JOURNAL OF REGIONAL SCIENCE, VOL. 44, NO. 4, 2004, pp. 713–741

A REEXAMINATION OF OUTPUT CONVERGENCE IN THE U.S. STATES: TOWARD WHICH LEVEL(S) ARE THEY CONVERGING?

Chi-Young Choi*

Department of Economics, University of New Hampshire, NH 03824.
E-mail: cy.choi@unh.edu

ABSTRACT. This paper reexamines the issue of output convergence among the 48 states in the continental United States. Implementing multiple panel data techniques to state per capita output during the period 1929–2001 reveals little evidence of stochastic convergence in all 48 states, but some evidence among collections of states at the regional level. This observation may suggest that output convergence in the United States has proceeded among geographically neighboring states rather than among distant states, notwithstanding the nearly complete integration of product and factor markets. Our findings appear to be robust to a subsample analysis, although the intensity of convergence varies with the choice of output measure and deflator. Industrial structures and geographic proximity are considered as potential explanations for the regional pattern of output growth dynamics.

1. INTRODUCTION

Are U.S. states converging? Due to the almost homogeneous institutional environments and the highly integrated markets for products and factors, the states are believed to satisfy the underlying conditions of the convergence hypothesis in the standard neoclassical growth model. This belief has been, in general, strengthened by ample empirical evidence in the convergence literature (e.g., Barro, 1991; Barro and Sala-i-Martin, 1992), but more recent studies (e.g., Johnson and Takeyama, 2000; Rey and Montouri, 1999; Tsionas, 2001) have provided somewhat challenging evidence to this conventional view.

The purpose of this paper is to reexamine the issue of output convergence among 48 continental U.S. states during the period 1929–2001 in the context of stochastic convergence.[1] By investigating persistence of shocks to relative

*The author is grateful to Stephen Cecchetti, Paul Evans, Nelson C. Mark, Suzanne McCoskey, Masao Ogaki, Jack Strauss, and Arnold Zellner for helpful comments and wishes to thank two anonymous referees and Andrew F. Haughwout (co-editor) for constructive suggestions. Any remaining errors are the author's.

Received: February 2003; Revised: December 2003; Accepted: March 2004.

[1]According to Carlino and Mills (1993), stochastic convergence is a prerequisite for cross-section convergence.

Blackwell Publishing, Inc., 350 Main Street, Malden, MA 02148, USA and 9600 Garsington Road, Oxford, OX4 2DQ, UK.

per capita outputs in a dynamic and stochastic environment, we attempt to determine whether economies are converging over time. Temporary shocks characterized by stationary relative outputs indicate that economies are stochastically converging, whereas relative outputs with either a deterministic term or a unit-root component invalidate the definition of stochastic convergence. In this vein, unit-root or cointegration testing procedures are often used to assess stochastic convergence.

The current study is distinctive from the existing literature in several dimensions. First, on the methodological plane, we employ multiple panel data techniques in the framework of confirmatory analysis. Although it is broadly agreed that panel time-series techniques achieve a clear power gain over univariate counterparts, there remain a couple of critical issues in association with the use of popular panel techniques. First, many earlier panel tests are constructed under the restrictive assumption of cross-sectional independence, which knowingly leads to serious size distortions when the assumption is violated (O'Connell, 1998).[2] Given that sub-economies within a nation are subject to common output shocks, cross-sectional dependence is present almost by nature in output data. Hence the size distortion problem due to the failure to account for cross-sectional dependence could overshadow the potential benefit of power gain in panel data analysis. Second, in light of the all-or-nothing structure of null hypotheses, it is not clear what we learn from a rejection by panel tests because the rejection is consistent with numerous alternatives. Take a popular panel unit-root test for instance. When we reject the null hypothesis that all series in the panel are unit-roots, one cannot determine whether the rejection is driven by all stationary series or as few as only one stationary series. Unfortunately, flipping the null hypothesis around does not facilitate interpretation. To circumvent this problem, we follow Choi (2002) to adopt *confirmatory analysis* that involves comparing outcomes of panel tests under competing null hypotheses. The basic idea of this strategy is that joint testing can improve the reliability of inference over single testing alone, particularly when an outcome from one panel test is reinforced by that from another test. The current study employs three panel testing procedures toward this end. The panel G-test under the null hypothesis of stationarity due to Choi (2002) is matched with two popular panel tests under the unit-root null hypothesis proposed by Levin, Lin, and Chu (2002) and Im, Pesaran, and Shin (2003) that have been popularly used for studying output growth rates and levels of real exchange rates. Because these tests are constructed under the assumption of cross sectional independence, we utilize a nonparametric bootstrap method to draw inferences from the bootstrapped distribution to account for contemporaneous correlations.

[2]More recent panel tests are designed to accommodate cross-sectional dependence by construction (e.g., Bai and Ng, 2002 and Phillips and Sul, 2003).

Next, we use state-level data instead of aggregated regional data more commonly adopted in the study of regional convergence. Regional economies are comprised of diverse state economies with potential cross-state variations in the behavior of output just as the national economy is a composite of heterogeneous regional subeconomies. Therefore, regionally aggregated data may mask the dynamic interactions of individual states by ignoring a large amount of valuable information entrenched in state specific characteristics. In addition, state-level data enable the distinction between convergence at the *national* level and convergence at the *regional* level. This distinction is intuitively important on the grounds that localized permanent shocks may hamper nationwide convergence but not necessarily convergence among states that form particular regions. Of course, using state-level data is not devoid of critics. As administrative units that do not account for functional aspects of geographical classification, U.S. states may not be perfect observational units for the analysis of regional convergence. Nonetheless they must be the most relevant units of policy making with certain degree of autonomy in fiscal policy and legal systems that are believed to have significant economic implications. Furthermore, U.S. states offer a long span of output data. A related issue then emerges as to the choice of specific state-level output measures. Among several output measures available, state personal income and state personal earnings are considered in the present study. Since the two measures differ in the subcomponents, they are expected to take different profiles on convergence. Carlino and Mills (1996) find that evidence of stochastic convergence wanes using earnings data that exclude state distribution of transfer payments, which may smooth the effects of deviations of state incomes.

Third, we consider both the national gross national product (GNP) deflator and the metropolitan area Consumer Price Indices (CPIs) as a proxy for state price index to deflate nominal state output. As pointed out by Barro and Sala-i-Martin (1992), use of the national deflator may lead to mismeasurements of real state and regional incomes unless the purchasing power parity (PPP) holds across states. In a recent study on the dynamics of price indices of 19 U.S. cities, Cecchetti, Mark, and Sonora (2002) find that relative price levels among major U.S. cities exhibit a very slow mean reversion pattern over the period from 1918 to 1995. Taken together, use of the national deflator could misguide inference on the dynamic properties of real state per capita incomes by disregarding dynamics embedded in relative prices between states. In this vein, metropolitan area CPIs may better reflect the interstate price differences, but unfortunately they are available only for 15 states during the entire sample period. To compromise, we compare the results based on both price deflators in our empirical analysis.

Before proceeding, one needs to recognize that stochastic convergence is based on a strong assumption that sample moments of data are interpretable as population moments for the underlying stochastic process. This assumption, however, can be problematic if economies are in transition toward a steady state as characterized by structural breaks. In this case, standard testing tools may fail to capture the transition property in the perspective of stochastic convergence even though economies are in fact converging. Considering that the possible presence of break points in the U.S. regional convergence process has been suggested by a number of authors (e.g., Carlino and Mills, 1993, 1996; Lowey and Papell, 1996), this feature of stochastic convergence may pose a serious limitation in the reliability of our analysis. To sidestep this potential problem, we conduct a subsample analysis for the period 1947–2001 following Carlino and Mills to posit 1946 as a break point.[3] A marked difference in the results between subsample and full sample analyses may signify the presence of a structural change around 1946 in the convergence process.

Our main empirical findings can be summarized as follows. First, we find little evidence of stochastic convergence in all 48 continental states, but some evidence among collections of states at the regional level. Evidence of regional convergence becomes stronger in the subsample analysis, but only marginally. Second, the degree of regional convergence appears to vary with the choice of output measures and deflators. Use of personal earnings instead of personal incomes tends to weaken overall evidence of convergence and so does consideration of metropolitan area CPIs as a proxy for state price index.

The remainder of this paper is organized as follows. The next section reviews the definition of stochastic convergence in the panel framework, and discusses econometric methods used in the current study. Section 3 describes the data and presents related issues. Section 4 reports the empirical results of our study. Section 5 explores potential explanations for the empirical results with concentration on the roles of industrial structures and geographic distance. Finally, Section 6 concludes.

2. METHODOLOGY

Stochastic convergence focuses on time-series properties of output data in a dynamic and stochastic environment taking initial conditions as given. By investigating persistence of shocks to relative real per capita outputs, we attempt to determine whether economies are converging over time. If relative outputs or output disparities between two economies follow a mean stationary stochastic process, it indicates that economies move over time toward a constant time-invariant differentials equilibrium, or stochastically converge. If output disparities contain either a deterministic term or a unit-root component, possibly due to permanent technology shocks as asserted in the endogenous growth model, the definition of stochastic convergence will be violated. Bernard

[3]Based on relative per capita earnings, Carlino and Mills (1996) claim that the U.S. states and regions have achieved convergence by 1946. Lowey and Papell (1996) also find two break points in the 1940s. Some studies suggest 1978 as a candidate for another break point in trend around which regional per capita incomes appeared to diverge sharply.

and Durlauf (1995) consider the following time-invariant Wold representation of income differences to test for the convergence hypothesis

$$(1) \qquad y_{i,t} - y_{j,t} = \kappa_{i,j} + \sum_{r=0}^{\infty} \pi_{i,j,r}\varepsilon_{i,j,t-r} \quad i,j = 1,2,3,\ldots N, \quad i \neq j$$

where $\pi_{i,j,r}$ is square-summable. Convergence holds if the income difference $y_{i,t} - y_{j,t}$ is stationary for all i,j pairs in a group and it does not hold if the difference persists permanently. Because of the stationarity requirement, the existence of stochastic convergence is related to the unit-root nonstationarity hypothesis and hence is tested by relevant time-series techniques such as unit-root and cointegration testing procedures. In general, studies on stochastic convergence (e.g., Bernard and Durlauf, 1995) have provided little evidence of international convergence as the time-series tests fail to reject the conventional null hypothesis of no convergence. However, since this lack of evidence is often attributed to the poor discriminatory power of the standard time-series techniques in small samples, subsequent studies have sought to overcome the problem by employing panel data techniques in evaluating the convergence hypothesis. Evans and Karras (1996) provide a suitable modification of Equation (1) to test for stochastic convergence in panel framework.

