# TO WHAT EXTENT ARE US REGIONAL INCOMES CONVERGING? ♦

Mark J. Holmes
Department of Economics
Waikato University
New Zealand
holmesmj@waikato.ac.nz

Jesús Otero
Facultad de Economía
Universidad del Rosario
Colombia
jesus.otero@urosario.edu.co

Theodore Panagiotidis
Department of Economics
University of Macedonia
Greece
tpanag@uom.gr

June 2011

Abstract

Long-run income convergence is investigated in the context of US regional data. We employ a novel pair-wise econometric procedure based on a probabilistic definition of convergence. The idea behind this is that the time-series properties of all the possible regional income pairs are examined by means of unit root and non-cointegration tests where inference is based on the fraction of rejections. We distinguish between the cases of strong convergence, where the implied cointegrating vector is [1,-1], and weak convergence, where long-run homogeneity is relaxed. In order to address cross-sectional dependence, we employ a bootstrap methodology to derive the empirical distribution of the fraction of rejections. Overall, the evidence in favour of convergence at state-level is weak insofar as it is only based on cointegration without homogeneity. We find that the strength of convergence between states decreases with distance and initial income disparity. Using MSA-level data, the evidence for convergence is stronger.

JEL Classification: C2, C3, R1, R2, R3.

Keywords: Panel data, cross-section dependence, pair-wise approach, income, convergence.

♦ We would like to thank Jeremy Smith for useful comments and suggestions on an earlier version of this paper. We also like to thank participants at the 2011 meeting of the New Zealand Econometric Study Group with particular reference to Andrew Coleman and Peter Phillips, as well as participants at the 71st International Atlantic Economic Conference held in Athens, Greece. The usual disclaimer applies.

## 1. Introduction

In recent years, economists have keenly debated the neoclassical growth model prediction of per capita income convergence. A wide range of studies that includes early work by Barro (1991), Barro and Sala-i-Martin (1991,1992), Baumol (1986), Bernard and Durlauf (1996), Carlino and Mills (1993), Mankiw et al. (1992) and Sala-i-Martin (1996) has considered convergence across countries, US states and European regions and provided mixed evidence in favour of convergence. The empirical tests of the convergence hypothesis have been based on both cross-sectional and time-series approaches. The cross-sectional approach is often encapsulated in the notion of β-convergence, which requires that 'poor' regions grow faster than 'rich' ones. However, several criticisms have been raised against the conclusions reached in many of these studies on account of Galton's fallacy or 'regression towards the mean' (Quah, 1993). In contrast, the time series approach is built on a stochastic definition of convergence where the per-capita disparities are expected to be stationary. This is exemplified by studies such as Bernard and Durlauf (1996) who examine the number of common stochastic long-run trends in real per capita personal income. They find little evidence of long-run convergence among OECD countries.

Using per capita income data across many decades, a number of studies have examined stochastic convergence in the case of the US states. As pointed out by Choi (2004) and others, due to the almost homogeneous institutional environments and the highly integrated markets for products and factors, the US states satisfy the underlying conditions of the convergence hypothesis in the standard neoclassical growth model. The existing evidence, however, is mixed. For example, Carlino and Mills (1993) find in favour of stochastic convergence insofar as shocks to relative regional per capita income are temporary, but only after allowing for a structural break in 1946. Evans and Karras (1996) employ a panel unit root test based on Levin et al. (2002). While this approach only allows for fixed effects and common slopes, they reject the null hypothesis of joint non-stationarity of relative per capita incomes. Tsionas (2001) employs vector error correction modelling and finds that multiple common trends are driving the income series thereby concluding against the convergence of real per capita incomes. Choi (2004) applies multiple panel data techniques to state per capita output and finds that output convergence in the United States has proceeded among geographically neighbouring states rather than among distant states, notwithstanding the nearly complete integration of product and factor markets. More recently, Mello (2011) examines relative incomes and considers whether low power of unit root tests as well as high persistence have led researchers to find evidence against convergence. Using a methodology based on fractional integration and interval estimation, support is found for stochastic convergence.

In this paper, we contribute to the debate concerning long-run income convergence among US states. In doing so, we analyse the interaction between non-stationary state income series and draw on the time series approach, but in a way that also utilises cross-sectional information. The novelty of our approach is the adoption and development of an econometric procedure advocated by Pesaran (2007). Within this framework, a probabilistic definition of regional convergence is proposed and forms the basis of our empirical testing strategy. This is an important deviation from the stochastic definition of convergence that the literature has focused on so far. The idea behind this is that for a sample of $N$ states, unit root and non-cointegration tests are conducted on all $N(N-1)/2$ real per-capita income differentials and pairs. Under the null hypothesis of non-stationarity or non-convergence, one would normally expect the fraction of real per capita income differentials or pairs for which the unit-root or non-cointegration hypothesis is rejected to be close to the size of the underlying unit-root or non-cointegration tests, which we denote as $\alpha$. However, we can argue that the null of non-stationarity or non-cointegration for all state pairs could be rejected if the fraction of rejections exceeds the chosen nominal size $\alpha$. The presence of cross-sectional dependence can make inference based on the fraction of rejections difficult, so we employ a bootstrap methodology to derive the empirical distribution of the fraction of rejections.

There already exist a very limited number of studies that investigate stochastic convergence using a pair-wise approach. Pesaran (2007) considers data for 101 countries and geographical sub-groups within. Relying on the use of pair-wise unit root tests provides little evidence of convergence at a global level, though there is some evidence of club convergence (Quah, 1997). Mello (2011) considers the case of income convergence across 48 contiguous US states. The pair-wise unit root testing procedure indicates that the non-stationary null is rejected in 8.6% of the cases where $\alpha$ is set equal to 5%. Le Pen (2011) offers a pair-wise study of output convergence between 195 European regions. While this particular study integrates structural breaks into the analysis, the evidence is not supportive of stochastic convergence. However, in the examination of income convergence, none of these studies consider an empirical distribution of the fraction of rejections, nor do they conduct an analysis of pair-wise cointegration (relaxing the [1, -1] assumption for the cointegrating vector). In our investigation of regional income convergence in the US, we address both of these important issues.

The paper is organised as follows. The following section describes the pair-wise approach for convergence. The initial paper by Pesaran (2007) only considers pair-wise unit root tests. In this paper, we introduce and develop a pair-wise cointegration test based on the application of the Johansen maximum likelihood test on all regional income pairs. It is argued that a pair-wise approach towards convergence testing offers important advantages over existing panel data unit root and cointegration methods in terms of addressing the proportion of the sample that is stationary or cointegrated, the presence of cross-sectional dependence across states, and the selection of a base or reference state. The third section discusses the data and the results of the empirical analysis. While the pair-wise unit root testing is not supportive of long-run convergence among 48 US states, our pair-wise cointegration approach provides some weaker evidence. We find that the strength of convergence, as measured by the long-run slope coefficient, is negatively related to both distance and initial income disparity. When we consider a more disaggregated dataset for 346 Metropolitan Statistical Areas (MSAs), then stronger evidence of convergence is also found. The final section offers some concluding remarks.

## 2. A pair-wise approach to testing for convergence

The unit root and cointegration tests employed in the past to assess stochastic convergence have typically applied to regional income benchmarked against national income. However this approach could be sensitive to the choice of base region or state. For example, real per capita income in states *i* and *j* might be found as non-stationary when measured against a third numeraire state *k*, but stationary when measured against one another. This would be the case when there is a highly persistent factor that is common to states *i* and *j*, but that is not shared by state *k*. The pair-wise methodology considers all possible bivariate relationships and does not involve what can be a problematic choice of a single reference state across the sample.

