



ISSN 1791-3144

**University of Macedonia
Department of Economics**

Discussion Paper Series

**A note on the extent of US regional income
convergence**

Mark J Holmes, Jesus Otero and Theodore Panagiotidis

Discussion Paper No. 3/2013

Department of Economics, University of Macedonia, 156 Egnatia str, 540 06
Thessaloniki, Greece, Fax: + 30 (0) 2310 891292
<http://www.uom.gr/index.php?newlang=eng&tmima=3&categorymenu=2>

A NOTE ON THE EXTENT OF US REGIONAL INCOME CONVERGENCE[♦]

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

Abstract

Long-run income convergence is investigated in the US context. We employ a novel pair-wise econometric procedure based on a probabilistic definition of convergence. 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. To address cross-sectional dependence, we employ a bootstrap methodology to derive the empirical distribution of the fraction of rejections. We find supporting evidence of US states sharing a common stochastic trend consistent with a definition of convergence based on long-run forecasts of state incomes being proportional rather than equal. We find that the strength of convergence between states decreases with distance and initial income disparity. Using Metropolitan Statistical Areas data, 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 thank participants at the 2011 meeting of the New Zealand Econometric Study Group with particular reference to Andrew Coleman and Peter Phillips, seminar participants at the University of Canterbury as well as participants at the 71st International Atlantic Economic Conference (Athens, Greece), the 18th International Conference on Panel Data (Paris, France), and the 66th European Meeting of the Econometric Society (Málaga, Spain). The Editor, an Associate Editor, and an anonymous referee also provided valuable comments and suggestions that helped improve our paper. The usual disclaimer applies.

1. Introduction

In recent years, economists have keenly debated the neoclassical growth model prediction of per-capita income convergence. Allowing for the same structural characteristics, countries or regions with relatively less capital would be expected to accumulate more capital and grow faster than those with relatively more capital, eventually converging to the same steady state. In contrast, models of endogenous growth or the “new” theories of international trade argue that such convergence might fail on account of increasing returns and externalities. Indeed, as pointed out by Krugman (1991) and Krugman and Venables (1995), in the presence of increasing returns economic activity might be expected to concentrate geographically in a few areas. A wide range of studies that includes early work by Barro (1991), Barro and Sala-i-Martin (1991, 1992), Baumol (1986), Bernard and Durlauf (1995), Carlino and Mills (1993), Mankiw et al. (1992) and Sala-i-Martin (1996) and more recently, studies such as Papageorgiou and Perez-Sebastian (2004), Inklaar and Timmer (2009) and Deckers and Hanck (2013) have considered convergence across countries, US states, European regions and industries, and provided mixed evidence in favour of convergence. The empirical investigation of the convergence hypothesis has 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. Moreover, temporary shocks to key structural variables such as saving rates, population growth, and technological progress are characterised by stationary relative outputs thereby indicating that economies are

stochastically converging. This definition is exemplified by studies such as Bernard and Durlauf (1995), who 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 among US states. As pointed out by Choi (2004) among 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) provide evidence 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 test 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 development of an econometric procedure advocated by Pesaran

(2007). Within this framework, a pair-wise probabilistic definition of regional convergence forms the basis of our empirical testing strategy. The idea here is that for a sample of N states, unit root tests are conducted on all $N(N - 1)/2$ real per-capita income differentials (pairs or gaps). Under the null of non-stationarity or non-convergence, one would expect the fraction of real per-capita income pairs for which the unit-root hypothesis is rejected to be close to the size of the underlying unit-root tests, denoted as α . Thus, we can argue that the null of non-stationarity for all state pairs could be rejected if this fraction of rejections exceeds α . However, the presence of cross-sectional dependence can make inference based on the fraction of rejections difficult, so the bootstrap methodology is employed to derive the empirical distribution of the fraction of rejections.

There already exist a limited number of studies that investigate stochastic convergence using a pair-wise approach. Pesaran (2007) considers data for 101 countries and geographical sub-groups. 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). Deckers and Hanck (2013) do not support the notion of convergence using a sample of 51 countries. In terms of studies of regional income convergence, Mello (2011) considers the case of the 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 α 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. Using these studies as a starting point, we extend the analysis in two important directions. First, none of these regional studies consider an empirical distribution of the fraction of rejections. In our investigation, we construct confidence intervals around the point estimates of the percentage of rejections. Second, and more importantly, the above mentioned studies only consider pair-wise stationarity with an implied cointegrating coefficient of unity. In sharp contrast to the existing literature, we conduct an analysis

that also permits for the possibility of pair-wise cointegration, that is relaxing the [1, -1] assumption for the cointegrating vector, as distinct to pair-wise stationarity. This extension allows for pair-wise states to share a common stochastic trend such that state responses to shocks are proportional rather than equal.