Consider a collection of economies $1,2,\ldots,N$ in a stochastic world that have eventual access to the same body of technological knowledge. Then, economies $1,2,\ldots,N$ are said to converge if, and only if, a common trend a_t and finite parameters $\mu_1,\mu_2,\ldots,\mu_N$ exist such that

$$(2) \qquad \lim_{k\to\infty} E_t(y_{i,t+k} - a_{t+k}) = \mu_i \quad \text{for} \quad i = 1,2,\ldots,N$$

where $y_{i,t}$ represents the logarithm of per capita output for economy i during period t and a_t denotes the common trend followed by the economies. Evans and Karras reformulate Equation (2) as follows to incorporate the feature of conditional convergence in panel framework[4]

$$(3) \qquad \Delta(y_{i,t} - \bar{y}_t) = \delta_i + \rho_i(y_{i,t-1} - \bar{y}_{t-1}) + \sum_{k=1}^{p_i} \phi_{i,k}\Delta(y_{i,t-k} - \bar{y}_{t-k}) + u_{i,t}$$

where $i = 1,2,\ldots,N$, $t = 1,2,\ldots,T$, and ϕs are parameters such that all roots of $\phi_{i,k}L^j$ lie outside the unit circle where L denotes the lag operator. Since stochastic convergence is stipulated as the log of per capita income in one economy relative to that of the group average being stationary, testing for convergence in equation (3) is analogous to testing whether $y_{i,t} - \bar{y}_t$ is stationary for *all* series in the panel. In this context, standard panel unit-root tests can be applied to ask whether shocks to per capita income disparities $y_{i,t} - \bar{y}_t$ are permanent.

[4]See the Appendix for the derivation.

For all its appeal, however, popular panel unit-root tests suffer from a couple of critical issues that need to be addressed in practice. First, many earlier panel tests are constructed under the restrictive assumption of cross-sectional independence (u_{it} in (3) is contemporaneously uncorrelated), which is easily violated in many empirical applications of interest. As noted by O'Connell (1998), the failure to account for contemporaneous correlation across individual series results in serious size distortions that could outweigh the potential benefits of power gain from enlarged sample size. Second and more important, given the maintained structures of null hypotheses, it is not clear what we learn from rejection of the null hypotheses because the rejection is consistent with numerous alternatives. In a standard panel unit-root test for instance, rejection of the unit-root null can be triggered either by *all* stationary series in the panel or by as few as *one* stationary series. Unfortunately, flipping the null hypothesis around does not facilitate interpretation. In view of the substantial cross-section variation often observed in popular panel data sets, it is not desirable to evaluate the convergence hypothesis under the all-or-nothing criterion as we cannot exclude the possibility of partial convergence, or partial divergence, in which convergence holds for some economies while not for all.

A major methodological contribution of the current study rests in adopting strategies to bypass these problems of panel data techniques in the study of the convergence hypothesis. To be more concrete, we use bootstrap methods recommended by Maddala and Wu (1999), Mark and Sul (2001), and Cecchetti et al. (2002) to take into account cross-sectional dependence across individual series. To deal with the ambiguity in interpreting rejection of null hypotheses, we employ the strategy of *confirmatory analysis* advocated by Choi (2002) who reports that simple comparison of the outcomes of two panel tests under competing null hypotheses can substantially improve the reliability of inference over the standard practice of using one panel test alone especially when the two outcomes corroborate each other. To this end, we implement multiple panel tests. Two popular tests under the null hypothesis of all unit-root series in the panel, respectively, due to Levin, Lin, and Chu (2002, hereafter LLC) and to Im, Pesaran, and Shin (2003, hereafter IPS), are paired with a panel test under the null hypothesis of all stationary series, and the panel G-test proposed by Choi (2002, hereafter PG). The Appendix presents a brief description of the nonparametric bootstrap method and the panel testing procedures along with their finite sample performances in the context of confirmatory analysis.

The current study also distinguishes itself from previous ones by considering both *national* average and *regional* averages for $\bar{y}_t$ in Equation (3). This distinction is intuitively appealing in the sense that permanent shocks to the national economy also affect regional economies, but regional per capita income should not move away from the national average income unless permanent regional specific shocks, such as localized technology shocks, induce regional income deviations. If regional specific shocks prevail, per capita output should converge regionwide rather than nationwide.

3. DATA AND RELATED ISSUES

Our dataset comprises annual Bureau of Economic Analysis (BEA) data on personal income and earnings, population for 48 continental U.S. states during the period 1929–2001, and deflator data for the corresponding period. Real per capita personal incomes are constructed by dividing nominal per capita output for each state by the relevant deflators.[5] The underlying data have been collected from the *State Personal Income* CD-ROM (2002) for the nominal per capita state income, earnings, and population, from the *FRB Saint Louis Database (FRED)* for the GNP deflator, and from the BLS homepage (http://www.bls.gov/cpi/home.htm) for the metropolitan-area CPI data. The resulting series are real per capita personal income and earnings with annual observations covering 1929–2001.

A couple of issues with regard to data are worth further discussion. As pointed out by Barro and Sala-i-Martin (1992), use of a common deflator may lead to mismeasurement of the real per capita income if purchasing power parity (PPP) does not hold across the states. Due to the nonavailability of state price indices, this issue has been largely overlooked in the literature.[6] However, in their recent study on the dynamics of price indices of 19 U.S. cities, Cecchetti, Mark, and Jonora (2002) report that relative price levels among major U.S. cities exhibit an exceptionally slow mean reversion during the period 1918–1995. Combined, it seems fair to argue that the common deflator cannot properly capture the underlying dynamics of relative real per capita incomes across states. This motivates us to adopt the metropolitan area CPIs as an alternative. Unfortunately, only 15 states have suitable data for the entire sample period under study, which poses a serious restriction on our analysis.[7] Nonetheless it is still worthwhile to compare the dynamic properties of these 15 selected state real per capita incomes based on the two deflators to trace the impact of price dynamics on output convergence.

Another crucial data issue is the choice of time series for per capita output. In the convergence literature three output measures are popularly used: state personal income, state personal earnings, and gross state product (GSP).

[5]Note that deflation serves no purpose if the same national price index is used to deflate nominal state incomes. Specifically let $y_{i,t}$ be the deflated real per capita income of state i at time t such that $y_{i,t} = Y_{i,t} - P_t$ where $Y_{i,t}$ represents the log nominal per capita income of state i at time t and P_t denotes the log national price index at t. Then income differential $y_{i,t} - \bar{y}_t = (Y_{i,t} - P_t) - (\bar{Y}_t - P_t) = Y_{i,t} - \bar{Y}_t$ is nothing other than relative nominal incomes.

[6]Major exceptions include Mitchener and McLean (1999), Johnson and Takeyama (2000), and Carlino and Sill (2001), who used metropolitan CPI data as a proxy for state price index.

[7]For some states that have more than two CPIs available, CPIs are combined with equal weights. For example, to form the price index of Ohio, the CPIs for two metropolitan areas in Ohio, Cleveland and Cincinnati, are equally weighted.

Among them, GSP may be most compatible with the mathematics of convergence given that it measures the output produced within a given state by all factors used in the state regardless of their owners' residence. However, GSP is not considered here due to the comparatively short time-span. The remaining two measures differ in their subcomponents: personal earnings include wages and salaries, other labor income, and proprietors' income, whereas personal income is measured as the sum of personal earnings plus dividends, interest, rent, and transfer payments less personal contributions for social insurance. As a result, the two measures may significantly differ in states with large numbers of residents who earn income in other states. This difference is noted by Carlino and Mills (1996), who contend that personal earnings are more suitable to convergence resulting from factor migration. They find stronger evidence of stochastic convergence using per capita income series rather than per capita earnings, and attribute it to the role of state distribution of transfer payments in smoothing the effects of deviations of state per capita earnings. In the current paper, both output measures are studied for the purpose of comparison.

4. EMPIRICAL RESULTS

Using National Price Index as Deflator

We begin our empirical analysis with the diagnostic results from univariate and panel test techniques applied to the aggregate output data of the eight BEA regions. The first panel of Table 1 reports the results from two powerful univariate tests, the DF-GLS test due to Elliott, Rothenberg, and Stock (1996) under the unit-root null and a stationarity test proposed by Kwiatkowski, Phillips, Schmidt and Shin (1992), now commonly referred as KPSS. The two tests reach agreements in four out of eight regions at the 10 percent significance level, but the evidence of convergence is rather mixed.[8] This finding is mirrored in the panel data analysis where the PG and IPS tests jointly reject their respective nulls, implying a possible mix of convergent and divergent regions in the panel.[9] Overall evidence of stochastic convergence based on regional aggregates is therefore inconclusive, in accordance with the findings by Carlino and Mills (1993).[10]

This redirects us to the state level data to gain some insights into which states are responsible for the inconclusive results. Table 2 presents the results from the joint application of the KPSS and DF-GLS tests to the state

[8]Although significance level is fairly consequential to the test results, there are no definite guidelines on the appropriateness of the choice of significance level. Throughout the paper a 10 percent significance level is used for the sake of consistency.

[9]Given that panel tests are built under the null hypothesis that all series are either stationary or unit-root, the two nulls are subject to reject if a panel, in fact, consists of both stationary and unit-root series.

[10]Carlino and Mills find that stochastic convergence holds in three of eight regions after inserting an exogenously determined break point at 1946. However, no break point is allowed in our study.