For the purposes of our empirical analysis we employ the Pesaran (2007) pair-wise testing procedure to analyse probabilistic convergence across a large number of cross section units. Following Pesaran (2007), let $y_{it}$ be real per capita income data in US state $i$ at time $t$, where $i = 1,\ldots,N$ and $t = 1,\ldots,T$. Pesaran's pair-wise approach is based on the examination of the time series properties of all $N(N-1)/2$ possible real per capita income gaps (or differentials) between states $i$ and $j$, which we denote as $g_{ijt} = y_{it} - y_{jt}$, where $i = 1,\ldots,N-1$ and $j = i+1,\ldots,N$. Consider next the application of the augmented Dickey and Fuller (ADF) (1979) or the Elliott, Rothenberg and Stock (ERS) (1996) unit root tests of order $p$ to $g_{ijt} = y_{it} - y_{jt}$, and let $Z_{ij,T}$ be an indicator function equal to one if the corresponding unit-root test statistic is rejected at significance level $\alpha$. More formally, in the case of the ADF test, $Z_{ij,T} = 1$ if $ADF(p) < K_{T,p,\alpha}$, where $ADF(p)$ is the calculated test statistic of order $p$, $K_{T,p,\alpha}$ is the critical value for the $ADF(p)$ of size $\alpha$, using $T$ observations. Similarly, when applying the ERS test, we would have $Z_{ij,T} = 1$ if $ERS(p) < K_{T,p,\alpha}$.

Pesaran (2007) considers the fraction of the $N(N-1)/2$ gaps for which the unit-root hypothesis is rejected, which is given by:

$$\bar{Z}_{NT} = \frac{2}{N(N-1)} \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} Z_{ij,T} , \tag{1}$$

and shows that under the null hypothesis of non-stationarity, the expected value of $\bar{Z}_{NT}$ is equal to the nominal size of the unit root test statistic, more formally:

$$\lim_{T \to \infty} E\left(\bar{Z}_{NT} \middle| H_o\right) = \alpha . \tag{2}$$

In the case of a unit-root test (such as ADF or ERS), under the null hypothesis of convergence, one would expect the proportion of rejections to be high and tending towards 100% as $T \to \infty$; analogously, under the divergence alternative the proportion of rejections ought to be low and around the nominal size of the test $\alpha$. The pair-wise approach is applicable when $N$ is large relative to $T$ and it can be shown that under the null of non-convergence, the fraction of the rejections converges to $\alpha$, as $N, T \to \infty$. Pesaran (2007) indicates that there are some difficulties involved in developing a formal procedure to test whether the proportion of rejections $\bar{Z}_{NT}$ is statistically different from $\alpha$, because the derivation of the variance of $\bar{Z}_{NT}$ is complicated due to the fact that $Z_{ij,T}$ and $Z_{ik,T}$ are not independent from each other. Thus, in order to overcome this difficulty inference on $\bar{Z}_{NT}$ can be based on the derivation of the empirical distribution of the fraction of rejections using the bootstrap methodology.

The implementation of the bootstrap is not an issue considered in Pesaran (2007), but in a subsequent paper by Pesaran, Smith, Yamagata and Hvozdyk (PSYH) (2009) when applying the pair-wise approach to test for purchasing power parity. More specifically, the model setup considered by these authors is based on the following set of equations:

$$y_{it} = \boldsymbol{\alpha}_i^{'}\mathbf{d}_t + \boldsymbol{\gamma}_i^{'}\mathbf{f}_t + \varepsilon_{it} \tag{3}$$

$$\Delta\varepsilon_{it} = \eta_i + \lambda_i\varepsilon_{i,t-1} + \sum_{l=1}^{p_i}\psi_{il}\Delta\varepsilon_{i,t-l} + \upsilon_{it} \tag{4}$$

$$\Delta f_{st} = \boldsymbol{\mu}_s^{'}\mathbf{d}_t + \phi f_{s,t-1} + \sum_{l=1}^{p_s}\xi_{sl}\Delta f_{s,t-l} + e_{st} \tag{5}$$

where $s = 1, 2, \ldots, m$ is the number of common factors, $\mathbf{d}_t = (1, t)^{'}$ denotes a vector of deterministic components that includes intercept, and intercept and trend, $\mathbf{f}_t$ is a $m \times 1$ vector of unobserved common factors, with elements denoted $f_{st}$, and $\varepsilon_{it}$ denotes the corresponding idiosyncratic elements. The unobserved common factors $f_{st}$ and/or the idiosyncratic elements $\varepsilon_{it}$ may be $I(0)$ or $I(1)$.

Following PSYH, we use the cross-sectional average of $y_{it}$, denoted $\overline{y}_t = N^{-1}\sum_{i=1}^{N} y_{it}$, as an estimate of the common factor that may induce cross-section dependence across state incomes. To account for cross-section dependence, real per capita income in each state is then regressed on the estimated common factor, that is:

$$y_{it} = \hat{\alpha}_i + \hat{\delta}_i t + \hat{\gamma}_i\overline{y}_t + \hat{\varepsilon}_{it}, \tag{6}$$

where the trend term is included if the corresponding estimated coefficient, $\hat{\delta}_i$, is statistically significant. Appendix 1, at the end of the document, reports the results of estimating the factor equations for the dataset used in the paper (details of which are provided in the next section). It should be noted that in these factor equations the linear trend term is included if it turns out to be statistically significant at the 5% level.

The next step is to examine the time series properties of the estimate of the common factor $\bar{y}_t$, which may be $I(0)$ or $I(1)$. This involves estimating the following $\mathrm{ADF}(p)$ regression for $\bar{y}_t$:

$$\Delta \bar{y}_t = \hat{\mu} + \hat{\phi} \bar{y}_{t-1} + \sum_{l=1}^{p} \hat{b}_l \Delta \bar{y}_{t-l} + \hat{e}_t, \tag{7}$$

which may also include a trend term if it is statistically significant, and where the optimal number of lags of the dependent variable $p$ may be determined e.g. using the Akaike information criterion (AIC). To illustrate the implementation of the bootstrap, let us consider for instance the case in which $\bar{y}_t$ has a unit root with a drift and no deterministic trend. Imposing a unit root on (7) and allowing for a drift, which is equivalent to setting $\hat{\phi} = 0$, implies the following restricted version of (7):

$$\Delta \bar{y}_t = \hat{\mu} + \sum_{l=1}^{p} \hat{c}_l \Delta \bar{y}_{t-l} + \hat{u}_t . \tag{8}$$

Thus, when a unit root and a drift term are imposed on the factor $\bar{y}_t$, the bootstrap samples of $\bar{y}_t$, denoted $\bar{y}_t^{(b)}$, can be computed using the following generating mechanism:

$$\bar{y}_t^{(b)} = \hat{\mu} + \bar{y}_{t-1}^{(b)} + \sum_{l=1}^{p} \hat{c}_l \Delta \bar{y}_{t-l}^{(b)} + \hat{u}_t^{(b)}, \tag{9}$$

where bootstrap residuals $\hat{u}_t^{(b)}$ are generated by randomly drawing with replacement from the set of estimated and centred residuals $\hat{u}_t$ in (8), and where the first $(p+1)$ values of $\bar{y}_t$ are used to initialise the process $\bar{y}_t^{(b)}$.