The paper is organised as follows. The following section briefly describes the pair-wise testing approach to convergence. 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 more supportive evidence. Further, 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), confirmatory evidence of convergence is found (more disaggregation implies stronger convergence). 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 literature to assess stochastic convergence have been typically applied to regional income benchmarked against national income. However, this approach could be sensitive to the choice of base region or state and susceptible to aggregation bias.¹ 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. Let y_{it} be real per capita income data in US state i at time t , where $i = 1, \dots, N$ and $t = 1, \dots, T$. Pesaran's pair-wise approach is based on the examination of the time series properties of all $N(N-1)/2$ possible real

¹ 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.

per-capita income gaps between states i and j denoted $g_{jt} = y_{it} - y_{jt}$ where $i = 1, \dots, N-1$ and $j = i+1, \dots, N$. Defining \bar{Z}_{NT} as the fraction of gaps for which the unit-root hypothesis is rejected, Pesaran (2007) shows that as $T \rightarrow \infty$, the expected value of \bar{Z}_{NT} is equal to the nominal size of the unit root test statistic, α . Thus, evidence of stochastic convergence is found whenever $\bar{Z}_{NT} > \alpha$, which is consistent with definition 2.1 in Bernard and Durlauf (1995). According to this definition, in order for two states to converge their incomes must be cointegrated, and the cointegrating vector must be equal to $[1, -1]'$. This means that in the long run the incomes of the two states contain a common stochastic trend (i.e. they share the same determinants) and respond to permanent shocks with the same weights.

Under the pair-wise approach, it is possible that convergence occurs across all states even if cointegration cannot be detected for every pair we choose. An important issue that arises is whether \bar{Z}_{NT} is statistically different from α . In a significant departure from Mello (2011), we compute confidence intervals for \bar{Z}_{NT} . These are based on the derivation of the empirical distribution of the fraction of rejections using the bootstrap methodology, because the derivation of the variance of \bar{Z}_{NT} is complicated due to the fact that individual unit root tests are not independent from each other. Moreover, the pair-wise approach so far outlined implicitly assumes that all income pairs are cointegrated with a known cointegrating vector equal to $[1, -1]'$. This implies that the output forecasts of any state pair will be equal as $T \rightarrow \infty$. In practice, this might be regarded as a somewhat strong assumption. Thus, we might expect a case for weaker convergence whenever there is cointegration between two income series but the cointegrated vector is not known. For this reason, we further deviate from Mello (2011) by extending Pesaran's pair-wise approach through 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, we alternatively examine whether y_{it} and y_{jt} are cointegrated, possibly with a slope different from unity. This weaker form of convergence is consistent with definition 2.2 in Bernard and Durlauf (1995) where long-run income forecasts are proportional. This means that the two series contain a common stochastic trend insofar as they share the same determinants in the long-run, but in the long-run they respond to permanent shocks with different weights. To test for pair-wise cointegration between state incomes, we employ the Johansen (1988) maximum likelihood estimator of cointegrated vector autoregressive (VAR) models, which offers the advantage that normalisation on a particular state within each bivariate relationship is not an issue, insofar as it does not matter whether one looks at cointegration between y_{it} and y_{jt} , or between y_{jt} and y_{it} .

At this point it is worth mentioning that attempting to estimate a single Vector Error Correction model (VEC) that incorporates the per capita income series for all the US states would be highly problematic because of the large number of states and lags that would be involved in the modelling exercise, and the number of time observations that are available. 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 three key advantages over existing panel techniques in terms of (i) addressing the proportion of cases that are stationary (or cointegrated); (ii) allowing for the presence of cross-sectional dependence across states; and (iii) not being affected by the selection of a base or reference state.

3. Empirical analysis

3.1 Data description

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). Perhaps it is worth noticing that 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 national 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.

In addition to the state-level data described above, we also analyse per-capita personal income data obtained from the FRED dataset for 346 MSAs over the study period 1969-2009. The reason for analysing this alternative data source is not only to exploit the information provided by a highly disaggregated dataset, but also the fact that the pair-wise approach is applicable even when the number of individual series in the panel, N , is large relative to its time dimension, T , as shown by the analytical and Monte Carlo simulation results reported in Pesaran (2007). In particular, the simulation results presented by this author in Tables 1b and 2b illustrate that under the null of

² Alaska and Hawaii are excluded 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.

non-convergence, the rejection frequencies vary little across N for a given T , and increase rapidly with T . Overall, the pair-wise testing approach performs well for $T \geq 30$.