TABLE 1: Test Results for Regional Aggregates

1929–2001				1947–2001		
			A. Univariate Tests			
DF-GLS	KPSS	Inference	Regions	DF-GLS	KPSS	Inference
−2.2826	0.2465	I(0)	New England	−1.3593	0.2504	–
−3.9236	0.4386	–	Mideast	−1.4701	0.2274	–
−1.0184	0.5916	I(1)	Great Lakes	−1.1463	0.4634	I(1)
−1.8817	0.3988	–	Plains	−2.0293	0.0937	I(0)
−1.7021	0.5992	–	Southeast	−2.0444	0.4830	–
−2.1069	0.4528	–	Southwest	−2.0565	0.2388	I(0)
−2.1460	0.1433	I(0)	Rocky Mtns	−2.3880	0.3742	–
−1.1522	0.5984	I(1)	Far West	−1.2544	0.5103	I(1)
			B. Panel Tests			
PG	IPS	Inference	Panel	PG	IPS	Inference
0.000	0.002	–	8 Regions	0.000	0.036	–

Note: The null hypothesis of the KPSS [the DF-GLS] test is that the series under study is I(0) [I(1)]. Hence nonrejection of the null by the KPSS test and rejection of the null by the DF-GLS test suggest evidence of stochastic convergence. Inferences are made at the 10 percent significance level and the corresponding critical values of the DF-GLS and the KPSS tests statistics are −1.62 and 0.348, respectively. Lag lengths for the DF-GLS test are selected by using the MBIC method proposed by Ng and Perron (2001) and the Lag lengths for the KPSS test are chosen by setting the maximum lag length to be integer $12(T/100)^{1/4}$. Andrews-Monahan's (1992) prewhitening method is used for long run variance calculation in the KPSS test. Entries for panel tests are p-values.

per capita incomes. Notice that both national and regional averages are considered here for the common trends $\bar{y}_t$.[11] The results in Table 2 illustrate a couple of interesting points. First, the two tests seldom yield harmonious outcomes on convergence at the regional level, strengthening our initial intuition that a regional economy is a composite of diverse state economies. Second, confirmatory evidence favorable to convergence can be found in 14 states using national average and in 12 states using regional average during the entire sample period. Interestingly the number declines to 10 in the subsample using the national average while it rises to 16 using regional averages implying a marginal gain in convergence evidence when regional averages are used for the common trend.[12]

[11]It is important to make a distinction between the two cases because the test results are based on the identical test procedures but different common trends. As a consequence, the test results have different implications on convergence. Evidence of convergence using the national average implies that per capita incomes of states converge toward the national average level, whereas evidence of convergence based on regional averages suggests that per capita income of states converge toward respective regional average income.

[12]Based on a single univariate unit-root test, Carlino and Mills (1996) find evidence of stochastic convergence in 18 states without a break point in the convergence process, but in 29 states with a break point.

TABLE 2: Univariate Tests on State Level Data

BEA Regions	State	1929–2001						1947–2001					
		National Average			Regional Average			National Average			Regional Average		
		DF-GLS	KPSS	Joint	DF-GLS	KPSS	Joint	DF-GLS	KPSS	Joint	DF-GLS	KPSS	Joint
New England	CT	−2.3967	0.3344	I(0)	−2.3093	0.3325	I(0)	−1.2637	0.1097	–	−1.8259	0.1489	I(0)
	ME	−1.8545	0.3422	I(0)	−3.4237	0.1005	I(0)	−2.5089	0.1501	I(0)	−2.9057	0.2630	I(0)
	MA	−2.7773	0.2721	I(0)	−1.9212	0.2044	I(0)	−1.1711	0.2223	–	−0.9970	0.2325	–
	NH	−2.5039	0.1294	I(0)	−1.3902	0.5576	I(1)	−1.1487	0.3861	I(1)	−2.2306	0.4839	–
	RI	−1.9423	0.4345	–	−1.5202	0.5745	I(1)	−1.7400	0.2548	I(0)	−1.3338	0.5296	I(1)
	VT	−2.2974	0.1422	I(0)	−1.4350	0.4513	I(1)	−1.2196	0.3587	I(1)	−1.3322	0.1981	–
Mid-East	DE	−1.9389	0.5582	–	−1.0664	0.5061	I(1)	−1.3371	0.4301	I(1)	−0.8673	0.4572	I(1)
	MD	−1.8361	0.2596	I(0)	−2.4239	0.5759	–	−1.8479	0.4104	–	−1.3036	0.4441	I(1)
	NJ	−1.9654	0.3765	–	−1.9144	0.5455	–	−1.6250	0.0941	I(0)	−1.3920	0.4466	I(1)
	NY	−3.3892	0.4830	–	−3.0757	0.5108	–	−1.5022	0.3528	I(1)	−1.6420	0.3517	–
	PA	−2.4743	0.4819	–	−1.7179	0.4126	–	−1.2720	0.3723	I(1)	−2.3733	0.1626	I(0)
Great Lakes	IL	−3.3614	0.6022	–	−3.3486	0.4813	–	−1.0773	0.4545	I(1)	−2.6684	0.4116	–
	IN	−0.9625	0.3831	I(1)	−3.3440	0.3763	–	−1.2992	0.4544	I(1)	−2.1480	0.1559	I(0)
	MI	−1.3061	0.5889	I(1)	−1.8437	0.4949	–	−1.3287	0.4472	I(1)	−3.0143	0.3920	–
	OH	−1.3292	0.6084	I(1)	−1.6829	0.4596	–	−0.9938	0.4696	I(1)	−1.5990	0.2142	–
	WI	−1.9413	0.5308	–	−1.2385	0.5446	I(1)	−1.0832	0.4205	I(1)	−1.2322	0.4547	I(1)
Plains	IA	−2.1550	0.1399	I(0)	−1.4384	0.5545	I(1)	−1.7457	0.3920	–	−1.4052	0.4259	I(1)
	KS	−2.0254	0.2882	I(0)	−4.0064	0.0931	I(0)	−1.3594	0.2304	–	−2.9604	0.0958	I(0)
	MN	−0.4784	0.5044	I(1)	−1.8531	0.1300	I(0)	−0.3886	0.5144	I(1)	−1.6186	0.4922	I(1)
	MO	−2.5308	0.4191	–	−2.1495	0.3852	–	−1.1319	0.3758	I(1)	−2.1251	0.1280	I(0)
	NE	−2.0611	0.1853	I(0)	−3.5585	0.4312	–	−2.4904	0.3419	I(0)	−2.4061	0.2616	I(0)
	ND	−2.1725	0.3180	I(0)	−2.5404	0.3182	I(0)	−3.2552	0.0921	I(0)	−3.2931	0.0814	I(0)
	SD	−2.2563	0.3402	I(0)	−2.3536	0.3544	–	−2.9299	0.0865	I(0)	−1.8617	0.1587	I(0)

	AL	−1.5469	0.5759	I(1)	−1.7941	0.4773	–	−1.5463	0.4884	I(1)	−1.9930	0.3804	–
	AR	−1.6933	0.5725	–	−1.9175	0.4936	–	−1.5546	0.4443	I(1)	−1.6610	0.2896	I(0)
	FL	−1.8426	0.4790	–	−1.6107	0.5241	I(1)	−1.6133	0.3841	I(1)	−1.3232	0.5222	I(1)
	GA	−1.8479	0.6020	–	−1.6332	0.5245	–	−1.3519	0.5039	I(1)	−1.1429	0.3616	I(1)
	KY	−1.5272	0.5433	I(1)	−1.3253	0.2885	–	−2.9204	0.4626	–	−0.9933	0.3480	–
South-East	LA	−1.9348	0.5344	–	−2.1339	0.5862	–	−2.1302	0.4048	–	−1.5886	0.4089	I(1)
	MS	−1.4535	0.5861	I(1)	−2.0241	0.5531	–	−1.7455	0.4616	–	−0.9440	0.4013	I(1)
	NC	−1.3791	0.6126	I(1)	−1.3680	0.5171	I(1)	−1.7276	0.4963	–	−1.4223	0.3150	–
	SC	−2.3515	0.5985	–	−3.1451	0.5175	–	−0.8534	0.4832	I(1)	−1.2683	0.4133	I(1)
	TN	−1.5647	0.6113	I(1)	−1.3805	0.2567	–	−0.5782	0.4880	I(1)	−0.1966	0.2957	–
	VA	−1.3559	0.6145	I(1)	−2.5429	0.1974	I(0)	−1.1074	0.4980	I(1)	−1.7623	0.3037	I(0)
	WV	−3.1804	0.1287	I(0)	−1.9640	0.5821	–	−1.8456	0.1567	I(0)	−1.5191	0.5086	I(1)
	AZ	−2.3491	0.4744	–	−2.0502	0.5115	–	−1.7448	0.3408	I(0)	−2.2468	0.4612	–
South-West	NM	−2.0468	0.2697	I(0)	−2.5325	0.1506	I(0)	−1.4650	0.2883	–	−1.4228	0.3003	–
	OK	−1.7299	0.4047	–	−1.9240	0.4684	–	−1.9284	0.1509	I(0)	−1.9002	0.2667	I(0)
	TX	−2.2754	0.4519	–	−1.0962	0.4948	I(1)	−1.8942	0.2288	I(0)	−0.6561	0.4323	I(1)
	CO	−2.2231	0.2506	I(0)	−0.4739	0.4936	I(1)	−1.5843	0.2212	–	−1.4019	0.5089	I(1)
Rocky Mtns	ID	−1.0192	0.3645	I(1)	−2.6184	0.2084	I(0)	−2.1222	0.4625	–	−2.6182	0.0781	I(0)
	MT	−0.3333	0.4872	I(1)	−0.4968	0.4880	I(1)	−1.2958	0.4861	I(1)	−1.4484	0.5040	I(1)
	UT	−0.9173	0.4852	I(1)	−2.0844	0.1505	I(0)	−1.0553	0.4601	I(1)	−1.6098	0.1246	–
	WY	−1.5813	0.4434	I(1)	−2.3289	0.2051	I(0)	−1.2765	0.3509	I(1)	−2.1698	0.0926	I(0)
	CA	−2.1955	0.5853	–	−2.8001	0.2929	I(0)	−1.5666	0.4970	I(1)	−1.6484	0.1026	I(0)
Far West	NV	−0.9354	0.5978	I(1)	−1.5892	0.5841	I(1)	−1.1097	0.5005	I(1)	−0.9317	0.4892	I(1)
	OR	−1.2037	0.5375	I(1)	−2.5778	0.3793	–	−1.9585	0.4867	–	−2.4376	0.1638	I(0)
	WA	−1.3875	0.5335	I(1)	−1.7204	0.5738	–	−1.7821	0.4490	–	−0.7521	0.5081	I(1)

Note: See the notes for Table 1.