In turn, the bootstrap samples of $y_{it}$, denoted as $y_{it}^{(b)}$, are generated as:

$$y_{it}^{(b)} = \hat{\alpha}_i + \hat{\delta}_i t + \hat{\gamma}_i \overline{y}_{it}^{(b)} + \hat{\varepsilon}_{it}^{(b)}, \tag{10}$$

where $\hat{\alpha}_i$, $\hat{\delta}_i$ and $\hat{\gamma}_i$ are the OLS estimates of $\alpha_i$, $\delta_i$ and $\gamma_i$ in (6), respectively, and

$$\varepsilon_{it}^{(b)} = \hat{\eta}_i + \left(1 + \hat{\lambda}_i\right) \varepsilon_{i,t-1}^{(b)} + \sum_{l=1}^{p_i} \hat{\psi}_{il} \Delta \varepsilon_{i,t-1}^{(b)} + \upsilon_{it}^{(b)}, \tag{11}$$

where bootstrap residuals $\upsilon_{it}^{(b)}$ are generated by randomly drawing with replacement from the set of estimated residuals $\upsilon_{it}$ in equation (4), and the first $(p+1)$ values of $\hat{\varepsilon}_{it}$ are used to initialise the process $\varepsilon_{it}^{(b)}$. The AIC is used to select the optimal lag order $p_i$.

Having obtained $y_{it}^{(b)}$, it is possible to compute all possible bootstrap income gaps (or differentials) between states $i$ and $j$, that is $g_{ijt}^{(b)} = y_{it}^{(b)} - y_{jt}^{(b)}$, so that one can then calculate the fraction of these income gaps for which the unit root hypothesis can be rejected using either the $\mathrm{ADF}(p)$ or $\mathrm{ERS}(p)$ test. The procedure already described is repeated $b = 1, \ldots, B$ times to derive the empirical distribution of the bootstrapped fraction of rejections.

The pair-wise approach outlined above implicitly assumes that all income pairs are cointegrated with a known cointegrating vector equal to $[1, -1]'$. In practice, this might be regarded as a somewhat strong assumption. For this reason, we extend Pesaran's pair-wise approach by considering a test for a weaker form of convergence, according to which income pairs are cointegrated with an unknown cointegrating vector. In other words, instead of testing whether $g_{ijt} = y_{it} - y_{jt}$ is stationary, one could alternatively test whether $y_{it}$ and $y_{jt}$ are cointegrated. There are several cointegration tests available in the literature, including single equation methods such as two-step OLS (Engle and Granger, 1987), fully modified OLS (Phillips and Hansen, 1990), and dynamic OLS (Saikkonen, 1992; Stock and Watson, 1993), and system methods such as the maximum likelihood estimator of cointegrated vector autoregressive (VAR) models (Johansen, 1988). In this paper we opt for the Johansen procedure, which offers the advantage that normalisation on a particular state within each bivariate relationship is not an issue.[1] Thus, the pair-wise Johansen approach that we propose here is very much in the spirit of the Pesaran's pair-wise approach, in the sense that it does not matter whether one looks at cointegration between $y_{it}$ and $y_{jt}$, or between $y_{jt}$ and $y_{it}$. Therefore, following Johansen (1988) we write a *k*-dimensional Vector Error Correction model (VEC) as (abstracting from deterministic components):

$$\Delta Y_t = \sum_{i=1}^{p-1} \Gamma_i \Delta Y_{t-i} + \Pi Y_{t-1} + \varepsilon_t , \qquad (12)$$

where $Y_t$ is a vector containing *k* endogenous variables, and $\varepsilon_t \sim iid\ (0, \Sigma)$. In the VEC model (12), $\Pi$ is a $(k \times k)$ matrix of long-run coefficients, which in the presence of cointegration can be factorised according to the number *r* of linearly independent cointegrating vectors as $\Pi = \alpha\beta'$, where $\alpha$ and $\beta$ are both $(k \times r)$ matrices of full rank, with $\beta$ containing the *r* cointegrating vectors and $\alpha$ carrying the corresponding loadings in each of the *r* vectors. In the Johansen procedure a test for the null hypothesis of *r* cointegrating relations is equivalent to a test of the hypothesis that $\Pi$ has less than full rank. For our specific purposes, $Y_t$ is a two-dimensional vector containing two state income series, let us say $y_{it}$ and $y_{jt}$, and our main interest is to determine if these maintain a long-run equilibrium relationship, that is if they are cointegrated. Note that when incorporating the Johansen reduced-rank regression approach within the Pesaran pair-wise procedure, our main focus is on whether there is evidence of cointegration between the variables, rather than on estimating the elements of the cointegrating vector themselves.

[1] Indeed, it should be recalled that in single equation methods one could find evidence of cointegration in one direction (e.g. when regressing $y_{it}$ on $y_{jt}$) but not in the other (i.e. when regressing $y_{jt}$ on $y_{it}$).

Attempting to estimate a single VEC that incorporates the per capita income series for all the US states would be highly problematic. The pair-wise Johansen approach provides the opportunity to incorporate all bivariate state income relationships that exist. Of course, there already exist panel unit root and non-cointegration tests such as Maddala and Wu (1999), Levin et al. (2002), Pedroni (2001) and Im et al. (2003) as potential ways of overcoming the low test power attached to univariate methods. However, following Pesaran et al. (2009), it can be argued that the pair-wise methodology provides two key advantages over existing panel techniques. First, the joint null hypothesis of these panel tests is the existence of a unit root across all series, or non-cointegration across all relationships. This hypothesis can be rejected even when the proportion of the cases for which the unit root or non-cointegration null is rejected is small. The pair-wise approach directly addresses the question of what proportion of cases are stationary or cointegrated. Second, the presence of unobserved common factors complicates the application of these tests where the presence of cross-section dependence can lead to size distortion. The so-called second generation panel unit root tests (following the terminology proposed in the survey by Breitung and Pesaran (2008)) have attempted to allow for possible cross-section dependence through unobserved common factors, but their applications are complicated by the uncertainties surrounding the number of unobserved factors, the nature of the unit root process (whether it is common or country specific), and the fact that longer data spans are required for modelling the cross-section dependence. The pair-wise method is robust to cross-sectional dependence since it operates on the time series dimension.

## 3. Data and empirical analysis

We employ Per Capita Personal Income (PCPI) data for 48 US states in dollars.[2] The data, expressed in natural logarithm form, are annual, cover the study period 1929 to 2009 for a total of 81 observations, and were downloaded from the Federal Reserve Economic Dataset (FRED) assembled by the Federal Reserve Bank of St. Louis. In the dataset, each income series is coded as the state abbreviation plus the suffix PCPI; thus, for instance, ARPCPI is Per Capita Personal Income in Arkansas, and so on. Because reliable data on state price levels are not available, the PCPI series for each state is then deflated by the overall consumer price index; see for example Sala-i-Martin (1996) and Barro and Sala-i-Martin (1999, ch.10). We also analyse per capita personal income data obtained from the FRED dataset for 346 MSAs over the study period 1969-2009. In the original pair-wise approach advocated by Pesaran (2007), which examines differentials between pairs of series, the results are not affected by the choice of data in nominal or real terms (as long as all series are deflated by the same deflator, as in this paper). However, when one considers cointegration between pairs of series, whether or not the series are nominal or real turns out to be important.