3.2 Discussion of results

We begin our empirical analysis of convergence by considering a standard approach based on unit root tests applied to income differentials in each state. For this, one could calculate relative income with respect to national real per-capita income (US), as in e.g. Mello (2011), or with respect to a particular state. For the latter, and in what can be regarded as a somewhat arbitrary choice, we select four benchmark states with respect to which all other incomes are measured, namely California (CA), Florida (FL), Illinois (IL) and New York (NY). As can be seen from the results reported in Table 1, there is considerable variation in the percentage of rejections, where in many cases the order of integration of the income differentials depends upon the chosen benchmark. For instance, in the cases of the states of Texas (TX) and Washington (WA) we reject the unit root hypothesis when their incomes are measured relative to the national average (US), Illinois (IL) and New York (NY), but not when measured relative to the states of California (CA) and Florida (FL). The results in Table 1 thus illustrate the importance of considering the additional information that emerges when one looks at state incomes in pairs of states, as opposed to the relative income of a given state in US income, or relative to the income of any other (somewhat arbitrarily chosen) state.

<Table 1 about here>

Table 2 reports the percentage of rejections of the augmented Dickey and Fuller (1979) (ADF) tests based on all 1128 bivariate income differentials. These tests are conducted at the 5 and 10% significance levels, the number of lags of the test regression is determined using the Akaike Information Criterion (AIC) with $p_{\max} = 4$, and a trend term is included if it is statistically

significant at the 5% level.³ 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 Elliott, Rothenberg and Stock (1996) (ERS) unit root test (these results are not reported here, though). If we follow Barro (1991) and others in making a distinction between Southern and Non-Southern states, the corresponding point fraction of rejections is slightly higher for Southern states (38.18%) compared to the 35.29% obtained for Non-Southern states.⁴ 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 move the analysis forward by employing the factor augmented sieve bootstrap approach outlined in Appendix 1. 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).

<Table 2 about here>

To implement the bootstrap, 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.⁵ 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

³ As indicated by Le Pen (2011), authors such as Carlino and Mills (1993), among others, argue that including a deterministic component in the output gap is not always incompatible with convergence. This situation could appear when regions or countries have not yet reached their long-run equilibrium. The deterministic trend would thus reflect a catch-up process. At the end of the catch-up phase, the deterministic trend becomes insignificant.

⁴ We classify states as Southern and Non-Southern following Barro and Sala-i-Martin (1992); see the notes in Table 1.

⁵ An application of the Bai and Ng (2002) test confirmed the presence of a single common factor driving US state incomes.

$ADF(1) = -3.075$ 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.

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. As can be seen from Table 3, qualitatively similar results are also obtained when considering Southern and Non-Southern states.

<Table 3 about here>

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 definition 2.2 in Bernard and Durlauf (1995), it can be argued 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 $[1, -1]$.⁶ 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 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.⁷

The results in Table 4 indicate that the proportion of rejections of $H_0 : r = 0$ (against $H_a : r \geq 1$) is 71.72% (i.e. 809 out of 1128 possible real income gaps) when setting $\alpha = 10\%$, while the corresponding number of rejections for $\alpha = 5\%$ is 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 thus 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

⁶ 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,

⁷ 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.

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).

<Tables 4 and 5 about here>

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 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 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 $|\beta_2^{(ij)} - 1|$, 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 $|\log y_{i0} - \log y_{j0}|$. 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 2000.⁸

⁸ We are most grateful to Gary Wagner who kindly provided these data, which were used in Garrett, Wagner and Wheelock (2007).

Estimation by OLS for the 809 cases where the trace test rejects the null hypothesis of no cointegration at the 10% significance level yields the following regression result:

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

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

where White's heteroscedasticity-consistent standard errors are reported in parentheses.⁹ The estimated positive coefficient on $\left| \log y_{i0} - \log y_{j0} \right|$ supports the view 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.¹⁰ To illustrate graphically the convergence results, Figure 1 presents the regression plane estimated above, along with the corresponding three-dimensional scatter plot of the data, where in order to improve the visualisation of the results the z -axis is restricted to the $[0, 1]$ interval. As can be seen from the figure, the shorter the distance between any two states, and the smaller the absolute difference between their initial incomes, the closer the cointegrating slope will be to unity; that is, the more likely is convergence definition 2.1 in Bernard and Durlauf (1995) to hold.