For the remaining states, the two tests yield either evidence against convergence or conflicting outcomes. Because several states reveal no convincing evidence on convergence, there seems no rigorous way to pinpoint which states are converging or diverging based on univariate tests, which is consistent with the general perception that one cannot develop conclusive evidence on stochastic convergence using univariate tests.

We then move on to panel data analysis. Table 3 displays the results of the LLC and IPS tests which are built upon the same null hypothesis that

TABLE 3: Panel Unit-Root Test Results

		LLC-test			IPS test		
		$\hat{\rho}$	τ	p-value	$\hat{\rho}$	$\bar{t}$	p-value
	48 States	0.930	−15.603	0.001	0.938	−2.456	0.001
	New England (6)	0.878	−9.016	0.001	0.874	−3.747	0.001
	Mideast (5)	0.898	−9.379	0.001	0.946	−3.755	0.001
1929–2001	Great Lakes (5)	0.952	−2.727	0.563	1.002	−1.208	0.692
	Plains (7)	0.774	−8.660	0.001	0.805	−3.200	0.001
National	Southeast (12)	0.962	−5.637	0.135	0.985	−1.873	0.110
Average	Southwest (4)	0.900	−4.881	0.021	0.944	−2.368	0.036
	Rocky Mountains (5)	0.920	−3.075	0.368	0.894	−1.708	0.254
	Far West (4)	0.944	−3.685	0.168	1.003	−1.626	0.304
	15 Selected States	0.927	−10.062	0.001	0.959	−2.588	0.001
		[0.995]	[−1.923]	[0.799]	[1.052]	[−0.438]	[0.847]
	New England (6)	0.895	−6.084	0.012	0.925	−2.514	0.010
	Mideast (5)	0.898	−7.008	0.002	0.944	−2.826	0.004
1929–2001	Great Lakes (5)	0.837	−7.292	0.001	0.851	−3.394	0.001
	Plains (7)	0.680	−11.342	0.001	0.620	−4.805	0.001
Regional	Southeast (12)	0.911	−6.988	0.018	0.929	−2.186	0.011
Average	Southwest (4)	0.901	−3.860	0.118	0.972	−1.721	0.222
	Rocky Mountains (5)	0.841	−4.689	0.049	0.735	−2.953	0.002
	Far West (4)	0.834	−5.416	0.018	0.891	−2.526	0.037
	48 States	0.941	−11.533	0.035	0.946	−1.996	0.046
	New England (6)	0.914	−4.399	0.214	0.949	−1.965	0.163
	Mideast (5)	0.920	−4.317	0.142	0.958	−1.924	0.143
1947–2001	Great Lakes (5)	0.952	−3.054	0.333	1.011	−1.435	0.329
	Plains (7)	0.801	−5.751	0.021	0.722	−2.694	0.001
National	Southeast (12)	0.949	−6.815	0.061	0.989	−2.217	0.061
Average	Southwest (4)	0.893	−4.279	0.081	0.964	−2.023	0.099
	Rocky Mountains (5)	0.916	−4.596	0.061	0.951	−2.196	0.052
	Far West (4)	0.975	−1.826	0.484	1.029	−1.141	0.419
	15 Selected States	0.950	−5.351	0.176	0.985	−1.530	0.217
		[0.998]	[−1.498]	[0.953]	[1.074]	[−0.347]	[0.984]
	New England (6)	0.902	−4.683	0.187	0.932	−2.075	0.141
	Mideast (5)	0.930	−3.795	0.187	0.957	−1.803	0.161

		LLC-test			IPS test		
		$\hat{\rho}$	τ	p-value	$\hat{\rho}$	$\bar{t}$	p-value
1947–2001	Great Lakes (5)	0.823	−4.680	0.113	0.811	−2.324	0.078
	Plains (7)	0.723	−6.530	0.002	0.713	−2.652	0.002
Regional	Southeast (12)	0.900	−6.399	0.025	0.962	−1.835	0.070
Average	Southwest (4)	0.866	−5.099	0.069	0.921	−2.560	0.053
	Rocky Mountains (5)	0.920	−3.290	0.240	0.906	−1.995	0.086
	Far West (4)	0.884	−3.714	0.137	0.937	−1.970	0.130

Note: Numbers in the parentheses represent the number of states belonging to each BEA region. All entries are based on national price index as deflator except for the numbers in the brackets which represent the results from CPIs as deflator. ρ represents the persistence parameter of the speed of convergence. In the LLC test, ρ is restricted to be equal across individual series, hence, simple estimated value is reported. In the IPS test, since ρ_i differs across i, bias adjusted estimates of ρ are presented using the formula recommended by Cecchetti et al. (2002). τ and $\bar{t}$ denote test statistics for LLC and IPS, respectively. See the Appendix for further details.

all series in the panel are unit-root processes. Hence nonrejection of the null is, in general, interpreted as evidence against convergence. The two tests produce concurrent outcomes in almost all cases considered. During the entire sample period, they jointly fail to reject the unit-root null for four regions using the national average, but they do so for only one region, Southwest, when regional averages are used, implying that convergence may hold better at the regional level than at the national level. Although this finding is not repeated in the subsample analysis, 1947–2001, as the convergence evidence does not explicitly improve with regional averages over national average, it is still interesting to see that the convergence speed measured by $\hat{\rho}$ is much faster at the regional level.

Table 4 summarizes the results of confirmatory analysis in panel data by pairing the PG test with the IPS test.[13] For all 48 states, we fail to obtain confirmatory inference about convergence as the two tests solidly reject their respective null hypotheses. According to Choi (2002), however, this joint rejection may be indicative of a possible mix of stationary series and unit-root series in the panel, which is confirmed shortly by our analysis of regional economies. At the regional level, the two tests reach agreements in four (full sample) to five (subsample) regions, but the evidence of convergence is still mixed: evidence of convergence in some regions, but evidence of divergence in others.

Overall, it is interesting to witness stronger evidence of stochastic convergence using regional averages rather than a national average. To highlight, during the full sample period, the number of regions with confirmatory evidence on convergence is merely one (Plains) when the national average is used, while it rises to three (Plains, Rocky Mountains, and Far West) when

[13]Here we focus on the combination of the IPS test and the PG test because the LLC and IPS tests produce roughly comparable results in our study.

TABLE 4: Confirmatory Analysis Using Panel Tests (Personal Income)

		National			Regional		
		PG	IPS	C.A.	PG	IPS	C.A.
1929–2001	48 States	0.000	0.001	–			
	New England (6)	0.000	0.001	–	0.000	0.010	–
	Mideast (5)	0.000	0.001	–	0.000	0.004	–
	Great Lakes (5)	0.000	0.692	I(1)	0.000	0.001	–
	Plains (7)	0.271	0.001	I(0)	0.407	0.001	I(0)
	Southeast (12)	0.000	0.110	I(1)	0.000	0.011	–
	Southwest (4)	0.000	0.036	–	0.000	0.222	I(1)
	Rocky Mtns (5)	0.253	0.254	–	0.797	0.002	I(0)
	Far West (4)	0.000	0.304	I(1)	0.166	0.037	I(0)
	15 Selected States	0.000	0.847	I(1)			
1947–2001	48 States	0.000	0.046	–			
	New England (6)	0.284	0.163	–	0.252	0.141	–
	Mideast (5)	0.126	0.147	–	0.021	0.161	I(1)
	Great Lakes (5)	0.000	0.329	I(1)	0.542	0.078	I(0)
	Plains (7)	0.992	0.001	I(0)	0.972	0.002	I(0)
	Southeast (12)	0.000	0.061	–	0.001	0.070	–
	Southwest (4)	0.809	0.099	I(0)	0.437	0.053	I(0)
	Rocky Mtns (5)	0.000	0.052	–	0.292	0.086	I(0)
	Far West (4)	0.000	0.419	I(1)	0.231	0.130	–
	15 Selected States	0.000	0.984	I(1)			

Note: C.A. denotes inference based on confirmatory analysis. Inferences are made at the significance level of 10%. Numbers in the parentheses represent the number of states belong to each BEA region.

using regional averages. At the same time, the number of regions with confirmatory evidence on divergence is reduced from three to one. This pattern is more evident in the subsample analysis where the number of regions with convergence evidence increases from two (Plains and Southwest) to four (Great Lakes, Plains, Southwest, Rocky Mountains) during the period 1947–2001. Another observation worth noting from Table 4 is that the states in the *Plains* region consistently exhibit convergence toward both regional and national averages regardless of the sample period. This may be partly because the per capita income in that region is about as volatile as the national average (Carlino and Sill, 2001) and partly because the states in the region are typically agricultural states in which fluctuations of output are often less persistent than those of manufacturing output (Sorensen and Yosha, 2000).

To summarize, confirmatory analysis in panel data strongly suggests that convergence does not hold in all 48 states. Instead, some evidence of convergence can be obtained at the regional level. Moreover, evidence of convergence marginally improves during the period 1947–2001.

Convergence Clubs?

The main finding in the previous section shares a similar spirit with the concept of convergence clubs, taken as a stylized fact in international data.[14] Some recent studies in the convergence literature suggest a possible existence of convergence clubs in the United States. For example, Johnson and Takeyama (2000) claim that the economic development of the U.S. states since 1950 can be best characterized by convergence clubs. Based on states' initial conditions they find three convergence clubs that contain certain geographical characteristics.