To begin with, we consider a more standard approach where unit root tests are applied to state differentials calculated with respect to national real per capita income. This involves performing 48 ADF unit root tests. Table 1 presents the results where we can reject the non-stationarity null at the 5% significance level in only 13 out of 48 cases, i.e. approximately 27.08% (the corresponding percentage of rejections at the 10% significance level is 35.42%). As argued above, this finding says little about bivariate convergence involving individual states. Table 2 reports the percentage of rejections of the ADF tests based on all 1128 bivariate income differentials. These tests are

[2] We exclude Alaska and Hawaii from our analysis on the grounds that these states are not geographically contiguous with any other state in the US, so the mechanisms that may underpin long-run constancy of income ratios across states within the US may not operate in these cases.

conducted at the 5 and 10% significance levels, and the order of augmentation of the test regression is determined using the AIC with $p_{max} = 4$. As can be seen, the percentage of rejections exceeds the size of the unit root test statistics, being equal to 33.78% (46.72%) at the 5% (10%) significance level; qualitatively similar results are obtained when employing the more powerful ERS unit root test (these results are not reported here, though). The results just described, however, only focus on the point estimate of the proportion of the pair-wise tests that reject the null hypothesis of no convergence. It is important to consider the precision of these estimates because potential cross-section dependence between the test outcomes is likely to increase the uncertainty considerably. We therefore employ the factor augmented sieve bootstrap approach outlined in the previous section. In doing so, the cross-sectional dependence is interpreted in terms of a factor model. As explained, the parameters of an underlying factor model are estimated directly, and we subsequently use these estimates to bootstrap the pair-wise rejection rates, treating this factor model as an approximation to the true data generation process (the bootstrap results are based on 5,000 replications).

We start off by considering the time series properties of the cross-sectional mean of all income series in real terms, denoted $\bar{y}_t$ in the previous section, as an estimate of the common factor.[3] The results indicate that the ADF and ERS tests (including constant and trend) provide mixed evidence regarding the order of integration of $\bar{y}_t$; that is, while $ADF(1) = -3.075$ suggests that the null hypothesis of a unit root is not rejected at the 10% significance level, $ERS(1) = -3.263$ provides evidence in favour of stationarity. Thus, for the purposes of the implementation of the bootstrap we consider two cases, one in which a unit root is imposed on $\bar{y}_t$, and another one in which a unit root is not imposed.

[3] An application of the Bai and Ng (2002) test confirmed the presence of a single common factor driving US state incomes.

Table 3 reports the respective distributions of the bootstrapped fraction of rejections for the income gaps in real terms. Focusing on the case where a unit root is imposed on the common factor, the results of the ADF test reveal that the mean of the bootstrap distribution is 16.03% for $\alpha = 10\%$ , a value that is much lower than the corresponding point estimate of 46.72% reported in Table 2. The lower bound of the 90% bootstrap confidence interval is 7.27%, which includes 10%. It should be recalled that for convergence one would expect a fraction of rejections larger than 10%, which is the significance level at which the tests are conducted. If a unit root is not imposed on the common factor, the lower bound of the bootstrap distribution is in the boundary of $\alpha = 10\%$ for the ADF test (i.e. 10.81%). Similar results are observed when using $\alpha = 5\%$ as significance level, or when instead of the ADF unit root test the ERS test is employed in the analysis. It is therefore clear that cross-section dependence introduces a large degree of uncertainty into the point estimate of the proportion of rejections.

Our findings so far do not provide strong support for the view of long-run convergence between US state incomes in real terms. The possibility we have considered so far is one of strong convergence where the implied cointegrating vector is restricted to [1,-1]. In the spirit of Bernard and Durlauf (1996), it is conceivable that a weaker form of convergence is more relevant whereby state income pairs are cointegrated, but with an unknown cointegrating vector not necessarily equal to [1, -1].[4] In order to explore this possibility, we employ and develop the Johansen (1988) maximum likelihood estimator of cointegrated VAR models within the Pesaran pair-wise setting. The starting point of the analysis is to estimate for each possible state income pairs a VAR model with an unrestricted constant term (since the variables exhibit a positive drift). The optimal lag length of the VAR models is determined using the AIC with $p_{max} = 4$ . Then, we use the trace and the maximum

[4] See also the early work by Quintos (1995) and others in the context of the relationship between government revenue and expenditure and the sustainability of the budget deficit,

eigenvalue tests to determine the number of cointegrating vectors, which is denoted $r$. The former involves testing the null hypothesis that there are $r = 0$ cointegrating vectors against the alternative that $r \geq 1$. The latter involves testing the null hypothesis that there are $r = 0$ cointegrating vectors against the alternative that $r = 1$. In both cases, if the null hypothesis is rejected, then this would provide support for the view that the two real income series share the same stochastic trend.[5]

The results in Table 4 indicate that the number of rejections of $Ho: r = 0$ (against $Ha: r \geq 1$) is 809 out of 1128 possible real income gaps, i.e. 71.72% when setting $\alpha = 10\%$ while the corresponding number of rejections for $\alpha = 5\%$ is 707, i.e. 62.68%. Once again, these initial point estimates of the percentage of rejections of non-cointegration fail to account for the presence of potential cross section dependence so we implement the bootstrap procedure. These results offer support for the presence of a weaker form of convergence; see Table 5. Indeed, let us again consider the results obtained for the trace test when a unit root is imposed on $\bar{y}_t$. When looking at $\alpha = 10\%$, the mean proportion of rejections is 37.31%, and the 90% bootstrap confidence interval around this mean estimate ranges from 12.94% to 69.59%. Therefore, this 90% confidence interval does not cover values below 10%. Qualitatively similar results are obtained when using the maximum eigenvalue test, or when setting $\alpha = 5\%$, or when a unit root is imposed on the common factor $\bar{y}_t$ (irrespective of the significance level).

These results are more favourable towards the presence of cointegration between bivariate state pairs. It can be argued that strong or weak convergence is reflected in the long-run slope coefficient that depicts each long-run relationship. Conditioning on the cases for which the trace test provides evidence in favour of cointegration, that is 809 when $\alpha = 10\%$ and 707 when $\alpha = 5\%$, the null hypothesis that the cointegrating vector can be set equal to [1, -1] cannot be rejected in less than

[5] It is worth mentioning that a test of the null hypothesis that $r \leq 1$ against the alternative that $r = 2$ is not undertaken as it opens up the possibility of obtaining counterintuitive results. Indeed, this second test rejection of the null would indicate the presence of two cointegrating vectors or, in other words, that each real income series is stationary in levels.

half of the possible cases, or more precisely $303/809 = 0.37$ when $\alpha = 10\%$ and $310/707 = 0.44$ when $\alpha = 5\%$. Given the presence of weak as opposed to strong convergence across state pairs and heterogeneity in the estimated long-run slopes, it is of interest to consider what factors might drive the estimated values of the slopes themselves and whether it is possible to define a basis for (weak) convergence clusters. Denoting $\beta_2^{(ij)}$ as the cointegrating slope, we measure the strength of convergence between real per capita personal incomes in states *i* and *j* as the absolute value of the difference between $\beta_2^{(ij)}$ and one, that is $\left|\beta_2^{(ij)} - 1\right|$, and consider the roles played by two potential drivers. The first is the absolute value of the difference between (the logs of) initial per capita income in states *i* and *j*, denoted by $\left|\log y_{i0} - \log y_{j0}\right|$. If $\beta$-convergence predicts that poorer states should grow faster than richer states, then $\left|\log y_{i0} - \log y_{j0}\right|$ should be characterised by a negative and significant coefficient. The second driver is (the log of) distance between states *i* and *j*, which we denote $\log D_{ij}$. For this, we employ the Euclidian distance between the population centres of any two states, based on the geographic coordinates (latitude and longitude) obtained from the Census Bureau for the year 2000.[6]

Estimation by OLS for the 809 cases where the Trace test rejected the null hypothesis of no cointegration at the 10% significance level provides the following results:

$$\left|\beta_2^{(ij)} - 1\right| = -\underset{(0.076)}{0.032} + \underset{(0.028)}{0.110}\left|\log y_{i0} - \log y_{j0}\right| + \underset{(0.011)}{0.024}\log D_{ij}$$

$$\hat{\sigma} = 0.251$$

where White's heteroscedasticity-consistent standard errors are reported in parentheses.[7] The

[6] We are most grateful to Gary Wagner who kindly provided these data, which were used in Garrett, Wagner and Wheelock (2007).

