Convergence of international output Time series evidence for 16 OECD countries

Qing Li^{a,*}, David Papell^b

^a*Department of Economics, Wichita State University, Wichita, KS 67260-0078, USA*

^b*Department of Economics, University of Houston, Houston, TX 77204-5882, USA*

Received 27 January 1998; accepted 18 June 1998

Abstract

This article examines convergence of per capita output for 16 OECD (Organization for Economic Cooperation and Development) countries. Conventional tests on conditional and time series convergence have given mixed results for similar economies. Utilizing the concepts of deterministic and stochastic convergence, we develop techniques which incorporate endogenously determined break points to test the unit root hypothesis in relative per capita income. The tests provide evidence of deterministic convergence for 10, and stochastic convergence for 14, of the 16 OECD countries. Our findings reveal that World War II is the major cause of the structural shifts in relative output. © 1999 Elsevier Science Inc. All rights reserved.

JEL classification: C32; O40

Keywords: Stochastic convergence; Deterministic convergence; ADF test

1. Introduction

Convergence, the tendency for per capita income of different economies to equalize over time, is one of the predictions of Solow's (1956) neoclassical growth model. Over the past decade, much theoretical and empirical work has been done in this area. The implications of convergence, or lack of convergence, for long-run relationships between different countries has led to a surge of interest and debate.

Solow's model predicts that convergence exists among different economies regardless of initial conditions once the determinants of aggregate production functions are

* Corresponding author. Tel.: 316-978-3220; fax: 316-978-3308.

E-mail address: Li@twsumv.uc.twsu.edu (Q. Li)

controlled for. It therefore requires a negative correlation between initial per capita output and its growth rate, so that poorer countries will catch up with wealthier countries. Pioneered by Baumol (1986), numerous studies exploring convergence have been developed. While Romer (1986) and Delong (1988) challenge the hypothesis of cross-country convergence, Barro (1991) and Mankiw, Romer, and Weil (1992) find that convergence can be achieved among economies that exhibit similar characteristics and when human capital variables such as education and savings rates are controlled for. They refer to this cross-section notion of convergence as conditional convergence.

Another form of convergence examines long-run output movements. Bernard and Durlauf (1995) define convergence between two (or more) countries when the long-run forecasts of output differences tend to zero as the forecasting horizon tends to infinity. Tests for the time series notion of convergence require cross-country per capita output differences to be stationary. In the bivariate case, this requires that the outputs be cointegrated with cointegrating vector $[1, -1]$. We refer to this notion of convergence as time series convergence. If they are cointegrated with cointegrating vector $[1, -\lambda]$, there are common trends in output. Thus cointegration between economies is a necessary, but not a sufficient condition for convergence.

The time series evidence has not been supportive of the convergence hypothesis. Quah (1990) and Ben-David (1994) do not find general evidence of convergence among a large number of countries using the Summers-Heston (1988) data.¹ Campbell and Mankiw (1989) fail to find convergence among OECD countries which display similar economic characteristics. Bernard and Durlauf (1995), in a study of 15 OECD countries from 1900 to 1987, reject convergence but find substantial evidence of common trends.²

Although these two testing frameworks have contributed to our understanding of the growth process, they also are the cause of confusion among different studies. Bernard and Durlauf (1996) show that conditional convergence is a weaker notion of convergence than time series convergence. They find that cross-section tests tend to spuriously reject the null of no convergence when economies have different long run steady states and that failure to reject the no convergence null using time series tests can be due to transitional dynamics in the data.

Because the time series tests for convergence depend on unit root and cointegration tests which are known to have relatively more power using data over a long time period, it is important to develop an economic framework that incorporates potential structural change in the deterministic component of the trend function. Failure to do so might lead to bias towards the acceptance of no convergence and to an erroneous interpretation of output movements.

A weaker definition of convergence in the time series context, stochastic convergence, which postulates convergence if the log of relative output is trend stationary, has been proposed by Carlino and Mills (1993). This definition, however, is open to criticism because the presence of a time trend allows for permanent per capita output differences. A stronger definition of convergence, which we call deterministic convergence, is that the log of relative output is level stationary.³ The concept of time series convergence used by Bernard and Durlauf (1995) further requires that the log of relative output be level stationary with zero mean.⁴

Carlino and Mills (1993) are interested in regions of the United States, and so consider regional per capita income relative to the United States as a whole. They adopt conventional Augmented-Dickey-Fuller (ADF) tests, which fail to reject the unit root hypothesis in the log of relative regional per capita output for all eight U.S. regions. They then incorporate an exogenously imposed trend break into the tests and are able to reject the unit root null for three out of eight regions, providing some evidence of stochastic convergence. Loewy and Papell (1996) reexamine the issue by allowing endogenously determined break points and lag lengths. They are able to find evidence of stochastic convergence in seven out of eight U.S. regions, which significantly strengthens Carlino and Mills' results and provides a benchmark case for convergence among similar economies.