The natural question then arises as to why regional convergence holds in some regions and not in others. An immediate but tentative answer is that the degree of homogeneity among states varies across regions. If states in a region are more homogeneous in terms of size or industrial structures, their relative output will be less volatile and thus more likely to converge. Another explanation could be that since the classification of regions, regional specifications, by the BEA is based on simple geographical location without specific economic consideration, states in the same region do not necessarily share common characteristics of per capita income. Indeed, there is a good reason to believe that two neighboring states in different regions like Connecticut (New England) and New York (Mideast) are more likely to share similar social and economic conditions than two states in the same region but farther apart, such as Connecticut and Vermont.

Robustness to the Choice of Output Measures

Personal income data are often criticized for improperly reflecting the income defined in the neoclassical growth model. For instance, Carlino and Mills (1996) contend that the neoclassical growth model relies on factor migration leading to convergence of wages and earnings but not necessarily per capita income, which includes transfer payments that may reinforce or counter any convergence trend in per capita earnings. By comparing the time-series properties of both personal income and earnings in major regions of the United States, they conclude that the evidence of stochastic convergence is stronger with the per capita income series than the per capita earnings version.

Our panel techniques confirm their finding. As reported in Table 5, confirmatory analysis using personal earnings data provide weaker evidence of convergence in all cases considered although the evidence is still marginally stronger at the regional level relative to the national level.

[14]Baumol (1986) found no explicit convergence pattern among 72 countries considered, but convergence was detected in certain country groups, such as industrialized countries, centrally planned economies, and middle-income market economies. He termed these country groups "convergence clubs."

TABLE 5: Confirmatory Analysis Using Panel Tests (Personal Earnings)

		National			Regional		
		PG	IPS	C.A.	PG	IPS	C.A.
1929–2001	48 States	0.000	0.013	–			
	New England (6)	0.000	0.018	–	0.000	0.242	I(1)
	Mideast (5)	0.000	0.010	–	0.000	0.016	–
	Great Lakes (5)	0.000	0.414	I(1)	0.060	0.001	–
	Plains (7)	0.826	0.001	I(0)	0.763	0.001	I(0)
	Southeast (12)	0.000	0.164	I(1)	0.000	0.035	–
	Southwest (4)	0.000	0.137	I(1)	0.000	0.602	I(1)
	Rocky Mtns (5)	0.000	0.737	I(1)	0.000	0.090	–
	Far West (4)	0.000	0.251	I(1)	0.601	0.051	I(0)
1947–2001	48 States	0.000	0.044	–			
	New England (6)	0.746	0.322	–	0.271	0.388	–
	Mideast (5)	0.756	0.166	–	0.124	0.181	–
	Great Lakes (5)	0.000	0.249	I(1)	0.801	0.063	I(0)
	Plains (7)	0.971	0.001	I(0)	0.948	0.001	I(0)
	Southeast (12)	0.000	0.050	–	0.000	0.102	I(1)
	Southwest (4)	0.776	0.192	–	0.000	0.153	I(1)
	Rocky Mtns (5)	0.000	0.256	I(1)	0.000	0.457	I(1)
	Far West (4)	0.000	0.252	I(1)	0.416	0.173	–

Note: See notes in Table 4.

Using the Metropolitan CPIs as Deflator

The results so far are based on the state per capita output deflated by the common national price index that ignores persistence of relative price across states. Here we take into account the persistence of relative price in the analysis by exercising the panel techniques to real per capita incomes of the 15 selected states deflated by relevant metropolitan area CPIs. Because 15 states are not representative enough to derive a reliable inference on regional convergence, only their convergence results vis-à-vis the national average are reported here. As shown in Table 4, confirmatory analysis provides no evidence of convergence among the 15 states for both full and subsamples. Instead, there is strong evidence of divergence among them, signifying that the real per capita incomes in the 15 states are deviating over time from the national average level. A tentative explanation for this finding would be that the divergence might have been driven by high persistence of relative state prices embedded in the metropolitan CPIs. This argument is readily supported by the convergence speeds reported in Table 3. The estimates of the persistence parameter for the 15 states are far greater using metropolitan CPIs as deflators than using the common national deflator. Specifically, the parameter estimates from the LLC test for the 15 states are 0.995 (full sample) and 0.998 (subsample) when metropolitan area CPIs are used, compared to 0.927 and

0.950 based on the common national deflator. On one hand, this result implies that using a national deflator may underestimate the dynamic properties of real per capita incomes by ignoring dynamics rooted in the relative prices between states. For this reason, special care should be taken in interpreting the empirical results of previous studies based on a common national deflator. However, in view of the substantial spatial variation in prices within states, metropolitan area CPIs could exaggerate the state-wide price level. Therefore, it is safe to argue that the convergence evidence based on the true state price index may fall in somewhere between these two extreme cases. Comparing the two results appears to be a step in the right direction in this regard.

5. EXPLANATIONS FOR REGIONAL CONVERGENCE

Our analysis in the previous section delivers two main messages. First, despite economic integration and high factor mobility within a nation, differences in regional characteristics could lead to different convergence paths across regions in the nation. Second, the differential across regions may widen in the long run while it may narrow within regions. A question then naturally emerges about what factors might be accountable for the differences in regional convergence characteristics and to what extent they matter for output convergence. The empirical literature on regional growth presents a rich menu of potential explanations for regional differences in income growth pattern. A partial list includes: (1) differing industrial structures (Browne, 1978; Shea, 1996; Carlino and Sill, 2001); (2) spatial components in income convergence process (Rey and Montouri, 1999; Magrini, 2003); (3) geographical variations in policy effects (Mehay and Solnick, 1990; Hooker and Knetter, 1997; Carlino and DeFina, 1998); and (4) different risk-sharing patterns (Sorensen and Yosha, 2000).[15] Among them we study the first two factors as potential sources of our empirical finding.

Heterogeneities in Industrial Structure

Heterogeneities in industry mix arising from regional specializations are often blamed for regional income disparities because they are believed to impede income convergence across regions that might have achieved the convergence in factor returns through economic integration and trade. Based on the dataset covering approximately 150 years, Kim (1998) finds that regional industrial structures played a crucial role in the divergence and convergence of U.S. regional income per capita, albeit differences in industry mix are not solely responsible for all the variations in regional per capita income.

[15]Among other factors, Browne (1989), Carlino (1992), and Esteban (2000) attribute the interregional inequality in per capita income for the United States to the regional variability in labor market conditions such as unemployment or participation rates.

To gain further insights, we examine the role of industrial mix in the per capita output convergence by positing that dissimilarity in industrial structures between states leads to more volatile relative per capita income. Following Kim (1998), we quantify the differences in the industrial structure of states using the index of state specialization such that

$$S_{ij} = \sum_{k=1}^{n} \left| \frac{E_{ki}}{E_i} - \frac{E_{kj}}{E_j} \right|$$

where E_{ki} denotes the employment level of industry k in state i and E_i represents the total industrial employment for state i. The indexes are constructed based on one-digit sectoral employment data during 1969–2000 for agriculture, mining, construction, manufacturing, transportation, wholesale trade, retail trade, finance, services, and government.[16] The index ranges between zero, when two states have the identical industrial structures, and two, if two engaged states possess completely different industrial structures. Table 6 reports the average specialization indexes for each state with regard to other states within the same region as well as outside the region. In most cases considered, the average indexes within regions are smaller than those outside regions, confirming our original intuition that states within a region share more homogeneous industrial structure than outside the region.

We investigate the relationship between industrial structures and the volatility of relative output by regressing the income volatility onto the state specialization index as

$$V_{ij} = \gamma + \delta \tilde{S}_{ij} + \varepsilon_i \tag{4}$$

where V_{ij} is the standard deviation of income disparity between states i and j at time t such that $V_{ij} = \left[\frac{1}{T}\sum_{t=1}^{T}(d_{ijt} - \bar{d}_{ij})^2\right]^{1/2}$, $\bar{d}_{ij} = \frac{1}{T}\sum_{t=1}^{T} d_{ijt}$, $d_{ijt} = y_{jt} - y_{it}$, and $\tilde{S}_{ij} = \frac{1}{T}\sum_{t=1}^{T} S_{ij}$. As presented in Table 6, in the vast majority of states the point estimates of the coefficient δ are positive and statistically significantly different from zero at the 10 percent level, indicating that two states with more homogeneous industrial structures are likely to experience less volatile patterns of relative per capita output.

The Role of Spatial Effects

If markets for goods and production factors are geographically concentrated, intrinsic homogeneity among economies could be positively related to spatial proximity. Given that transportation costs induce differences in relative prices, geographically neighboring regions with lower transportation costs will be more likely to adjust quickly to certain relative price and output disturbances than regions that are far apart. Moreover, economies located

[16]The data are retrieved from the State Personal Income CD-ROM.