<Figure 1 about here>

In the final part of our investigation, we examine convergence using MSA-level data which

⁹ It should also be noted that when the null of homogeneity is not rejected at the 10% significance level (i.e. in 303 instances), the corresponding value of the dependent variable is set equal to zero.

¹⁰ To assess whether convergence is stronger when two states are both Southern, as defined in Barro and Sala-i-Martin (1992), it was also tried to include in the regression model a dummy variable that takes the value of one if this condition is satisfied, and zero otherwise. However, the estimated coefficient on this additional regressor was not found to be statistically different from zero.

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 (49) instances 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.¹¹ 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 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 $H_0 : r = 0$ (against $H_a : r \geq 1$) 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, when setting $\alpha = 5\%$, or when a unit root is imposed on the common factor \bar{y}_t (irrespective of the significance level).

¹¹ 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.

<Tables 6, 7, 8 and 9 about here>

3.3 Comparison to existing findings

The earlier pair-wise findings reported by Mello (2011) and Le Pen (2011) only focus on strong convergence and are based on point estimates making no allowance for confidence intervals. Our investigation, based on confidence intervals, suggests that in fact, a weaker form of convergence at state level is likely to hold in terms of state incomes sharing a common stochastic trend. While Le Pen (2011) provides evidence against convergence in the case of a European sample, it is most likely that the point pair-wise estimate for US states reported by Mello (2011) lies within a confidence interval based on the size of the individual unit root tests. The results that we report here are supportive of convergence across US states, but we find that relative state incomes might respond proportionately to shocks such as those based on saving, demography, technology and so on. Therefore, the long-run forecasts of relative state incomes are proportional rather than equal. This is a form of convergence that is not considered by Mello (2011) or by Le Pen (2011). The earlier study by Tsionas (2001) investigates common trends, but concludes that multiple common trends, rather than a single common trend, drive real incomes across states. On this basis, Tsionas (2001) is not able to conclude that convergence is present. Our finding of weak convergence is robust to a pair-wise grouping that permits us to distinguish between Southern and non-Southern states, and we do not find evidence that supports the presence of convergence clubs along these particular lines.

We do find that strong convergence is present when examining data at MSA-level. In convergence studies, the concept of aggregation bias suggests that it is more likely that convergence will be found at state-level. It should, however, be remembered that the MSA-level study period used in this study starts much later in 1969 rather than in 1929. On this basis, our study suggests that

strong convergence could be a phenomenon associated with the more recent decades. After all, the support for stochastic convergence offered by Carlino and Mills (1993) is reliant on allowing for a structural break in 1946. Lastly, we find that distance is an important consideration when reflecting on the strength of convergence. Indeed, we support the findings of Choi (2004) insofar as there is convergence with respect to geographically neighbouring states. While we find that convergence is stronger for contiguous states, our results also suggest that convergence is present with respect to the states separated by larger distances.

4. Concluding remarks

This paper employs 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 weaker form of convergence characterised by cointegration between state incomes where the elements of the cointegrating vector are unrestricted. However, we find that the likelihood of convergence between any two states is strongest for those states that have similar initial per-capita incomes and are closest in terms of distance. Additional analysis at a more disaggregated level using metropolitan statistical area data using a more recent study period provides stronger evidence of long-run convergence

characterised by stationary income differentials.

