{\tau|T} - \tilde{\mu}_{T|T})} \tag{6.6}$$

with a standard normal distribution. The test may be generalised to multivariate series of differences.

The use of non-stationary models to test convergence is apparently contradictory since once convergence has taken place the series are stationary. The next sub-section sets out stable models for μ_t and argues that they will often be well-approximated by stochastic trends.

6.3. Modelling the dynamics of converging economies

An error correction mechanism (ECM) can be used to capture convergence dynamics. Time series tests of whether economies are converging can then be formulated within this framework. This reflects a view that modelling comes before testing and that testing in a vacuum is not very informative. Tests can be based on estimated error correction parameters, but, as we will see, such tests are closely related to the unit root tests earlier. However, the appropriate way to use unit root tests only becomes apparent within the context of the model.

We initially look at the simplest model with an ECM. The model is (3.2) but without the time trend, that is

$$y_t = \alpha + \mu_t, \qquad \mu_t = \phi\mu_{t-1} + \eta_t, \quad t = 1, ..., T, \tag{6.7}$$

with a fixed initial value, μ_0. The crucial point is that this is not constructed as a model of a stable contrast but rather as a model of transitional dynamics in a situation where the initial value is some way from zero. If $\phi < 1$, the gap tends to narrow over time. Indeed, when this holds, the model satisfies all the definitions of convergence given sub-section 6.1. Of course when the initial conditions have worked themselves out, the series becomes stationary.

The equivalent model

$$\Delta y_t = \rho(y_{t-1} - \alpha) + \eta_t, \quad t = 2, ..., T, \tag{6.8}$$

where $\rho = \phi - 1$, can be interpreted as saying that the expected growth rate in the current period is a negative fraction, ρ, of the gap between the two economies, corrected for the permanent difference, α. This accords with the notion of convergence in the cross-sectional literature.

The ECM may be generalised to allow for richer dynamics. Thus the AR form, (6.8), may be augmented with lagged values of differenced observations. OLS estimation is possible if no account is taken of the constraint in the constant term.

An alternative route is to formulate a UC model in which the error correction mechanism is in the trend. This avoids confounding transitional dynamics of convergence with short-term steady-state dynamics. Thus

$$y_t = \alpha + \mu_t + \psi_t + \varepsilon_t, \qquad \mu_t = \phi\mu_{t-1} + \eta_t, \quad t = 1, ..., T. \tag{6.9}$$

Estimation is effected using the state space form with a diffuse prior for μ_t (as though it were nonstationary) and α treated as a fixed parameter. Care must be taken as α is not identified when ϕ is unity; it may be wise to carry out numerical optimisation with a transformation, such as $-log\rho$ lying between 0 and ∞, which keeps ρ strictly less than one.

Smoother transitional dynamics can be achieved by specifying μ_t in (6.9) as a second-order ECM, that is

$$\begin{aligned}\mu_t &= \phi\mu_{t-1} + \beta_{t-1}, \quad t = 1, ..., T, \\ \beta_t &= \phi\beta_{t-1} + \zeta_t,\end{aligned}$$

or, letting $\rho = \phi - 1$,

$$\begin{aligned}\Delta\mu_t &= \rho\mu_{t-1} + \beta_t, \quad t = 1, ..., T, \\ \Delta\beta_t &= \rho\beta_{t-1} + \zeta_t,\end{aligned} \tag{6.10}$$

This is equivalent to an AR(2) process with both roots equal to ϕ. With a value of ϕ close to one μ_t will behave in a similar way to the smooth trend fitted to the Japan-US difference. On the other hand, the first-order ECM behaves rather like a random walk specification and tracks the observations closely, leaving little scope for the addition of short-term non-transitional components.

The first-order model can more easily be extended within an autoregressive framework, by adding lagged differences. Fitting such a model to the Japan-US series by OLS gave

$$\widehat{\Delta y_t} = -0.010y_{t-1} + 0.319\Delta y_{t-1} - 0.173\Delta y_{t-2} + 0.196\Delta y_{t-3} - 0.040\Delta y_{t-4},$$

with $\widehat{\sigma}_\eta = 0.045$. Without the lagged differences, the coefficient of y_{t-1} was -0.012 with $\widehat{\sigma}_\eta = 0.047$. The addition of a constant to the right hand side makes little difference to the results; as might be expected it was not significantly different from zero, its '$t-$statistic' being -0.78. However, the implied value of α, 0.6, is some way from zero.