TABLE 6: State Specialization Index And Income Volatility

		Specialization Index		Income Volatility and Specialization ($V_i = \gamma + \delta\tilde{S}_i + \varepsilon_i$)			
Region	States	Inside Region	Outside Region	$\hat{\delta}$	s.e.	t-ratio	R^2
New England	CT	0.155	0.249	10.14	2.05	4.96*	0.35
	ME	0.167	0.178	2.38	3.14	0.76	0.01
	MA	0.160	0.243	14.05	2.01	7.00*	0.52
	NH	0.139	0.228	6.45	1.67	3.87*	0.25
	RI	0.145	0.236	10.72	2.79	3.85*	0.25
	VT	0.161	0.196	8.19	2.42	3.38*	0.20
Mid-East	DE	0.201	0.204	6.48	4.06	1.60	0.05
	MD	0.241	0.232	7.55	3.09	2.45*	0.12
	NJ	0.170	0.226	10.81	2.24	4.83*	0.34
	NY	0.213	0.263	17.46	3.44	5.07*	0.36
	PA	0.181	0.204	7.82	2.00	3.91*	0.25
Great Lakes	IL	0.140	0.203	7.42	2.29	3.24*	0.19
	IN	0.107	0.231	−0.12	1.71	−0.07	0.00
	MI	0.087	0.221	3.35	1.92	1.75*	0.06
	OH	0.081	0.211	4.57	2.01	2.27*	0.10
	WI	0.106	0.212	0.83	1.82	0.46	0.00
Plains	IA	0.148	0.204	4.19	1.69	2.49*	0.12
	KS	0.156	0.192	8.26	1.57	5.25*	0.38
	MN	0.178	0.173	3.54	1.46	2.43*	0.12
	MO	0.176	0.169	4.22	1.79	2.35*	0.11
	NE	0.142	0.210	6.85	1.45	4.74*	0.33
	ND	0.232	0.303	12.42	2.29	5.42*	0.40
	SD	0.182	0.261	10.53	2.12	4.97*	0.35
South-East	AL	0.160	0.184	12.29	2.79	4.41*	0.30
	AR	0.186	0.190	10.68	2.88	3.71*	0.23
	FL	0.318	0.187	4.24	1.87	2.27*	0.10
	GA	0.176	0.158	6.24	3.84	1.62	0.06
	KY	0.188	0.179	9.74	2.35	4.15*	0.28
	LA	0.244	0.184	9.18	1.95	4.72*	0.33
	MS	0.197	0.220	15.84	2.75	5.76*	0.42
	NC	0.201	0.211	6.92	2.63	2.63*	0.13
	SC	0.205	0.222	9.21	3.00	3.07*	0.17
	TN	0.180	0.163	2.98	2.43	1.23	0.03
	VA	0.222	0.180	6.51	3.15	2.07*	0.09
	WV	0.224	0.197	5.36	2.53	2.11*	0.09
South-West	AZ	0.188	0.219	6.58	2.00	3.29*	0.19
	NM	0.212	0.291	8.05	2.22	3.62*	0.23
	OK	0.185	0.227	11.29	1.93	5.84*	0.43
	TX	0.165	0.196	8.24	2.00	4.12*	0.27
Rocky Mtns	CO	0.205	0.208	5.55	1.60	3.47*	0.21
	ID	0.211	0.211	1.26	1.94	0.65	0.01

Region	States	Specialization Index: Inside Region	Specialization Index: Outside Region	Income Volatility and Specialization ($V_i = \gamma + \delta\tilde{S}_i + \varepsilon_i$): $\hat{\delta}$	s.e.	t-ratio	R^2
	MT	0.194	0.256	1.25	2.86	0.44	0.00
	UT	0.190	0.194	5.02	2.48	2.02*	0.08
	WY	0.298	0.366	−1.30	3.31	−0.39	0.00
Far West	CA	0.193	0.195	11.00	3.46	3.18*	0.18
	NV	0.398	0.411	24.95	7.18	3.48*	0.21
	OR	0.209	0.173	4.12	2.43	1.69*	0.06
	WA	0.204	0.175	3.70	2.78	1.33	0.04

Note: Specialization index represents the average of $S_{ij} = \sum_{k=1}^{n} |\frac{E_{ki}}{E_i} - \frac{E_{kj}}{E_j}|$, where E_{ki} denotes the employment of industry k in state i. Data on one-digit sectoral employment are used for the period between 1969 and 2000, which include agriculture, mining, construction, manufacturing, transportation, wholesale trade, retail trade, finance, services, and government. s.e. represents standard errors. T-values are calculated under the null hypothesis of no relationship between state specialization index (S_{ij}) and per capita income volatility, or H_0: $\delta = 0$ in $V_j = \gamma + \delta S_j + \varepsilon_i$. Statistical significance is indicated by an asterisk (*) for the 10% level.

close to each other are more subject to common random shocks than those geographically far apart. According to Audretsch and Feldman (1996), the ability to receive knowledge and technology spillovers is influenced by the distance from the knowledge source and, thus, geographic concentration perhaps plays an essential role in the spillovers. Rey and Montouri (1999) find strong evidence of spatial clusters of states that are homogeneous in terms of output convergence process.

In this context, it is instructive to consider the role of geographic space in which economic relationships take place. We explore the relationship between the volatility of income disparities and geographic distance, conceived as a proxy for transportation costs. A positive relationship between the two variables would nuance that geographic proximity exerts a positive impact on the convergence process of per capita output. The following regression equation is estimated under a *prior* that the two variables of interest are positively related such that

$$V_i = \alpha + \beta D_i + \varepsilon_i \quad i = 1, 2, \ldots, N$$

where V_i denotes the standard deviation of income disparity as in Equation (4) and D_i represents the logarithm of spatial distance of state i relative to the geographic center.[17] Table 7 reports the regression results for the entire 48 states based on the national price index. In all cases considered, the point

[17]Five states are randomly chosen as geographic centers and the distances between state capitals are used as a proxy for the distances between states. For example, the physical distance between Ohio and New Hampshire is measured by the distance between Columbus and Concord.

TABLE 7: Geographic Distance and Income Volatility

48 states				Geographic Center	15 states			
$\hat{\beta}$	s.e	t-ratio	R^2		$\hat{\beta}$	s.e	t-ratio	R^2
7.1155	1.3698	5.1947*	0.3749	Georgia	−1.0373	2.3358	−0.4441	0.0162
4.7251	1.0003	4.7239*	0.3315	Kansas	−1.6918	1.8517	−0.9136	0.0650
3.6124	0.9088	3.9749*	0.2599	Massachusetts	1.0773	2.4285	0.4436	0.0161
6.9302	1.2706	5.4544*	0.3980	Texas	3.2553	4.3317	0.7515	0.0449
3.3419	1.4230	2.3485*	0.1092	Washington	−0.2608	1.5943	−0.1636	0.0022

Note: Per capita incomes of 48 states are deflated by the national price index while those of 15 selected states are deflated by metropolitan area CPIs. S.e. represents standard errors. T-values are calculated under the null of no relationship between geographic distance and per capita income volatility. That is, $\beta = 0$, where $y_{it} - y_t = \hat{\alpha} + \hat{\beta} D_i$. Statistical significance is indicated by an asterisk (*) for the one percent level.

estimates of β are statistically significantly different from zero at the one percent level within the range of 0.91 and 1.23, suggesting that an increase in the physical distance by 2.718 times [ln(2.718) = 1] is likely to make the relative output more volatile by about 0.91 to 1.23 percentage points per annum. As a consequence, two states far apart are liable to exhibit more volatile relative per capita income and less likely to converge over time. This stable relationship is not observed for the 15 states based on metropolitan area CPIs in which some point estimates for β have unexpected signs although the coefficients are not statistically significant.

6. SUMMARY AND CONCLUDING REMARKS

Economic theory suggests that economies with integrated markets for products and production factors, and the same institutional environments, should exhibit convergence in real per capita output. The states in the U.S. are believed to satisfy these underlying conditions of convergence better than other observational units. The current study has reexamined the issue of output convergence among the 48 continental U.S. states by implementing multiple panel data techniques to state per capita output in the framework of confirmatory analysis. We find little evidence of stochastic convergence at the national level, but some evidence at the regional level during the period 1929–2001. The evidence of regional convergence slightly improves during the post-World War II period. However, the intensity of the regional convergence pattern appears to vary with the choice of output measures and deflators. Use of personal earnings as an output measure and use of metropolitan area CPIs as a deflator tend to weaken the evidence. Overall, despite the nearly identical institutional environments and highly integrated product and factor markets, there is little evidence of convergence among the 48 contiguous states. Instead, convergence is likely to hold among geographically adjacent states that share certain common regional features such as climate and industrial structures.

These findings apparently conflict with the central tenet of the neoclassical growth model as well as with the common perception based on earlier empirical evidence. A search of potential accounts for our empirical results shows that industrial structures and geographic distance play important roles in the process of output convergence. Indeed, other noteworthy factors contributing to a regional convergence patterns share certain spatial properties.

REFERENCES

Andrews, Donald W. K. and J. Christopher Monahan. 1992. "An Improved Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimator," *Econometrica*, 60, 953–966.

Audretsch, David B. and Maryann P. Feldman. 1996. "R&D Spillovers and the Geography of Innovation and Production," *American Economic Review*, 86, 630–640.

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, 407–443.

Barro, Robert J. and Xavier Sala-i-Martin. 1992. "Convergence," *Journal of Political Economy*, 100, 223–251.

Baumol, William J. 1986. "Productivity Growth, Convergence, and Welfare: What the Long-Run Data Show," *American Economic Review*, 76, 1072–1085.

Berkowitz, Jeremy and Lutz Kilian. 2000. "Recent Development in Bootstrapping Time Series," *Econometric Reviews*, 19, 1–48.

Bernard, Andrew B. and Steven N. Durlauf. 1995. "Convergence in International Output," *Journal of Applied Econometrics*, 10, 97–108.

Browne, Lynn E. 1978. "Regional Industry Mix and the Business Cycle," *New England Economic Review*, November/December, 35–53.

———. 1989. "Shifting Regional Fortunes: The Wheel Turns," *New England Economic Review*, Federal Reserve Bank of Boston, May/June, 27–40.

Carlino, Gerald A. 1992. "Are Regional Per Capita Earnings Diverging?," *Business Review*, Federal Reserve Bank of Philadelphia, 3–12.

Carlino, Gerald A. and Robert F. DeFina. 1998. "The Differential Regional Effects of Monetary Policy," *The Review of Economics and Statistics*, 80, 572–587.

Carlino, Gerald A. and Leonard Mills. 1993. "Are U.S. Regional Incomes Converging? A Time Series Analysis," *Journal of Monetary Economics*, 32, 335–346.

———. 1996. "Convergence and the U.S. States: A Time-Series Analysis," *Journal of Regional Science*, 36, 597–616.

Carlino, Gerald A. and Keith Sill. 2001. "Regional Income Fluctuations: Common Trends and Common Cycles?" *The Review of Economics and Statistics*, 83, 446–456.

Cecchetti, Stephen G., Nelson C. Mark, and Robert Sonora. 2002. "Price Index Convergence among United States Cities," *International Economic Review*, 43, 1081–1099.