Table 1. ADF unit root t -tests on relative per capita personal income

State	Relative to US	Relative to CA	Relative to FL	Relative to IL	Relative to NY
AL	-2.432	-2.469	-1.636	-1.933	-3.533*
AR	-3.364*	-3.037*	-1.807	-1.741	-2.040
AZ	-3.365*	-3.288*	-1.669	-3.225*	-3.225*
CA	-1.376	n.a.	-2.287	-2.131	-2.876*
CO	-2.101	-2.164	-3.360*	-1.957	-2.471
CT	-2.081	-0.760	-2.662*	-2.030	-2.981*
DE	-1.256	-2.989*	-1.924	-1.666	-2.794*
FL	-2.193	-2.287	n.a.	-2.935*	-3.494*
GA	-3.624*	-2.634*	-2.047	-2.268	-3.846*
IA	-2.341	-1.750	-2.695*	-2.499	-2.309
ID	-3.368*	-3.409*	-1.202	-4.984*	-4.281*
IL	-2.606*	-2.131	-2.935*	n.a.	-1.765
IN	-2.473	-2.587	-0.844	-4.685*	-2.482
KS	-3.262*	-1.804	-2.276	-2.633*	-3.687*
KY	-1.628	-2.255	-1.809	-2.481	-2.310
LA	-1.717	-1.423	-2.424	-1.477	-2.460
MA	-2.468	-0.524	-3.905*	-1.653	-1.686
MD	-2.463	-0.801	-3.434*	-0.685	-2.155
ME	-2.028	-0.473	-4.471*	-2.305	-2.713*
MI	-1.258	-2.504	-0.222	-3.933*	-3.067*
MN	-1.942	-1.265	-3.752*	-1.699	-2.510
MO	-2.621*	-1.670	-2.915*	-1.941	-2.869*
MS	-2.665*	-2.570	-1.675	-2.141	-3.742*
MT	-1.842	-3.525*	-1.009	-3.566*	-2.557
NC	-3.644*	-2.479	-1.942	-3.139*	-4.096*
ND	-2.057	-1.852	-2.191	-2.071	-2.232
NE	-1.891	-1.435	-2.867*	-1.842	-2.060
NH	-1.539	-0.668	-3.024*	-1.291	-2.292
NJ	-2.328	-0.074	-2.650*	-0.782	-2.443
NM	-2.186	-1.970	-1.810	-2.739*	-3.802*
NV	-0.858	-3.079*	-0.978	-1.965	-3.001*
NY	-4.083*	-2.876*	-3.494*	-1.765	n.a.
OH	-0.982	-2.324	-0.822	-5.341*	-3.056*
OK	-1.570	-1.365	-2.274	-2.759*	-1.968
OR	-1.515	-2.575	-1.338	-3.827*	-3.370*
PA	-3.299*	-0.612	-3.195*	-1.225	-2.560
RI	-2.376	-1.663	-3.165*	-1.766	-5.185*
SC	-3.597*	-3.151*	-2.364	-2.178	-3.583*
SD	-2.250	-2.084	-2.098	-2.318	-2.727*
TN	-3.805*	-2.178	-1.422	-1.690	-3.223*
TX	-3.265*	-2.074	-2.013	-2.723*	-3.929*
UT	-2.466	-3.332*	-1.161	-4.255*	-3.750*
VA	-1.497	-3.020*	-1.316	-2.620*	-3.484*
VT	-1.033	-0.086	-4.082*	-0.702	-2.872*
WA	-3.068*	-2.524	-1.845	-3.734*	-3.585*
WI	-2.809*	-1.436	-2.632*	-2.205	-3.481*
WV	-2.398	-1.489	-3.527*	-1.917	-2.842*
WY	-3.204*	-2.635*	-2.205	-1.855	-2.965*

Notes: * denotes significance at the 10% level; n.a. indicates not applicable. Unit root tests include an intercept term and the number of lags is determined 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.

α	All states	Non-Southern states	Southern states
5%	33.78	35.29	38.18
10%	46.72	47.15	58.18

Notes: In line with Barro and Sala-i-Martin (1992), the following eleven states are classified as Southern: Alabama, Arkansas, Florida, Georgia, Kentucky, Louisiana, Mississippi, North Carolina, South Carolina, Tennessee and Virginia. The remaining states are Non-Southern. 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										
States	α	Mean	Median	SD	2.5%	5%	10%	90%	95%	97.5%
All	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
Non-Southern	5%	9.09	8.26	4.33	<u>3.00</u>	3.45	4.35	15.02	17.27	<u>19.52</u>
	10%	16.16	15.32	6.37	6.46	<u>7.36</u>	8.71	24.78	<u>27.78</u>	30.78
Southern	5%	10.63	9.09	6.43	<u>1.82</u>	1.82	3.64	18.18	21.82	<u>25.46</u>
	10%	19.08	18.18	8.85	5.46	<u>7.27</u>	9.09	30.91	<u>34.55</u>	40.00

Not imposing a unit root on common factor										
States	α	Mean	Median	SD	2.5%	5%	10%	90%	95%	97.5%
All	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
Non-Southern	5%	13.17	12.16	6.20	<u>4.35</u>	5.11	6.29	21.47	24.93	<u>28.38</u>
	10%	21.83	21.02	7.99	9.16	<u>10.21</u>	12.16	32.58	<u>36.49</u>	40.24
Southern	5%	13.08	12.73	7.08	<u>1.82</u>	3.64	5.46	21.82	25.46	<u>29.09</u>
	10%	22.82	21.82	9.40	7.27	<u>9.09</u>	10.91	34.55	<u>40.00</u>	43.64