[7] It should also be noted that when the null of homogeneity is not rejected at the 10% significance level (i.e. 303 instances), the corresponding value of the dependent variable is set equal to zero.

estimated positive coefficient on $\left|\log y_{i0} - \log y_{j0}\right|$ is not consistent with $\beta$-convergence insofar as poorer states do not necessarily catch-up with richer states. Instead, we find that the likelihood of strong convergence or $\left|\beta_2^{(ij)} - 1\right| = 0$ is enhanced if two pair-wise states are characterised by a similar initial per capita incomes. The estimated coefficient on (the log of) distance is positive and statistically significant, supporting the view that convergence between any two states is strongest, the closest they are in terms of distance. Thus, although our findings are supportive of cointegrating relationships across state pairs, it is on this basis that convergence clubs or groupings may arise. In this respect, our findings are consistent with Choi (2004) who finds that convergence is strongest among geographically neighbouring states that share certain common regional features such as climate and industrial structures.

In the final part of our investigation, we consider convergence using MSA-level data which provides 346 annual income series over the study period 1969-2009. As with the case of State-level data, we are using the maximum study period dictated by data availability. Once again, we begin our analysis by applying the ADF unit root test to MSA income differentials calculated with respect to national real per capita income. Results, not reported here, indicate that the null hypothesis of non-stationarity can be rejected only in 26 instances (49) when using a 5% (10%) significance level. Table 6 reports the percentage of rejections of the ADF unit root tests based on all 59585 bivariate MSA-level income differentials.[8] The percentage of rejections exceeds the size of the unit root test statistics, being equal to 15.48% (25.38%) at the 5% (10%) significance level. Table 7 reports the respective distributions of the bootstrapped fraction of rejections. These results are supportive of strong convergence when MSA-level data are analysed. Focusing on the case where a unit root is

[8] The significant increase in the number of income differentials that accompanies an analysis at MSA-level poses no additional difficulties for the pair-wise convergence approach, other than the requirement for substantially more computing time.

imposed on the common factor for example, the lower bound of the 90% bootstrap confidence interval is 16.26%. In contrast to the earlier results based on state-level data, this is greater than 10%. This finding is qualitatively unchanged when using $\alpha = 5\%$ as the significance level, or when a unit root is not imposed on the common factor. Further results based on the application of the pair-wise Johansen test are reported in Tables 8 and 9. Focussing on the Trace test, the fraction of rejections of $Ho: r = 0$ (against $Ha: r \geq 1$) is 24.64% when setting $\alpha = 10\%$, while the corresponding fraction of rejections for $\alpha = 5\%$ is 15.32%. If a unit root is imposed on $\bar{y}_t$, when looking at $\alpha = 10\%$ the mean proportion of rejections is 27.40%, and the 90% bootstrap confidence interval around this mean estimate ranges from 17.98% to 42.66%.  This 90% confidence interval does not cover values below 10%, and so offers further support for income convergence. Qualitatively similar results are obtained when using the maximum eigenvalue test, or when setting $\alpha = 5\%$, or when a unit root is imposed on the common factor $\bar{y}_t$ (irrespective of the significance level).

## 4. Concluding remarks

This paper uses time series annual information for US states to assess one of the key predictions of the neoclassical growth model, namely that of real per-capita income convergence. Our empirical modelling exercise uses a pair-wise probabilistic approach to examine stochastic convergence. This approach is based on the fraction of rejections of non-stationarity or non-cointegration across all bivariate state per capita income pairs. According to our results, we confirm convergence over a long time period as well as convergence with highly disaggregated data. While we reject strong convergence at state-level insofar as testing the non-stationarity of pair-wise state income differentials, these tests are characterised by implied cointegrating vectors of the form [1,-1] under the alternative hypothesis. Further results based on the development and application of a pair-wise

Johansen cointegration test offer more empirical support. In this respect, there is a weak form of convergence characterised by cointegration between state incomes where the elements of the cointegrating vector are unrestricted. However, we find that convergence between any two states is strongest for those states that have similar per capita incomes and are closest in terms of distance. Additional analysis at a more disaggregated level using metropolitan statistical area data provides stronger evidence of long-run convergence characterised by stationary income differentials.

Table 1. ADF unit root tests on state income relative to national real per capita income

| State | Obs | Lags | t-Stat | *p*-value |
|---|---|---|---|---|
| AL | 76 | 4 | -2.432 | [0.136] |
| AR | 77 | 3 | -3.364 | [0.015] |
| AZ | 77 | 3 | -3.365 | [0.015] |
| CA | 79 | 1 | -1.376 | [0.590] |
| CO | 79 | 1 | -2.101 | [0.245] |
| CT | 79 | 1 | -2.081 | [0.253] |
| DE | 77 | 3 | -1.256 | [0.646] |
| FL | 77 | 3 | -2.193 | [0.211] |
| GA | 78 | 2 | -3.624 | [0.007] |
| IA | 79 | 1 | -2.341 | [0.162] |
| ID | 77 | 3 | -3.368 | [0.015] |
| IL | 80 | 0 | -2.606 | [0.096] |
| IN | 78 | 2 | -2.473 | [0.126] |
| KS | 76 | 4 | -3.262 | [0.020] |
| KY | 80 | 0 | -1.628 | [0.464] |
| LA | 77 | 3 | -1.717 | [0.419] |
| MA | 79 | 1 | -2.468 | [0.127] |
| MD | 79 | 1 | -2.463 | [0.128] |
| ME | 80 | 0 | -2.028 | [0.274] |
| MI | 80 | 0 | -1.258 | [0.646] |
| MN | 76 | 4 | -1.942 | [0.312] |
| MO | 80 | 0 | -2.621 | [0.093] |
| MS | 76 | 4 | -2.665 | [0.085] |
| MT | 77 | 3 | -1.842 | [0.358] |
| NC | 78 | 2 | -3.644 | [0.007] |
| ND | 79 | 1 | -2.057 | [0.262] |
| NE | 78 | 2 | -1.891 | [0.335] |
| NH | 79 | 1 | -1.539 | [0.509] |
| NJ | 79 | 1 | -2.328 | [0.166] |
| NM | 80 | 0 | -2.186 | [0.213] |
| NV | 76 | 4 | -0.858 | [0.796] |
| NY | 77 | 3 | -4.083 | [0.002] |
| OH | 80 | 0 | -0.982 | [0.756] |
| OK | 80 | 0 | -1.570 | [0.493] |
| OR | 80 | 0 | -1.515 | [0.521] |
| PA | 78 | 2 | -3.299 | [0.018] |
| RI | 78 | 2 | -2.376 | [0.152] |
| SC | 80 | 0 | -3.597 | [0.008] |
| SD | 78 | 2 | -2.250 | [0.191] |
| TN | 77 | 3 | -3.805 | [0.004] |
| TX | 77 | 3 | -3.265 | [0.020] |
| UT | 77 | 3 | -2.466 | [0.128] |
| VA | 78 | 2 | -1.497 | [0.530] |
| VT | 80 | 0 | -1.033 | [0.738] |
| WA | 79 | 1 | -3.068 | [0.033] |
| WI | 76 | 4 | -2.809 | [0.062] |
| WV | 78 | 2 | -2.398 | [0.146] |
| WY | 78 | 2 | -3.204 | [0.023] |

Note: The number of lags is selected using the AIC with $p_{max} = 4$.