Chang, Yoosoon. 2002. "Nonlinear IV Unit Root Tests in Panels with Cross-Sectional Dependency," *Journal of Econometrics*, 110, 261–292.

Choi, Chi-Young. 2002. "A Variable Addition Panel Test for Stationarity and Confirmatory Analysis," mimeo Department of Economics, University of New Hampshire.

Elliott, Graham, Thomas J. Rothenberg, and James H. Stock. 1996. "Efficient Tests for an Autoregressive Unit Root," *Econometrica*, 64, 813–836.

Esteban, Joan. 2000. "Regional Convergence in Europe and the Industry Mix: A Shift-Share Analysis," *Regional Science and Urban Economics*, 30, 353–364.

Evans, Paul and Georgios Karras. 1996. "Convergence Revisited," *Journal of Monetary Economics*, 37, 249–265.

Hall, Alastair. 1994. "Testing for A Unit Root in Time Series with Pretest Data Based Model Selection," *Journal of Business and Economic Statistics*, 12, 461–470.

Hooker, Mark and Michael M. Knetter. 1997. "The Effects of Military Spending on Economic Activity: Evidence from State Procurement Spending," *Journal of Money, Credit, and Banking*, 29, 400–421.

Im, Kyung So, M. Hashem Pesaran, and Yongcheol Shin. 2003. "Testing for Unit Roots in Heterogeneous Panels," *Journal of Econometrics*, 115, 53–74.

Johnson, Paul A. and Lisa N. Takeyama. 2000. "Absolute, Conditional, or Club Convergence in the U.S. States?" mimeo Department of Economics, Vassar College.

Kilian, Lutz and Tao Zha. 2002. "Quantifying the Uncertainty about the Half-Life of Deviations from PPP," *Journal of Applied Econometrics*, 17, 107–125.

Kim, Sukkoo. 1998. "Economic Integration and Convergence: U.S. Regions, 1840–1987," *Journal of Economic History*, 58, 659–683.

Kwiatkowski, Denis, Peter C.B. Phillips, Peter Schmidt, and Yongcheol Shin. 1992. "Testing the Null Hypothesis of Stationary against the Alternative of a Unit Root," *Journal of Econometrics*, 54, 159–178.

Levin, Andrew, Chien-Fu Lin, and Chia-Shang James Chu. 2002. "Unit Root Tests in Panel Data: Asymptotic and Finite-Sample Properties," *Journal of Econometrics*, 108, 1–24.

Lowey, Michael B. and David H. Papell. 1996. "Are U.S. Regional Incomes Converging? Some Further Evidence," *Journal of Monetary Economics*, 38, 587–598.

Maddala, G.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.

Magrini, Stefano. 2003. "Regional (Di)Convergence," mimeo University Ca Foscari of Venice.

Mark, Nelson C. 2001. *International Macroeconomics and Finance: Theory and Empirical Methods*. Malden, MA: Blackwell Publishers.

Mark, Nelson C. and Donggyu Sul. 2001. "Nominal Exchange Rates and Monetary Fundamentals: Evidence from a Small Post-Bretton Woods Panel," *Journal of International Economics*, 53, 29–52.

Mehay, Stephen and Loren Solnick. 1990. "Defense Spending and State Economic Growth," *Journal of Regional Science*, 30, 477–487.

Mitchener, Kris J. and Ian W. McLean. 1999. "U.S. Regional Growth And Convergence, 1880–1980," *Journal of Economic History*, 59, 1016–1042.

Newey, Whitney K. and Kenneth D. West. 1987. "A Simple, Positive, Semidefinite Heteroskedasticity and Autocorrelation Consistent Covariance Matrix," *Econometrica*, 55, 703–708.

Ng, Serena and Pierre Perron. 2001. "Lag Length Selection and the Construction of Unit Root Tests with Good Size and Power," *Econometrica*, 69, 1519–1554.

O'Connell, Paul G. J. 1998. "The Overvaluation of Purchasing Power Parity," *Journal of International Economics*, 44, 1–19.

Park, Joon Y. 1990. "Testing for Unit Roots and Cointegration by Variable Addition," *Advances in Econometrics*, 8, 107–133.

Park, Joon Y. and Buhmsoo Choi. 1988. "A New Approach to Testing for a Unit Root," Working Paper No. 88–23, Department of Economics, Cornell University.

Phillips, Peter C. B. and Donggyu Sul. 2003. "Dynamic Panel Estimation and Homogeneity Testing under Cross-Section Dependence," *The Econometrics Journal*, 6, 217–259.

Rey, Sergio J. and Brett D. Montouri. 1999. "U.S. Regional Income Convergence: A Spatial Econometric Perspective," *Regional Studies*, 33, 143–156.

Shea, John. 1996. "Comovement in Cities," *Carnegie-Rochester Conference Series on Public Policy*, 44, 169–206.

Sorensen, Bent E. and Oved Yosha. 2000. "Is Risk Sharing in the United States a Regional Phenomenon?" Economic Review, Federal Reserve Bank of Kansas City, Second Quarter, 33–47.

State Personal Income (SPI) CD-ROM. 1929–2001. U.S. Department of Commerce, Washington D.C., 2002.

Tsionas, Efthymios G. 2001. "Regional Convergence and Common, Stochastic Long-run Trends: A Re-examination of the U.S. Regional Data," *Regional Studies*, 35, 689–696.

APPENDIX

A.1 REGRESSION EQUATION OF STOCHASTIC CONVERGENCE IN PANEL DATA

Averaging Equation (2) over the N economies yields

$$\lim_{k\to\infty} E_t(\bar{y}_{t+k} - a_{t+k}) = \frac{1}{N}\sum_{i=1}^{N}\mu_i$$

where $\bar{y}_t = \frac{1}{N}\sum_{i=1}^{N} y_{i,t}$. To remove a_t which is unobservable, subtract each member of the above equation from the corresponding member of Equation (2),

$$\lim_{k\to\infty} E_t(y_{i,t+k} - \bar{y}_{t+k}) = \mu_i - \frac{1}{N}\sum_{i=1}^{N}\mu_i$$

where economies $i = 1,2,\ldots,N$ are said to converge if and only if $y_{i,t} - \bar{y}_t$ is stationary for each i.

A typical regression to test conditional convergence is $\frac{1}{T}\log\left(y_{i,T}/y_{i,0}\right) = \alpha + \beta\log(y_{i,0}) + \gamma X_i + \varepsilon_i$, $i = 1,2,\ldots,N$, where X_i is a vector of observations on exogenous variables designed to control for the cross economy heterogeneity in levels and growth rates of per capita income, α and β are parameters, γ is a parameter vector, and ε is an error term with a zero mean and finite variance. The traditional conditional convergence equation can be reformulated to incorporate the feature of conditional convergence

$$y_{i,t} - \bar{y}_t = \delta_i + \lambda(y_{i,t-1} - \bar{y}_{t-1}) + \epsilon_{i,t}$$

where $\delta_i = [(\lambda - 1)\gamma/\beta]'X_i$, $\lambda \equiv (1 + \beta T)^{1/T}$, and $\varepsilon_{i,t}$ may be serially correlated with a zero mean and finite and constant variance. Then it can be transformed into

$$\Delta(y_{i,t} - \bar{y}_t) = \delta_i + \rho_i(y_{i,t-1} - \bar{y}_{t-1}) + \sum_{k=1}^{p}\phi_{i,k}\Delta(y_{i,t-k} - \bar{y}_{t-k}) + u_{i,t}$$

which is Equation (3) in the text.

A.2 THREE PANEL TESTS

(1) *The Levin, Lin, and Chu (LLC) test*

LLC (2002) propose several panel unit-root tests based on the following Augmented Dickey Fuller (ADF) model. Notice that a linear trend term is omitted as it is inconsistent with the convergence hypothesis we examine

$$\Delta y_{it} = \alpha_i + \beta y_{i,t-1} + \sum_{j=1}^{k_i} \phi_{ij} \Delta y_{i,t-j} + \varepsilon_{it}, \quad i = 1, 2, \ldots, N; t = 1, 2, \ldots T$$

In the paper, the lag length k_i in the ADF regression equation is chosen by Hall's (1994) general-to-specific method based on the recursive t-test. Under the null hypothesis that all series in the panel are unit-root (H_0: $\beta_1 = \ldots = \beta_N = \beta = 0$) against the alternative that all series are stationary (H_A: $\beta_1 = \ldots = \beta_N = \beta < 0$), the adjusted t-statistic (τ^*) obtained from pooled regression packages has a limiting distribution of standard normal

$$\tau^* \xrightarrow{d} N(0, 1)$$

Beware that a homogeneity restriction is imposed on the implicit alternative hypothesis such that all series, rather than at least one of them, are stationary. Despite this restriction, rejection of the null hypothesis can occur when a small number of stationary series are present in the panel (See Mark, 2001, page 44). For this reason, rejection of the null should be interpreted as at least one stationary series exists in the panel.

(2) *The Im, Pesaran, and Shin (IPS) test*

IPS (2003) develop a group mean panel unit-root test that allows for heterogeneities in intercept and serial correlation as well as convergence speed across individual series. In the following regression equation

$$\Delta y_{it} = \alpha_i + \rho_i y_{i,t-1} + \sum_{j=1}^{k_i} \phi_{i,j} \Delta y_{i,t-j} + \epsilon_{it} \quad i = 1, 2, \ldots, N; t = 1, 2, \ldots T$$

The univariate ADF test is applied to each individual series to construct the sample mean of the resulting t-statistics (τ_i), $\bar{\tau}_N = \frac{1}{N} \sum_{i=1}^{N} \tau_i$. Under the null hypothesis that all series in the panel are unit-root (H_0: $\rho_i = 0$ for all i) against the alternative that at least one of them is stationary (H_A: $\rho_i < 0$ for at least one i), the IPS test statistic has a standard normal limiting distribution,

$$\frac{\sqrt{N}(\bar{\tau}_N - E(\bar{\tau}_N))}{\sqrt{Var(\bar{\tau}_N)}} \xrightarrow{d} N(0, 1)$$

(3) *The Panel G-test*

The panel G-test (PG) is proposed by Choi (2002) as a panel extension of the univariate G-test originally developed by Park and Choi (1988) and Park (1990). As a variable addition test, it is based on the regression of a given time series onto time polynomials including superfluous time polynomial terms. Specifically, a time series is regressed on a time polynomial with order dictated by the null hypothesis and then some superfluous higher-order time polynomial terms are added. By testing the significance of these superfluous time polynomial terms with the standard test, such as the Wald test, we attempt to discern whether the series is stationary (around a deterministic trend) or unit-root. The superfluous regressors will be insignificant if the time series is stationary, whereas they will be significant if the series contains a unit-root component. Because the test allows for deterministic trend polynomials of arbitrary order, it is known to deal with a broad class of data generating processes exhibiting serial correlation and heteroskedasticity.