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	α	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	α	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	α									
$r = 0$	$r \geq 1$	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	α									
$r = 0$	$r = 1$	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	α									
$r = 0$	$r \geq 1$	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	α									
$r = 0$	$r = 1$	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 $[1, -1]$

α	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

α	Mean	Median	SD	2.5%	5%	10%	90%	95%	97.5%
5%	12.65	12.46	2.29	<u>8.65</u>	9.19	9.88	15.64	16.72	<u>17.84</u>
10%	21.16	20.99	3.12	15.39	<u>16.26</u>	17.32	25.21	<u>26.59</u>	27.93

Not imposing a unit root on common factor

α	Mean	Median	SD	2.5%	5%	10%	90%	95%	97.5%
5%	12.68	12.49	2.25	<u>8.80</u>	9.30	9.98	15.62	16.68	<u>17.74</u>
10%	21.20	21.01	3.07	15.71	<u>16.44</u>	17.42	25.24	<u>26.57</u>	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	α	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	α	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

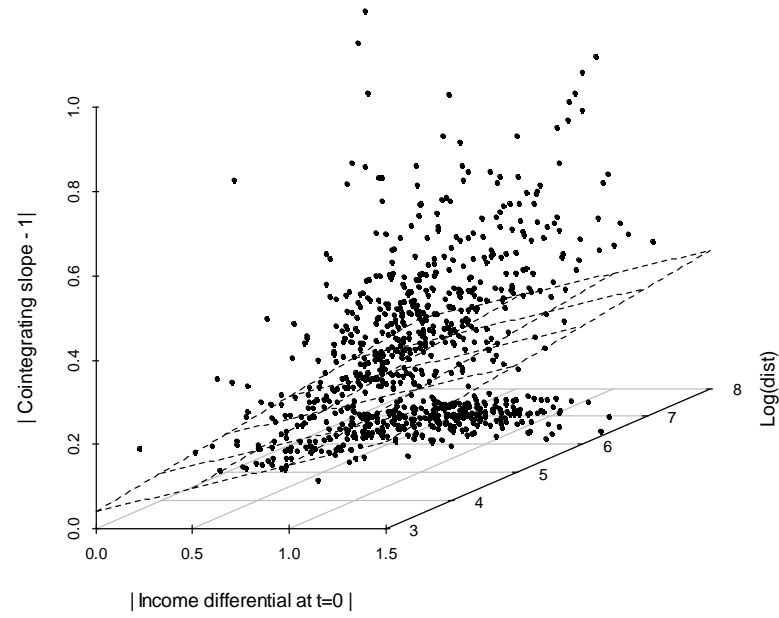
Johansen cointegration test	Mean	Median	SD	2.5%	5%	10%	90%	95%	97.5%		
Trace:											
Ho	Ha	α									
$r = 0$	$r \geq 1$	5%	19.57	18.23	6.32	<u>11.37</u>	12.33	13.33	27.38	31.34	<u>36.22</u>
$r = 0$	$r \geq 1$	10%	27.40	25.74	7.95	16.73	<u>17.98</u>	19.36	37.45	<u>42.66</u>	48.17
Max. Eigenvalue:											
Ho	Ha	α									
$r = 0$	$r = 1$	5%	17.72	16.79	4.65	<u>11.45</u>	12.15	13.01	23.12	26.31	<u>29.34</u>
$r = 0$	$r = 1$	10%	25.15	24.04	6.01	16.96	<u>17.91</u>	18.95	32.40	<u>36.21</u>	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	α									
$r = 0$	$r \geq 1$	5%	20.14	19.68	3.85	<u>13.97</u>	14.71	15.61	25.23	27.24	<u>28.91</u>
$r = 0$	$r \geq 1$	10%	28.28	27.78	4.83	20.25	<u>21.29</u>	22.55	34.58	<u>37.14</u>	38.99
Max. Eigenvalue:											
Ho	Ha	α									
$r = 0$	$r = 1$	5%	18.46	18.09	3.36	<u>13.06</u>	13.70	14.51	22.89	24.66	<u>26.16</u>
$r = 0$	$r = 1$	10%	25.96	25.59	4.23	19.00	<u>19.88</u>	20.97	31.56	<u>33.57</u>	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.

Figure 1. Three-dimensional scatter plot and regression plane of convergence determinants



Note: To improve the visualisation of the scatter plot, the observations for which the value of the z -axis is greater than one are excluded. Nonetheless, the regression plane is fitted using all the 809 observations for which the Johansen trace test rejects the null hypothesis of no cointegration at the 10% significance level.