6.4. Unit root tests of convergence

In the context of the models of the previous sub-section, a test of the null hypothesis that the economies are not converging, against the alternative that they are, is a test of $H_0 : \phi = 1$ against $H_1 : \phi < 1$ or, equivalently, $H_0 : \rho = 0$ against $H_1 : \rho < 0$.

First consider the case of absolute convergence so that α is set to zero and $\widehat{\rho}$ is the OLS estimator of ρ obtained by regressing Δy_t on y_{t-1}. When $\rho = 0$, the limiting distribution of the Dickey-Fuller (DF) test statistic $T\widehat{\rho}$ does not depend on y_0 and the 5% critical value is -8.039. However, while the limiting distribution is a good approximation to the finite sample distribution when $y_0 = 0,$ it becomes more concentrated as y_0/σ increases; see Evans and Savin (1981, p764). Thus using the asymptotic critical value will tend to give a conservative test. Based on the finite sample distribution, Evans and Savin (1981, p771, table IV) show that the power of the Dickey-Fuller test carried out using $T\widehat{\rho}$ (rather than its t-statistic) becomes higher as y_0/σ increases. This perhaps reflects the fact that forcing the regression line to pass through the origin when some of the values of y_t and y_{t-1} are large means that ϕ, and hence ρ, is estimated rather accurately. On the other hand, if a constant is included, the power of the Dickey-Fuller test tends to fall as y_0/σ increases and can even be less than the size; see Evans and Savin (1984, table VI). This is important since most of the literature on unit root testing for convergence is based on equations with constant terms (and sometimes even trends). It is therefore not surprising that researchers such as Bernard and Durlauf (1996, p 172) say,..'the time series results accepting the no convergence null may be due to transitional dynamics..' The point is clearly illustrated for the Japan-US series by the fact that the ADF statistic based on four lags is -1.064 which is actually on the 'wrong side' of the median; the 5% (asymptotic) critical value is -14.1. Dropping the constant gives a DF statistic of -1.925 falling to -1.540 when four lagged differences are included. Although this is relatively further to the left of the distribution it is still nowhere near the asymptotic 5% critical value, -8.04.

Even with the constant excluded, the evidence in Evans and Savin (1981) suggests that the power of the DF test based on ρ is low. As they say on p771 'The most striking feature of these power functions is the large value of T required to give reasonable power at alternatives near the null hypothesis, particularly for values less than unity.' Furthermore, problem with the DF test with no constant is that the small sample critical values depend on y_0 and no table is currently published. An alternative test can be based on the statistic

$$\zeta = T^{-1}\sum_{t=1}^{T} {y_t}^2/\sum_{t=2}^{T}(y_t - y_{t-1})^2 \quad or \quad \zeta = T^{-1}\sum_{t=2}^{T} {y_t}^2/\sum_{t=2}^{T}(y_t - y_{t-1})^2$$

This is CvM_0 under the null, but it will tend to be large when the series is in the process of converging. However, like DF, the small sample distribution will depend on the initial value.

How do we test for relative convergence? If we knew the gap, α, then we could simply subtract this from all the observations and carry out absolute convergence tests. Since α is normally not known, we might apply DF with a constant or we might take the final observation as indicative of α and subtract this from the observations. Thus we use $y_t - y_T, t = 1, ..., T$, but it can be shown that this alters the asymptotic DF distribution and in any case the idea does not generalise well to situations with more complex dymamics. An alternative strategy is to adapt the ζ statistic to

$$\zeta^{(T)} = T^{-1}\sum_{t}(y_t - y_T)^2/\sum_{t}(y_t - y_{t-1})^2.$$

All that has been done is to subtract y_T from the observations rather than y_1, and the asymptotic distribution is still CvM_0 under the null of no convergence. If, at the end of the sample, the series has only just converged or is in the process of converging it will appear like an explosive process going backwards. Hence $\zeta^{(T)}$ will tend to be large and the appropriate critical region is in the *upper* tail. On the other hand if convergence has taken place $\zeta^{(T)}$ will tend to become smaller the longer the series has been in a steady-state. This may not happen with the forward ζ test.

The ζ and $\zeta^{(T)}$ tests can be extended to deal with additional unobserved components as outlined in sub-section 2.3. Having estimated the short-term nuisance parameters in (6.9) the Kalman filter is applied under the null without α included. The estimator of the level at time T is then an estimator of $\alpha + \mu_T$ in the original model and if convergence is assumed to have taken place it is an estimator of α. The innovations for the reverse unit root test, $\zeta^{(T)}$, are then constructed by a backward filter starting from this estimator of the level at time T. Multivariate tests, based on generalising either ζ or $\zeta^{(T)}$ may also prove useful.