Table 2. Fraction of rejections assuming state income pairs are cointegrated with known cointegrating vector $[1, -1]'$. Sample period 1929 – 2009.

| $\alpha$ | Fraction of rejections |
|---|---|
| 5% | 33.78 |
| 10% | 46.72 |

Notes: The ADF unit-root test regressions include a linear trend if it is statistically significant at the 5% level. The number of lags of the dependent variable is determined using the AIC with $p_{max} = 4$. Critical values for the ADF test are based on response surfaces estimated by Cheung and Lai (1995).

Table 3. Distribution of the bootstrapped fraction of rejections assuming state income pairs are cointegrated with known cointegrating vector $[1,-1]'$. Sample period 1929 – 2009.

Imposing a unit root on common factor

| $\alpha$ | Mean | Median | SD | 2.5% | 5% | 10% | 90% | 95% | 97.5% |
|---|---|---|---|---|---|---|---|---|---|
| 5% | 8.79 | 7.80 | 4.57 | <u>2.75</u> | 3.28 | 3.98 | 14.98 | 17.73 | <u>20.57</u> |
| 10% | 16.03 | 14.89 | 6.63 | 6.38 | <u>7.27</u> | 8.59 | 25.09 | <u>28.82</u> | 31.74 |

Not imposing a unit root on common factor

| $\alpha$ | Mean | Median | SD | 2.5% | 5% | 10% | 90% | 95% | 97.5% |
|---|---|---|---|---|---|---|---|---|---|
| 5% | 11.46 | 10.73 | 4.78 | <u>4.43</u> | 5.14 | 6.12 | 17.73 | 20.39 | <u>23.05</u> |
| 10% | 20.10 | 19.33 | 6.41 | 9.66 | <u>10.81</u> | 12.50 | 28.81 | <u>31.65</u> | 34.93 |

Notes: Pair-wise ADF unit root tests. The bounds of the confidence intervals are given by the underlined figures. The number of bootstrap replications used to derive the empirical distribution of the fraction of rejections is 5000.

Table 4. Fraction of rejections assuming state income pairs are cointegrated with unknown cointegrating vector. Sample period 1929 – 2009.

Johansen trace test

| Ho | Ha | $\alpha$ | Fraction of rejections |
|---|---|---|---|
| $r = 0$ | $r \geq 1$ | 5% | 62.68 |
| $r = 0$ | $r \geq 1$ | 10% | 71.72 |

Johansen maximum eigenvalue test

| Ho | Ha | $\alpha$ | Fraction of rejections |
|---|---|---|---|
| $r = 0$ | $r = 1$ | 5% | 53.55 |
| $r = 0$ | $r = 1$ | 10% | 65.60 |

Notes: The Johansen cointegration test results are based on the estimation of bivariate VAR models with a constant term that enters unrestrictedly. The number of lags of the VAR models is determined using the AIC with $p_{max} = 4$. $r$ denotes the number of cointegrating vectors. Critical values are based on response surfaces estimated by MacKinnon, Haug and Michelis (1999).

Table 5. Distribution of the bootstrapped fraction of rejections assuming state income pairs are cointegrated with unknown cointegrating vector. Sample period 1929 – 2009.

Imposing a unit root on common factor

| Johansen test | | | Mean | Median | SD | 2.5% | 5% | 10% | 90% | 95% | 97.5% |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Trace: | | | | | | | | | | | |
| Ho | Ha | $\alpha$ | | | | | | | | | |
| $r = 0$ | $r \geq 1$ | 5% | 25.87 | 22.52 | 14.80 | <u>6.20</u> | 7.71 | 9.93 | 47.08 | 55.33 | <u>62.41</u> |
| $r = 0$ | $r \geq 1$ | 10% | 37.31 | 34.62 | 17.31 | 10.90 | <u>12.94</u> | 16.58 | 62.59 | <u>69.59</u> | 74.91 |
| Max. Eigenvalue: | | | | | | | | | | | |
| Ho | Ha | $\alpha$ | | | | | | | | | |
| $r = 0$ | $r = 1$ | 5% | 21.54 | 17.64 | 13.66 | <u>5.41</u> | 6.56 | 8.07 | 40.78 | 50.36 | <u>56.83</u> |
| $r = 0$ | $r = 1$ | 10% | 31.65 | 28.10 | 16.22 | 9.66 | <u>11.44</u> | 13.74 | 55.50 | <u>64.27</u> | 69.42 |

Not imposing a unit root on common factor

| Johansen test | | | Mean | Median | SD | 2.5% | 5% | 10% | 90% | 95% | 97.5% |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Trace: | | | | | | | | | | | |
| Ho | Ha | $\alpha$ | | | | | | | | | |
| $r = 0$ | $r \geq 1$ | 5% | 27.80 | 25.53 | 12.77 | <u>9.22</u> | 11.26 | 13.56 | 45.48 | 52.75 | <u>58.95</u> |
| $r = 0$ | $r \geq 1$ | 10% | 36.10 | 34.31 | 13.57 | 14.71 | <u>17.29</u> | 20.21 | 55.14 | <u>61.53</u> | 67.11 |
| Max. Eigenvalue: | | | | | | | | | | | |
| Ho | Ha | $\alpha$ | | | | | | | | | |
| $r = 0$ | $r = 1$ | 5% | 20.90 | 18.44 | 11.22 | <u>6.56</u> | 7.80 | 9.49 | 35.90 | 43.71 | <u>50.44</u> |
| $r = 0$ | $r = 1$ | 10% | 27.92 | 25.44 | 12.22 | 11.26 | <u>12.86</u> | 14.89 | 44.60 | <u>52.58</u> | 59.22 |

Notes: $r$ denotes the number of cointegrating vectors. The bounds of the confidence intervals are given by the underlined figures. The number of bootstrap replications used to derive the empirical distribution of the fraction of rejections is 5000.

Table 6. Fraction of rejections assuming MSA income pairs are cointegrated with known cointegrating vector $[1,-1]'$

| $\alpha$ | Fraction of rejections |
|---|---|
| 5% | 15.48 |
| 10% | 25.38 |

Notes: The ADF unit-root test regressions include a linear trend if it is statistically significant at the 5% level. The number of lags of the dependent variable is determined using the AIC with $p_{max} = 4$. Critical values for the ADF test are based on response surfaces estimated by Cheung and Lai (1995).