References

- Bai, Jushan and Serena Ng, (2002) Determining the number of factors in approximate factor models. *Econometrica* 70, 191-221.
- Barro, Robert J. (1991) Economic growth in a cross section of countries. *Quarterly Journal of Economics* 106, 407-43.
- Barro, Robert. J. and Xavier Sala-i-Martin , (1991) Convergence across states and regions. *Brookings Papers on Economic Activity* Issue 1, 107-82.
- Barro, Robert. J. and Xavier Sala-i-Martin (1992) Convergence. *Journal of Political Economy* 100, 223-51.
- Barro, Robert. J. and Xavier Sala-i-Martin (1999) *Economic Growth*. MIT Press, Cambridge MA.
- Baumol, William. J. 1986) Productivity, growth, convergence and welfare: What the long-run data show, *American Economic Review* 76, 1072-85.
- Bernard, Andrew B. and Steven N. Durlauf. (1995) Convergence in international output. *Journal of Applied Econometrics* 10, 97-108.
- Carlino, Gerald A. and Leonard O. Mills (1993) Are U.S. regional incomes converging?: A time series analysis. *Journal of Monetary Economics* 32, 335-346.
- Cheung, Yin-Wong and Kon S. Lai (1995) Lag order and critical values of the augmented Dickey-Fuller test. *Journal of Business and Economic Statistics* 13, 277-280.
- Choi, Chi-Young (2004) A reexamination of output convergence in the US states: toward which level(s) are they converging? *Journal of Regional Science* 44, 713-741.
- Deckers, Thomas and Christoph Hanck (2013) Multiple testing for output convergence. *Macroeconomic Dynamics*, DOI: <http://dx.doi.org/10.1017/S1365100512000338>.
- Dickey David A. and Wayne A. Fuller (1979) Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American Statistical Association* 74, 427-431.

- Elliott, Graham, Rothenberg, Thomas J. and James H. Stock (1996) Efficient tests for an autoregressive unit root. *Econometrica* 64, 813-836.
- Evans, Paul and Georgios Karras (1996) Convergence revisited. *Journal of Monetary Economics* 37, 249-265.
- Garrett, Thomas A., Wagner, Gary A. and David C. Wheelock (2007) Regional disparities in the spatial correlation of state income growth, 1977–2002. *Annals of Regional Science* 41, 601-618.
- Im, Kyung S., Pesaran, M. Hashem and Yongcheol Shin (2003) Testing for unit roots in heterogeneous panels. *Journal of Econometrics* 115, 53-74.
- Inklaar, Robert and Marcel P. Timmer (2009) Productivity convergence across industries and countries: The importance of theory-based measurement. *Macroeconomic Dynamics* 13(S1), 218–240.
- Johansen, Soren (1988) Statistical analysis of cointegration vectors, *Journal of Economic Dynamics and Control* 12, 231-254.
- Krugman, Paul (1991) Increasing returns and economic geography. *Journal of Political Economy* 99, 433–499.
- Krugman, Paul and Anthony J. Venables (1995) Globalisation and the inequality of nations. *Quarterly Journal of Economics* 110, 857–880.
- Levin, Andrew, Lin, Chien-fu and Chia-Shang Chu (2002) Unit root tests in panel data: Asymptotic and finite-sample properties. *Journal of Econometrics* 108, 1-24.
- Le Pen, Yannick (2011) A pair-wise approach to output convergence between European regions. *Economic Modelling* 28, 955-64.
- MacKinnon, James G., Haug, Alfred and Leo Michelis. (1999) Numerical Distribution Functions of Likelihood Ratio Tests for Cointegration. *Journal of Applied Econometrics* 14, 563-577.

- Maddala, Gangadharrao .S. and Shaowen Wu (1999) A comparative study of unit root tests with panel data and a new simple test. *Oxford Bulletin of Economics and Statistics* 61, 631-652.
- Mankiw, N. Gregory., Romer, David and David N. Weil (1992) A contribution to the empirics of economic growth. *Quarterly Journal of Economics* 107, 407-37.
- Mello, Marcello, (2011) Stochastic convergence across US states. *Macroeconomic Dynamics* 15,160-183.
- Papageorgiou, Chris and Fidel Perez-Sebastian (2004) Can transition dynamics explain the international output data? *Macroeconomic Dynamics* 8, 466-492.
- Pedroni, Peter (2001) Purchasing power parity in cointegrated panels, *Review of Economics and Statistics* 83, 727-731.
- Pesaran, M. Hashem (2007) A pair-wise approach to testing for output and growth convergence. *Journal of Econometrics* 138, 312-355.
- Pesaran, M. Hashem, Smith, Ron P., Yamagata, Takashi and Liudmyla Hvozdyk (2009) Pairwise tests of purchasing power parity. *Econometric Reviews* 28, 495-521.
- Quah, Danny (1993) Galton's fallacy and tests of the convergence hypothesis. *Scandinavian Journal of Economics* 95, 427-43.
- Quah, Danny (1997) Empirics for growth and distribution: polarization, stratification and convergence club. *Journal of Economic Growth* 2, 27-59.
- Quintos, Carmela E. (1995) Sustainability of the deficit process with structural shifts. *Journal of Business and Economic Statistics* 13, 409-417.
- Sala-i-Martin, Xavier (1996) The classical approach to convergence analysis. *Economic Journal* 106, 1019-1036.
- Tsionas Efthymios 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. Bootstrapping the \bar{Z}_{NT} statistic