Testing in the presence of short-term dynamics can also be carried out using the ADF test. However, if a UC model is more appropriate the performance of

ADF tests may be unsatisfactory for reasons already documented in sub-section 3.2. Of course the converse may also be true.

6.5. Monte Carlo experiments

The usual ranges of values considered simulation experiments on in unit root tests may not appropriate for assessing their properties in the context of convergence. In the Japan-US case, regressing the difference on the lagged difference gives an estimate of ϕ of 0.988 with a disturbance standard deviation, σ_η, of 0.047. The corresponding estimate of ϕ obtained with the ADF(4) regression is similar, being 0.990, and the disturbance variance remains much the same if one simply fits a random walk. The initial (log) difference, in 1960, is 1.75, indicating that income per capita in the US was nearly six times the Japanese figure. Differences of 2, 1, .5 and .1 correspond to ratios of 7.39, 2.72, 1.65 and 1.11 respectively. For the Monte Carlo experiments we set σ_η to one, so rescaling the preceding four differences on the assumption that the original σ_η is 0.05 gives initial values, μ_0, of 40, 20, 10 and 2 respectively.

The expected difference in the growth rates of two economies is equal to $\rho = \phi - 1$ and so with $\phi = 0.98$ and a ratio of 1.65, the difference in growth rates is 1%. Some idea of what different values of ϕ imply about the closing of the gap can be obtained by noting that the $k-$step ahead forecast from an AR(1) model at $t = 0$ is ϕ^k times the initial value. Thus ϕ^k is the fraction of gap expected to remain after k time periods. As we shall see, the performance of various tests depends very much on the point reached in the convergence process as well as on the initial value. The results will be reported in a later version of this paper.

7. Models for converging economies

This section sets out a models for the levels of two converging economies and shows how it relates to the model of differences set up in sub-section 6.3.

A bivariate model for two converging economies can be set up in ECM form as

$$\begin{aligned} \Delta y_{1t} &= \rho_1(y_{1,t-1} - y_{2,t-1}) + \eta_{1t} \\ \Delta y_{2t} &= \rho_2(y_{2,t-1} - y_{1,t-1}) + \eta_{2t} \end{aligned} \tag{7.1}$$

where $y_{i,t}$ denotes, for example, per capita output for region i at time t. Absolute convergence is assumed for simplicity; otherwise a constant α has to be added or subtracted to the difference terms on the right hand side of (7.1). The model corresponds to the first-order vector autoregression

$$\begin{aligned} y_{1t} &= \phi_1 y_{1,t-1} + (1-\phi_1) y_{2,t-1} + \eta_{1t} \\ y_{2t} &= (1-\phi_2) y_{1,t-1} + \phi_2 y_{2,t-1} + \eta_{2t} \end{aligned}$$

where $\phi_i = 1 + \rho_i, i = 1, 2$. The roots of the transition matrix

$$\Phi = \begin{bmatrix} \phi_1 & 1-\phi_1 \\ 1-\phi_2 & \phi_2 \end{bmatrix}$$

are one and $\phi_1 + \phi_2 - 1$. The conditions for the second root to lie inside the unit circle are $0 \leq \phi_1 + \phi_2 \leq 2$ or, equivalently, $-2 \leq \rho_1 + \rho_2 \leq 0$.

If $\rho_1 + \rho_2 < 0$, then in the long-run the forecasts converge to the same value since, from theory of Markov chains,

$$\lim_{j\to\infty} \Phi^j = \begin{bmatrix} (1-\phi_2)/(2-\phi_1+\phi_2) & (1-\phi_1)/(2-\phi_1+\phi_2) \\ (1-\phi_2)/(2-\phi_1+\phi_2) & (1-\phi_1)/(2-\phi_1+\phi_2) \end{bmatrix} = \begin{bmatrix} \rho_2/(\rho_1+\rho_2) & \rho_1/(\rho_1+\rho_2) \\ \rho_2/(\rho_1+\rho_2) & \rho_1/(\rho_1+\rho_2) \end{bmatrix} \tag{7.2}$$