Table 7. Distribution of the bootstrapped fraction of rejections assuming MSA income pairs are cointegrated with known cointegrating vector $[1,-1]'$

Imposing a unit root on common factor

| $\alpha$ | Mean | Median | SD | 2.5% | 5% | 10% | 90% | 95% | 97.5% |
|---|---|---|---|---|---|---|---|---|---|
| 5% | 12.65 | 12.46 | 2.29 | 8.65 | 9.19 | 9.88 | 15.64 | 16.72 | 17.84 |
| 10% | 21.16 | 20.99 | 3.12 | 15.39 | 16.26 | 17.32 | 25.21 | 26.59 | 27.93 |

Not imposing a unit root on common factor

| $\alpha$ | Mean | Median | SD | 2.5% | 5% | 10% | 90% | 95% | 97.5% |
|---|---|---|---|---|---|---|---|---|---|
| 5% | 12.68 | 12.49 | 2.25 | 8.80 | 9.30 | 9.98 | 15.62 | 16.68 | 17.74 |
| 10% | 21.20 | 21.01 | 3.07 | 15.71 | 16.44 | 17.42 | 25.24 | 26.57 | 27.78 |

Notes: Pair-wise ADF unit root tests. The bounds of the confidence intervals are given by the underlined figures. The number of bootstrap replications used to derive the empirical distribution of the fraction of rejections is 5000.

Table 8. Fraction of rejections assuming MSA income pairs are cointegrated with unknown cointegrating vector

Johansen trace test

| Ho | Ha | $\alpha$ | Fraction of rejections |
|---|---|---|---|
| $r = 0$ | $r \geq 1$ | 5% | 15.32 |
| $r = 0$ | $r \geq 1$ | 10% | 24.64 |

Johansen maximum eigenvalue test

| Ho | Ha | $\alpha$ | Fraction of rejections |
|---|---|---|---|
| $r = 0$ | $r = 1$ | 5% | 12.57 |
| $r = 0$ | $r = 1$ | 10% | 20.39 |

Notes: The Johansen cointegration test results are based on the estimation of bivariate VAR models with a constant term that enters unrestrictedly. The number of lags of the VAR models is determined using the AIC with $p_{max} = 4$. $r$ denotes the number of cointegrating vectors. Critical values are based on response surfaces estimated by MacKinnon, Haug and Michelis (1999).

Table 9. Distribution of the bootstrapped fraction of rejections assuming MSA income pairs are cointegrated with unknown cointegrating vector

Imposing a unit root on common factor

| Johansen cointegration test | | | Mean | Median | SD | 2.5% | 5% | 10% | 90% | 95% | 97.5% |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Trace: | | | | | | | | | | | |
| Ho | Ha | $\alpha$ | | | | | | | | | |
| $r = 0$ | $r \geq 1$ | 5% | 19.57 | 18.23 | 6.32 | 11.37 | 12.33 | 13.33 | 27.38 | 31.34 | 36.22 |
| $r = 0$ | $r \geq 1$ | 10% | 27.40 | 25.74 | 7.95 | 16.73 | 17.98 | 19.36 | 37.45 | 42.66 | 48.17 |
| Max. Eigenvalue: | | | | | | | | | | | |
| Ho | Ha | $\alpha$ | | | | | | | | | |
| $r = 0$ | $r = 1$ | 5% | 17.72 | 16.79 | 4.65 | 11.45 | 12.15 | 13.01 | 23.12 | 26.31 | 29.34 |
| $r = 0$ | $r = 1$ | 10% | 25.15 | 24.04 | 6.01 | 16.96 | 17.91 | 18.95 | 32.40 | 36.21 | 40.26 |

Not imposing a unit root on common factor

| Johansen cointegration test | | | Mean | Median | SD | 2.5% | 5% | 10% | 90% | 95% | 97.5% |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Trace: | | | | | | | | | | | |
| Ho | Ha | $\alpha$ | | | | | | | | | |
| $r = 0$ | $r \geq 1$ | 5% | 20.14 | 19.68 | 3.85 | 13.97 | 14.71 | 15.61 | 25.23 | 27.24 | 28.91 |
| $r = 0$ | $r \geq 1$ | 10% | 28.28 | 27.78 | 4.83 | 20.25 | 21.29 | 22.55 | 34.58 | 37.14 | 38.99 |
| Max. Eigenvalue: | | | | | | | | | | | |
| Ho | Ha | $\alpha$ | | | | | | | | | |
| $r = 0$ | $r = 1$ | 5% | 18.46 | 18.09 | 3.36 | 13.06 | 13.70 | 14.51 | 22.89 | 24.66 | 26.16 |
| $r = 0$ | $r = 1$ | 10% | 25.96 | 25.59 | 4.23 | 19.00 | 19.88 | 20.97 | 31.56 | 33.57 | 35.39 |

Notes: *r* denotes the number of cointegrating vectors. The bounds of the confidence intervals are given by the underlined figures. The number of bootstrap replications used to derive the empirical distribution of the fraction of rejections is 5000.

## References


Bai, J., Ng, S., 2002. Determining the number of factors in approximate factor models. Econometrica 70, 191-221.

Barro, R. J., 1991. Economic growth in a cross section of countries. Quarterly Journal of Economics 106, 407-43.

Barro, R. J., Sala-I-Martin, X., 1991. Convergence across states and regions. Brookings Papers on Economic Activity Issue 1, 107-82.

Barro, R. J., Sala-I-Martin, X., 1992. Convergence. Journal of Political Economy 100, 223-51.

Barro, R., Sala-i-Martin, X., 1999. Economic Growth. MIT Press, Cambridge MA.

Baumol, W. J., (1986), Productivity, growth, convergence and welfare: What the long-run data show, American Economic Review 76, 1072-85.

Bernard, A. B., Durlauf., S.N., 1995.Convergence in international output. Journal of Applied Econometrics 10, 97-108.

Breitung, J., Pesaran, M.H., 2008. Unit roots and cointegration in panels. In: Mátýas, L., Sevestre, P. (Eds.), The Econometrics of Panel Data, Springer-Verlag, Berlin Heidelberg, 279-322.

Carlino, G.A., Mills, L.O., 1993. Are U.S. regional incomes converging?: A time series analysis. Journal of Monetary Economics 32, 335-346.

Cheung, Y.-W, Lai, K.S., 1995. Lag order and critical values of the augmented Dickey-Fuller test. Journal of Business and Economic Statistics 13, 277-280.

Choi, C.Y., 2004. A reexamination of output convergence in the US states: toward which level(s) are they converging? Journal of Regional Science 44, 713-741.

Dickey D.A., Fuller W.A., 1979. Distribution of the estimators for autoregressive time series with a unit root. Journal of the American Statistical Association 74, 427-431.

Elliott, G., Rothenberg, T.J., Stock, J.H., 1996. Efficient tests for an autoregressive unit root. Econometrica 64, 813-836.

Engle, R.F., Granger, C.W.J., 1987. Cointegration and error correction: representation, estimation and testing. Econometrica 55, 251-276.

Evans, P., Karras, G., 1996. Convergence revisited. Journal of Monetary Economics 37, 249-265.

Garrett, T.A., Wagner, G.A., Wheelock, D.C., 2007. Regional disparities in the spatial correlation of state income growth, 1977–2002. Annals of Regional Science 41, 601-618.

Hansen, B.E., Phillips, P.C.B., 1990. Estimation and inference in models of cointegration: A simulation study. In: Fomby, T.B., Rhodes, G.F. (Eds.), Advances in Econometrics: Cointegration, Unit Roots and Spurious Regression, JAI Press. Greenwich, CO.

Im K., Pesaran, M.H., Shin, Y., 2003. Testing for unit roots in heterogeneous panels. Journal of Econometrics 115, 53-74.

Johansen, S., 1988. Statistical analysis of cointegration vectors, Journal of Economic Dynamics and Control 12, 231-254.

Johansen, S., 1995. Likelihood-based Inference in Cointegrated Vector Autoregressive Models. Oxford University Press, Oxford.