To test whether all series in a panel $\{y_{it}\}_{t=1,\ldots,T}^{i=1,\ldots,N}$ are level stationary against the alternative that at least one of them is unit-root, the panel G-statistic is articulated as

$$G^*(0,2) = \frac{NT(\hat{\sigma}^2 - \tilde{\sigma}^2)}{\frac{1}{N}\sum_{i=1}^{N}\hat{\omega}_i^2} \quad \xrightarrow{d} \quad \chi^2_{2N}$$

where

$$\hat{\sigma}^2 = \frac{1}{N}\frac{1}{T}\sum_{i=1}^{N}\sum_{t=1}^{T}\hat{e}_{it}^2, \quad \hat{e}_{it} = y_{it} - \hat{\alpha}_{0i}$$

$$\tilde{\sigma}^2 = \frac{1}{N}\frac{1}{T}\sum_{i=1}^{N}\sum_{t=1}^{T}\tilde{e}_{it}^2, \quad \tilde{e}_{it} = y_{it} - \tilde{\alpha}_{0i} - \tilde{\alpha}_{1i}t - \tilde{\alpha}_{2i}t^2$$

$\hat{\omega}_i^2$ is the long run variance of $\hat{e}_{it}$, and α_1 and α_2 are the coefficients for the superfluous time polynomial terms. For the long run variance estimation, Newey and West's (1987) Bartlett kernel is used with the fixed bandwidth of $l = \text{integer}[5(T/100)^{1/4}]$ (for annual data).

The null hypothesis will be rejected in favor of the alternative if the test statistic exceeds certain critical values. In this paper inferences are drawn from p-values based on a nonparametric bootstrap method described below instead of the asymptotic distribution in order to control for cross sectional dependence.

A.3 NONPARAMETRIC BOOTSTRAP PROCEDURES

The residual-based nonparametric bootstrap method is practiced as follows.

First, by using the iterative seemingly unrelated regression (SUR) method, fit the following equation to get an estimator of the parameters $\hat{\alpha}_i$, $\hat{\rho}_i$, and $\hat{\gamma}_{ij}$ together with $\hat{\epsilon}_{it}$, the fitted residuals of ε_{it}

$$y_{it} = \alpha_i + \rho_i y_{i,t-1} + \sum_{j=1}^{k_i}\gamma_{ij}\Delta y_{i,t-j} + \epsilon_{it}$$

where k_i is chosen from data using Hall's (1994) method.

Second, to account for cross sectional dependence, estimate the variance and covariance of ε_{it}, Σ, by $\hat{\Sigma} = (1/T)\sum_{t=1}^{T} \underline{\hat{\epsilon}_t \hat{\epsilon}_t'}$ where $\underline{\hat{\epsilon}_t} = (\hat{\epsilon}_{1t}, \cdots, \hat{\epsilon}_{Nt})$ is the vector of residuals using the iterative SUR method.

Third, resample the estimated residuals with a cross section index fixed in order to preserve cross sectional dependence among individual series. This is nonparametric bootstrap since error terms are drawn from a moving-block method without further assumption on the error term distribution. Then, generate the pseudo-observations y^*_{it} for y_{it} following the above equation. Initial values of y^*_{it} are obtained from block resampling as described in Berkowitz and Kilian (2000) by dividing y_{it} into $T-k$ overlapping blocks with length $k+1$ and choose a block randomly with replacement for $y_{it}{}^*$

$$y_{it}{}^* = \hat{\alpha}_i + \hat{\rho}_i y_{i,t-1} + \sum_{j=1}^{k_i} \hat{\gamma}_{ij} \Delta y_{i,t-j} + \hat{\epsilon}_{it}{}^*$$

where $\hat{\alpha}_i$, $\hat{\gamma}_i$ and $\hat{\rho}_i$ are the SUR estimators obtained from the first step and $\hat{\epsilon}^*_{it}$ is a pseudo-innovation drawn from the resampling.

Finally, run the panel tests on the pseudo-data $y_{it}{}^*$ to derive the empirical distribution of the test statistics and the corresponding p-values. The number of replications used in each experiment is 5,000.

A.4. FINITE SAMPLE PERFORMANCE OF PANEL TESTS

The small sample properties of the panel test techniques used are well documented in their original work, but not much is known about their performances in the framework of confirmatory analysis particularly in comparison with the univariate counterparts. We conduct a Monte Carlo simulation experiment to evaluate the finite sample performance of confirmatory analysis when the panel G-test is paired with the IPS test. As the results from the PG-LLC combination are similar, they are not reported here.

Simulation is designed under the following maintained DGP for the panel sizes of $(N,T) = (10, 100)$ comparable to the actual dataset used here.

$$y_{it} = (1-\rho_i)\alpha_i + \rho_i y_{i,t-1} + u_{it}$$

where $i = 1,2,\ldots,N$, $t = 1,2,\ldots,T$ and y_{i0} is randomly selected. The y_{it} is unit-root nonstationary — hereafter I(1) — process if $\rho_i = 1$, whereas y_{it} will be stationary — hereafter I(0) — process when $\rho_i < 1$. ρ_i and α_i are randomly generated on U[0.8,0.95] and N(0,1) respectively and they are fixed at their realized values after the draw. The error term u_{it} is set to follow an AR(1) process such that $u_{it} = \theta_i u_{i,t-1} + \varepsilon_{it}$ where θ_iU[0.2,0.4], which varies over i but fixed for each model after selected. Cross-sectional dependence is incorporated in the error terms $\varepsilon_t = (\varepsilon_{1t},\ldots,\varepsilon_{Nt})'$, by drawing them from an N-dimensional multivariate normal distribution with mean zero and covariance matrix Σ following the steps outlined in Chang (2002). Cross-sectional heterogeneity is allowed in the DGP by the random generation of θ_i, ρ_i and α_i. The first 100 observations of y_{it} are discarded after generating extra observations. Each simulation run is carried out with 5,000 replications.

There are four possible outcomes when a panel test is paired with another panel test under the competing null hypothesis. Two agreement outcomes (A-R and R-A) help confirm conclusions from respective single testing because rejection of a panel test reinforces nonrejection of the other panel test, whereas two disagreement outcomes, joint rejection (R-R) and joint nonrejection (A-A), yield contradictions. The frequency of each outcome is calculated by the number of times out of 5,000 simulations that p-values of the panel G-test and the IPS test are above or below 0.05 and 0.1. As presented in Table A-1, the combination of the two panel tests has fairly good discriminatory power with reasonable sizes. For the panel of all I(1) series, the joint testing produces quite precise inferences. The probability of correct inference exceeds 92 percent at the 5 percent significance level and the frequency of disagreement outcomes is very close to the nominal sizes. The performance slightly deteriorates for the panel of all I(0) series as the correct inference is around 86 percent at the 5 percent significance level, mainly due to the increase in the R-R frequency. Nevertheless, joint inference based on two panel tests significantly improves the reliability of test inference over its univariate counterparts. For instance, the joint inference based on two popular univariate tests, the KPSS test and the DF-GLS test, correctly distinguishes I(0) series about 8 times out of 10, but merely 4 times out of 10 times for I(1) series.

A.5. CONTIGUOUS U.S. STATES BY BEA REGIONS

Group	States
New England (6)	CT, ME, MA*, NH, RI, VT
Mid-Atlantic (5)	DE, MD*, NJ, NY*, PA*
Great Lakes (5)	IL*, IN, MI*, OH*, WI
Plains (7)	IA, KS*, MN*, MO*, NE, ND, SD
Southeast (12)	AL, AR, FL, GA*, KY, LA, MS, NC, SC, TN, VA, WV
Southwest (4)	AZ, NM, OK, TX*
Rocky Mountains (5)	CO, ID, MT, UT, WY
Far West (4)	CA*, NV, OR*, WA*

Note: Each state is represented by its postal code. States with asterisk (*) represent the 15 states in which suitable metropolitan area CPI data are available.

TABLE A-1: Finite Sample Performance of Confirmatory Analysis ($N = 10$; $T = 100$)

test combinations	DGP		A-A	A-R	R-A	R-R
PG test and IPS test	I(0) panel	5%	0.0	86.2*	0.0	13.8
		10%	0.0	80.6*	0.0	19.4
	I(1) panel	5%	2.3	0.0	92.3*	5.4
		10%	1.3	0.0	88.8*	9.9
KPSS test and DF-GLS test	I(0)	5%	20.4	61.5*	7.2	10.9
		10%	11.9	62.2*	9.8	16.1
	I(1)	5%	32.8	14.8	47.9*	4.5
		10%	16.6	19.4	45.4*	18.6

Note: Entries are based on 5,000 replications with the panel size of $N = 10$, $T = 100$. In the DGP, the AR(1) coefficient ρ_i's for I(0) series are randomly generated on U[0.80,0.95]. A-A denotes the fraction of times when both tests under the conflicting null hypotheses fail to reject their respective nulls. A-R denotes the fraction of times when the first test fails to reject the null while the other test rejects the null. R-A denotes the fraction of times when thes first test rejects the null while the other test fails to reject the null. R-R denotes the fraction of times when both tests reject their respective nulls. Entries with asterisk (*) indicate the portion of correct inference. Ng-Perron's rule is used for the lag selection in the DF-GLS test. Nonparametric bootstrap method is used for the panel tests.