The model can be transformed to

$$\Delta y_{1t} - \Delta y_{2t} = \Delta(y_{1t} - y_{2t}) = (\rho_1 + \rho_2)(y_{1,t-1} - y_{2,t-1}) + \eta_{1t} - \eta_{2t}$$

$$\Delta y_{1t} + \Delta y_{2t} = \Delta(y_{1t} + y_{2t}) = (\rho_1 - \rho_2)(y_{1,t-1} - y_{2,t-1}) + \eta_{1t} + \eta_{2t}$$

The first equation corresponds to the univariate convergence equation of (6.8) and there is convergence if $\rho_1 + \rho_2$ is negative. If $\rho_1 - \rho_2 = 0$, the sum does not depend on the gap and is a random walk. This means that $\rho_1 = \rho_2 = \rho$ (and $\phi_1 = \phi_2 = \phi$) and so the second root of the transition matrix lies inside the unit circle if $-1 \leq \rho \leq 0$. As is clear from (7.2), the economies converge to the average of their two (current) levels. Another possibility is to set $\rho_2 = 0$ (or $\rho_1 = 0$) so that country one converges to country two (or vice versa).

Model (7.1) can be written in the equivalent form

$$\begin{aligned} \Delta y_{1t} &= 2\rho_1(y_{1,t-1} - \overline{y}_{t-1}) + \eta_{1t} \\ \Delta y_{2t} &= 2\rho_2(y_{2,t-1} - \overline{y}_{t-1}) + \eta_{2t} \end{aligned}$$

If $\rho_1 = \rho_2$ then $\overline{y}_t$ is a random walk. Subtracting the sum of the first differences from the individual equations in (7.1) leads to

$$\begin{aligned} \Delta(y_{1t} - \overline{y}_t) &= (\rho_1 + \rho_2)(y_{1,t-1} - \overline{y}_{t-1}) + (\eta_{1t} - \eta_{2t})/2 \\ \Delta(y_{2t} - \overline{y}_t) &= (\rho_1 + \rho_2)(y_{2,t-1} - \overline{y}_{t-1}) + (\eta_{2t} - \eta_{1t})/2 \end{aligned} \tag{7.3}$$

These two equations contain exactly the same information. A univariate test of convergence can be based on either one of them, but not both.

The model may be generalised in the UC framework by adding other stochastic components and introducing a constant α to allow for relative convergence. Thus

$$\begin{aligned} y_{1t} &= \mu_t^\beta + \mu_{1t} + \varepsilon_{1t} \\ y_{2t} &= \mu_t^\beta + \alpha + \mu_{2t} + \varepsilon_{2t} \end{aligned} \tag{7.4}$$

with the vector of levels $(\mu_{1t}, \mu_{2t})'$ modelled as in (7.1). The initial value of the common trend, μ_t^β, is set to zero so as to avoid confounding with the common level of μ_{1t} and μ_{2t}; ie the level to which they converge. The common trend may be a smooth stochastic trend, that is $\Delta^2\mu_t^\beta = \zeta_t$, though in the special case when it is deterministic it reduces to $\mu_t^\beta = \beta t$. The SSF is probably best set up by treating α and β as fixed and giving $(\mu_{1t}, \mu_{2t})'$ a diffuse prior.

Subtracting y_{2t} from y_{21t} in (7.4) gives a univariate model of the form (6.9). A second-order model for the convergence dynamics can be set up so as to give (6.10) for the series of differences. Multivariate models can also be constructed.

8. Conclusions

This paper has set out related classes of stationarity and unit root tests and extended them to deal with multivariate series. It was then argued that stationarity tests are appropriate for testing stability while unit root tests can be used, in a non-standard way, to test whether economies are in the process of converging. However, other types of time series tests may be more useful for answering certain questions on convergence.

A test of whether absolute convergence has taken place is best done by a direct test. Time series tests can be positively misleading. In any case, while it may be interesting to estimate the parameter ρ which indicates the rate of convergence, it is not clear that a test of the unit root hypothesis, $\rho = 0$, would serve any useful purpose. Further issues of interest might be whether ρ is constant as predicted by the theory underlying cross-sectional regressions of growth rates.

If absolute convergence has not taken place we may wish to test whether the economies are in the process of converging absolutely. In this case the unit root tests may have a useful role to play.

A direct test of whether relative convergence has taken place cannot be done since we do not know α. However, direct tests of whether convergence has ceases after a certain point can be carried out. We might also sensibly use unit root tests to see if convergence is taking place, although it would be prudent to actually estimate models of convergence as well.