To bootstrap the \bar{Z}_{NT} statistic we follow Pesaran, Smith, Yamagata and Hvozdnyk (PSYH) (2009).

The model setup considered by these authors is based on the following set of equations:

$$y_{it} = \boldsymbol{\alpha}'_i \mathbf{d}_t + \boldsymbol{\gamma}'_i \mathbf{f}_t + \varepsilon_{it} \quad (2)$$

$$\Delta \varepsilon_{it} = \eta_i + \lambda_i \varepsilon_{i,t-1} + \sum_{l=1}^{p_i} \psi_{il} \Delta \varepsilon_{i,t-l} + \nu_{it} \quad (3)$$

$$\Delta f_{st} = \boldsymbol{\mu}'_s \mathbf{d}_t + \phi f_{s,t-1} + \sum_{l=1}^{p_s} \xi_{sl} \Delta f_{s,t-l} + e_{st} \quad (4)$$

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

In line with PSYH, we use the cross-sectional average of y_{it} , denoted $\bar{y}_t = N^{-1} \sum_{i=1}^N y_{it}$, 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 regressed on \bar{y}_t :

$$y_{it} = \hat{\alpha}_i + \hat{\delta}_i t + \hat{\gamma}_i \bar{y}_t + \hat{\varepsilon}_{it}. \quad (5)$$

In Appendix 2 we report the results of estimating the factor equations for the state-level data used in the paper. It should be noted that in these factor equations the 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 \bar{y}_t , which may be $I(0)$ or $I(1)$.

This involves estimating the following ADF(p) regression for \bar{y}_t :

$$\Delta \bar{y}_t = \hat{\mu} + \hat{\phi} \bar{y}_{t-1} + \sum_{l=1}^p \hat{b}_l \Delta \bar{y}_{t-l} + \hat{e}_t, \quad (6)$$

which may also include a trend term if it is statistically significant, and where 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 (6), i.e. letting $\hat{\phi} = 0$, and allowing for a drift implies the following restricted version of (6):

$$\Delta \bar{y}_t = \hat{\mu} + \sum_{l=1}^p \hat{c}_l \Delta \bar{y}_{t-l} + \hat{u}_t. \quad (7)$$

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:

$$\bar{y}_t^{(b)} = \hat{\mu} + \bar{y}_{t-1}^{(b)} + \sum_{l=1}^p \hat{c}_l \Delta \bar{y}_{t-l}^{(b)} + \hat{u}_t^{(b)}, \quad (8)$$

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 (7), 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:

$$y_{it}^{(b)} = \hat{\alpha}_i + \hat{\delta}_i t + \hat{\gamma}_i \bar{y}_{it}^{(b)} + \hat{\varepsilon}_{it}^{(b)}, \quad (9)$$

where $\hat{\alpha}_i$, $\hat{\delta}_i$ and $\hat{\gamma}_i$ are the OLS estimates of α_i , δ_i and γ_i in (5), respectively, and

$$\varepsilon_{it}^{(b)} = \hat{\eta}_i + (1 + \hat{\lambda}_i) \varepsilon_{i,t-1}^{(b)} + \sum_{l=1}^{p_i} \hat{\psi}_{il} \Delta \varepsilon_{i,t-l}^{(b)} + \nu_{it}^{(b)}, \quad (10)$$

where bootstrap residuals $v_{it}^{(b)}$ are generated by randomly drawing with replacement from the set of estimated residuals v_{it} in (3), 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 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 $ADF(p)$ or $ERS(p)$ test. The procedure already described is repeated $b = 1, \dots, B$ times to derive the empirical distribution of the bootstrapped fraction of rejections.

Appendix 2. 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