We examine the unit root hypothesis in relative per capita income for 16 OECD economies from 1900 to 1989. Using both conventional ADF tests and Perron's (1997) sequential tests for unit roots with endogenously determined trend breaks to investigate Carlino and Mills' (1993) notion of stochastic convergence, we are able to reject the unit root hypothesis in 14 out of 16 countries. We also test for deterministic convergence. Using ADF and Perron and Vogelsang's (1992) tests with endogenously determined breaks in the mean, as well as tests, developed below, which force the break to eliminate the deterministic trend, we are able to reject the unit root hypothesis in 10 of the 16 countries. By incorporating structural change and considering stochastic, as well as deterministic, convergence, we are able to find more evidence of convergence than found by Bernard and Durlauf (1995).

2. Empirical results

Time series notions of convergence imply that per capita output disparities between converging economies follow a stationary process. Stochastic or deterministic convergence is therefore directly related to the unit root hypothesis in relative per capita output. We utilize both conventional ADF tests as well as tests which incorporate a one-time break in the deterministic trend. Rejection of the null hypothesis of a unit root, whether or not a break is included, provides evidence of convergence. Whether it constitutes evidence of stochastic or deterministic convergence depends on the exact form of the test.

The data are from Maddison (1991), adjusted to exclude the impact of boundary changes. Per capita GDPs are calculated by dividing aggregate GDPs by mid-year population levels. While the Maddison data begins in 1870 and ends in 1989, we truncate the data before 1900 since the data in the prior years are of lower quality and are not available for Japan, the Netherlands and Switzerland. By using the more recent Maddison data, we are able to add one additional country (Switzerland) and two more years to the data analyzed by Bernard and Durlauf (1995).

We first apply conventional ADF tests to investigate unit roots in relative per capita output. We run a regression on the first difference of the logarithm of relative per capita output (to that of the group) on a constant, a trend, the lagged level of the dependent variable and k lagged first differences as seen in the following [Eq. (1)].

$$\Delta RI_t = \mu + \beta t + \alpha RI_{t-1} + \sum_{j=1}^k c_j \Delta RI_{t-j} + \epsilon_t, \quad (1)$$

where RI_t stands for the logarithm of relative per capita output of individual country at time t , which is measured by individual country's per capita output as a percentage of the aggregate per capita output of the group. Specifically, we take the logarithm of the individual country's per capita real GDP divided by the aggregate per capita real GDP, which is calculated by dividing the aggregate real GDP of all 16 countries by the total population of these countries.

We use a data dependent method to select the value of k . Following Campbell and Perron (1991), we first start with an upper bound, $k_{\max}$ (8 in this case), on k . If the t -statistic on k^{th} lagged first differences ΔRI is significant, choose $k = k_{\max}$. If not, reduce k by one until the last included lag becomes significant. If no lags are significant, we set $k = 0$. We use the 10% value of the asymptotic normal distribution (1.6) as the significance criterion. We are able to reject the unit root hypothesis in favor of a trend stationary alternative if α is significantly different from zero. In that case, stationarity of relative per capita output constitutes evidence of stochastic convergence. We also run the ADF without a time trend. In that case, rejection of the unit root hypothesis in favor of the alternative of level stationarity constitutes evidence of deterministic convergence.

The results from the ADF tests are reported in Table 1. With a time trend, the unit root hypothesis can be rejected at the 5% level for nine of the 16 countries. Without a time trend, the null is rejected for five countries. Even without a trend break, we are able to find considerable evidence for stochastic convergence, and some evidence for deterministic convergence.

Although the ADF is the standard methodology for testing the unit root hypothesis, there is much evidence that it has serious power problems. As emphasized by Campbell and Perron (1991), misspecification of the deterministic trend can bias test results towards the nonrejection of the unit root hypothesis. This is especially important in the case of long time spans of data which are likely to be affected by major structural shifts. On the other hand, in the absence of a trend break allowing for structural change, as in the sequential Dickey-Fuller tests described below, will decrease the power of the tests.