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

Le Pen, Y., 2011. A pair-wise approach to output convergence between European regions. Economic Modelling 28, 955-64.

MacKinnon, J.G., Haug, A., and Michelis, L., 1999. Numerical Distribution Functions of Likelihood Ratio Tests for Cointegration. Journal of Applied Econometrics 14, 563-577.

Maddala, G.S., Wu, S., 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.

Mankiw, N.G., Romer, D., Weil, D.N, 1992. A contribution to the empirics of economic growth. Quarterly Journal of Economics 107, 407-37.

Mello M., 2011. Stochastic convergence across US states. Macroeconomics Dynamics 15,160-183.

Pedroni, P., 2001. Purchasing power parity in cointegrated panels, Review of Economics and Statistics 83, 727-731.

Pesaran, M.H., 2007. A pair-wise approach to testing for output and growth convergence. Journal of Econometrics 138, 312-355.

Pesaran, M.H., Smith, R.P., Yamagata, T., Hvozdyk, L., 2009. Pairwise tests of purchasing power parity. Econometric Reviews 28, 495-521.

Quah, D., 1993. Galton's fallacy and tests of the convergence hypothesis. Scandinavian Journal of Economics 95, 427-43.

Quah, D., 1997. Empirics for growth and distribution: polarization, stratification and convergence club. Journal of Economic Growth 2, 27–59.

Quintos, C., 1995. Sustainability of the deficit process with structural shifts. Journal of Business and Economic Statistics 13, 409-417.

Saikkonen, P., 1992. Estimation and testing of cointegrated systems by an autoregressive approximation. Econometric Theory 8, 1-27.

Sala-i-Martin, X., 1996. The classical approach to convergence analysis. Economic Journal 106, 1019-1036.

Stock, J.H., Watson, M.W., 1993. A simple estimator of cointegrating vectors in higher order integrated systems. Econometrica 61, 783-820.

Tsionas E.G., 2001. Regional convergence and common stochastic long-run trends: A re-examination of the US regional data. Regional Studies 35, 689-696.

Appendix 1. Factor estimate equations

| State | Intercept | (s.e.) | Trend | (s.e.) | $\bar{y}_t$ | (s.e.) | $\bar{R}^2$ |
|---|---|---|---|---|---|---|---|
| AL | -1.840 | 0.100 | -0.002 | 0.001 | 1.369 | 0.029 | 0.998 |
| AR | -2.005 | 0.121 | -0.003 | 0.001 | 1.396 | 0.035 | 0.997 |
| AZ | -0.095 | 0.083 | -0.003 | 0.001 | 1.039 | 0.024 | 0.997 |
| CA | 1.317 | 0.028 | | | 0.766 | 0.006 | 0.995 |
| CO | 0.480 | 0.085 | 0.003 | 0.001 | 0.884 | 0.024 | 0.997 |
| CT | 2.472 | 0.174 | 0.011 | 0.001 | 0.423 | 0.050 | 0.985 |
| DE | 2.028 | 0.136 | 0.003 | 0.001 | 0.574 | 0.039 | 0.986 |
| FL | -0.221 | 0.034 | | | 1.051 | 0.008 | 0.996 |
| GA | -1.276 | 0.038 | | | 1.248 | 0.009 | 0.996 |
| IA | -0.903 | 0.147 | -0.006 | 0.001 | 1.263 | 0.042 | 0.992 |
| ID | -0.731 | 0.141 | -0.006 | 0.001 | 1.204 | 0.041 | 0.992 |
| IL | 0.965 | 0.025 | | | 0.833 | 0.006 | 0.996 |
| IN | -0.546 | 0.083 | -0.007 | 0.001 | 1.188 | 0.024 | 0.997 |
| KS | -1.389 | 0.107 | -0.008 | 0.001 | 1.386 | 0.031 | 0.996 |
| KY | -0.969 | 0.037 | | | 1.165 | 0.008 | 0.996 |
| LA | -0.681 | 0.037 | | | 1.110 | 0.008 | 0.996 |
| MA | 2.875 | 0.156 | 0.015 | 0.001 | 0.264 | 0.045 | 0.988 |
| MD | 1.686 | 0.125 | 0.009 | 0.001 | 0.588 | 0.036 | 0.993 |
| ME | 1.334 | 0.144 | 0.007 | 0.001 | 0.618 | 0.042 | 0.991 |
| MI | 0.003 | 0.110 | -0.006 | 0.001 | 1.087 | 0.032 | 0.994 |
| MN | 0.514 | 0.079 | 0.005 | 0.001 | 0.855 | 0.023 | 0.998 |
| MO | 0.259 | 0.025 | | | 0.945 | 0.006 | 0.997 |
| MS | -2.297 | 0.123 | -0.002 | 0.001 | 1.435 | 0.035 | 0.997 |
| MT | -0.361 | 0.166 | -0.007 | 0.001 | 1.141 | 0.048 | 0.986 |
| NC | -1.293 | 0.038 | | | 1.247 | 0.009 | 0.996 |
| ND | -3.389 | 0.293 | -0.017 | 0.002 | 1.886 | 0.084 | 0.979 |
| NE | -0.898 | 0.148 | -0.005 | 0.001 | 1.250 | 0.043 | 0.992 |
| NH | 1.924 | 0.129 | 0.013 | 0.001 | 0.469 | 0.037 | 0.993 |
| NJ | 1.958 | 0.131 | 0.008 | 0.001 | 0.548 | 0.038 | 0.992 |
| NM | -1.357 | 0.153 | -0.006 | 0.001 | 1.319 | 0.044 | 0.992 |
| NV | 1.029 | 0.135 | -0.003 | 0.001 | 0.854 | 0.039 | 0.988 |
| NY | 2.549 | 0.156 | 0.009 | 0.001 | 0.413 | 0.045 | 0.984 |
| OH | 0.488 | 0.065 | -0.003 | 0.001 | 0.936 | 0.019 | 0.998 |
| OK | -1.459 | 0.154 | -0.006 | 0.001 | 1.351 | 0.044 | 0.993 |
| OR | -0.041 | 0.106 | -0.005 | 0.001 | 1.076 | 0.031 | 0.994 |
| PA | 1.278 | 0.079 | 0.004 | 0.001 | 0.697 | 0.023 | 0.997 |
| RI | 2.670 | 0.152 | 0.011 | 0.001 | 0.326 | 0.044 | 0.986 |
| SC | -1.546 | 0.043 | | | 1.281 | 0.010 | 0.996 |
| SD | -2.618 | 0.242 | -0.012 | 0.002 | 1.665 | 0.070 | 0.985 |
| TN | -1.147 | 0.026 | | | 1.213 | 0.006 | 0.998 |
| TX | -0.855 | 0.106 | -0.003 | 0.001 | 1.211 | 0.031 | 0.996 |
| UT | -0.414 | 0.102 | -0.005 | 0.001 | 1.121 | 0.030 | 0.995 |
| VA | 0.080 | 0.100 | 0.006 | 0.001 | 0.922 | 0.029 | 0.997 |
| VT | 1.190 | 0.130 | 0.008 | 0.001 | 0.645 | 0.038 | 0.993 |
| WA | 0.233 | 0.117 | -0.003 | 0.001 | 1.009 | 0.034 | 0.994 |
| WI | 0.316 | 0.019 | | | 0.941 | 0.004 | 0.998 |
| WV | 0.067 | 0.103 | 0.002 | 0.001 | 0.920 | 0.030 | 0.996 |
| WY | 0.422 | 0.063 | | | 0.929 | 0.014 | 0.982 |