The test procedures which derive from the LM principle all generalise to multivariate series and have the important property that they are independent of the type of transformations typically used to make comparisons between different economies.

Appendix

The eigenvalues of the $\mathbf{S}^{-1}\mathbf{C}$ matrix defined for (4.2) are unaffected by a non-singular transformation of the observations. Let the vector of observations be pre-multiplied by a square non-singular matrix $\mathbf{A}$. The eigenvalues of this matrix are obtained by solving the determinantal equation

$$\left|(\mathbf{ASA}')^{-1}\mathbf{ACA}' - \lambda\mathbf{I}\right| = 0$$

This may be rewritten

$$\left|\mathbf{A}'^{-1}\mathbf{S}^{-1}\mathbf{CA}' - \lambda\mathbf{A}'^{-1}\mathbf{A}'\right| = \left|\mathbf{A}'^{-1}\right|\left|\mathbf{S}^{-1}\mathbf{C} - \lambda\mathbf{I}\right|\left|\mathbf{A}'\right| = \left|\mathbf{S}^{-1}\mathbf{C} - \lambda\mathbf{I}\right|\left|\mathbf{A}'\right| / \left|\mathbf{A}'\right| = 0$$

Cancelling the determinants of $\mathbf{A}'$ leaves the determinant for obtaining the eigenvalues of the original matrix.

REFERENCES

Bernard, A.B. and S. Durlauf (1996). Interpreting tests of the convergence hypothesis, *Journal of Econometrics* 71, 161-73.

Bhargava, A.(1986), On the theory of testing for unit roots in observed time series, *Review of Economic Studies,* 53, 36-84.

Evans, G.B.A. and N.E.Savin (1981), Testing for unit roots 1, *Econometrica*, 49: 753-80.

Evans, G.B.A. and N.E.Savin (1984), Testing for unit roots 2, *Econometrica*, 52: 1241-70.

Harvey A.C. (2001). Testing in unobserved components models, *Journal of Forecasting,* **20,** 1-19.

Harvey, A.C. and J.Bernstein (2000). Measurement and testing of inequality from time series deciles with an application to US wages. Mimeo.

Harvey, A.C. and J. Nyblom (2001). Common trend and co-integration tests in an unobserved components framework, mimeo

Harvey, A.C. and M. Streibel (1997), Testing for nonstationary unobserved components, mimeo.

Hobijn, B. and P.H.Franses (2000), Asymptotically perfect and relative convergence of productivity. *Journal of Applied Econometrics*, 15: 59-81.

Koopman S.J., Harvey, A.C., Doornik, J.A and Shephard, N. (2000). *STAMP 6.0 Structural Time Series Analyser, Modeller and Predictor*, London: Chapman and Hall.

Kuo, B-S. and A. Mikkola (2001), how sure are we about purchasing power parity? Panel evidence with the null of stationary exchange rates. *Journal of Money Credit and Banking (to appear)*

Kwiatkowski, D., Phillips, P.C.B, Schmidt, P. 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* 44, 159-78.

Leybourne, S.J. and B.P.M. McCabe (1994), A consistent test for a unit root, *Journal of Business and Economic Statistics,* **12**, 157-66.

Nabeya, S. and K. Tanaka (1990), Limiting power of unit-root tests in time-series regression, *Journal of Econometrics,* **46,** 247-71.

Nyblom, J. and T. Mäkeläinen (1983), Comparison of tests for the presence of random walk coefficients in a simple linear model, *Journal of the American Statistical Association,* **78**, 856-64.

Nyblom, J., and A.C.Harvey (2000). Tests of common stochastic trends, *Econometric Theory,* 16, 176-99.

Pantula, S. (1991). Asymptotic distributions of unit root tests when the process is nearly stationary. *Journal of Business and Economic Statistics*, 9, 63-71.

Schmidt, P. and P.C.B. Phillips (1992): LM Tests for a Unit Root in the Presence of Deterministic Trends. *Oxford Bulletin of Economics and Statistics* 54, 257-287.

Schmidt, P. and J. Lee (1991): A modification of the Schmidt-Phillips unit root test, *Economics Letters,* **36,** 285-89.

Tanaka, K.(1996). *Time Series Analysis.* New York: John Wiley and Sons.

Taylor, M.P. and L. Sarno (1998), The behaviour of real exchange rates during the post-Bretton Woods period. *Journal of International Economics,* **46,** 281-312.