Sequential Dickey-Fuller tests which allow for a one-time change in both the intercept and the slope of the deterministic trend, with the time of break determined endogenously, have been developed by Banerjee, Lumsdaine, and Stock (1992), Christiano (1992), and Zivot and Andrews (1992). We use the specification in Perron (1997).⁵ The model is an example of an “innovational outlier” model because the change is assumed to occur gradually. The sequential ADF test involves estimating the following regressions which are given in Eq. (2):

$$\Delta RI_t = \mu + \beta t + \delta D(T_B)_t + \theta DU_t + \gamma DT_t + \alpha RI_{t-1} + \sum_{j=1}^k c_j \Delta RI_{t-j} + \epsilon_t, \quad (2)$$

where T_B is the break date. The “one-time” dummy $D(T_B)_t = 1$ if $t = T_B + 1, 0$

Table 1
ADF tests on relative per capita output

$$\Delta RI_t = \mu + \beta_t + \alpha RI_{t-1} + \sum_{j=1}^k c_j \Delta RI_{t-j} + \epsilon_t$$

Country	With trend			Without trend		
	t_α	α	k	t_α	α	k
Australia	−4.56*	−0.44	4	−0.44	−0.01	7
Austria	−2.39	−0.12	0	−2.47	−0.12	0
Belgium	−3.25***	−0.14	1	−2.53	−0.08	1
Canada	−3.48**	−0.24	5	−3.49**	−0.23	5
Denmark	−4.25*	−0.30	3	−3.00**	−0.16	1
Finland	−3.62**	−0.21	1	−0.75	−0.03	4
France	−3.55**	−0.19	3	−3.56**	−0.19	3
Germany	−3.12	−0.19	0	−1.36	−0.05	0
Italy	−2.62	−0.11	1	−2.49	−0.10	1
Japan	−1.50	−0.05	0	−0.39	−0.01	0
Netherlands	−3.67**	−0.21	1	−3.15**	−0.16	1
Norway	−3.95**	−0.22	1	−2.23	−0.08	1
Sweden	−2.95	−0.15	1	−1.43	−0.05	5
Switzerland	−3.77**	−0.17	1	−3.82*	−0.17	1
U.K.	−3.64**	−0.28	5	−0.92	−0.02	6
U.S.	−2.03	−0.08	4	−0.56	−0.01	4
Critical values for t_α						
	1%		5%		10%	
With trend	4.05		3.46		3.15	
Without trend	3.50		2.89		2.58	

* denotes statistical significance at the 1% level.

** denotes statistical significance at the 5% level.

*** denotes statistical significance at the 10% level.

Critical values are from MacKinnon (1991) with 90 observations.

otherwise, the “intercept” dummy $DU_t = 1$ if $t > T_B$, 0 otherwise, and the “slope” dummy $DT_t = t - T_B$ if $t > T_B$, 0 otherwise. With the break date chosen exogenously, this is the test in Perron (1989). In order to endogenize the break date selection, we run different regressions for $T_B = 2, 3, \dots, T-1$. The break date is chosen to minimize the t -statistic on α . The unit root null is rejected in favor of trend stationarity, and hence evidence is provided for stochastic convergence, if α is significantly different from zero. We use the same technique as described above for the ADF test to choose the lag length.⁶

The results for the sequential Dickey-Fuller tests are presented in Table 2. Incorporating trend breaks in the unit root tests significantly strengthens the findings of stochastic convergence. Using Perron’s (1997) finite sample critical values for $T = 100$, the unit root null can be rejected in favor of the trend stationary alternative at the 5% level for seven out of 16 countries: Belgium, Denmark, France, Germany, Japan, Switzerland, and the United States; and at the 10% level for one additional

Table 2
Sequential Dickey-Fuller tests for stochastic convergence
 $\Delta RI_t = \mu + \beta_t + \delta D(T_B)_t + \theta DU_t + \gamma DT_t + \alpha RI_{t-1} + \sum c_j \Delta RI_{t-j} + \epsilon_t$

Country	Year of break	α	μ	β	θ	γ	δ	k
Australia	1923	-0.53 (-5.21)	0.24 (4.83)	-0.003 (-2.20)	-0.03 (-1.03)	-0.001 (-0.25)	0.04 (1.14)	4
Austria	1943	-0.81 (-5.32)***	-0.09 (-1.88)	-0.006 (-2.90)	-1.00 (-4.79)	0.017 (4.55)	0.25 (2.23)	8
Belgium	1938	-0.39 (-6.53)*	0.05 (2.69)	-0.001 (-1.48)	-0.23 (-5.01)	0.002 (2.81)	0.13 (2.60)	1
Canada	1938	-0.26 (-4.10)	0.05 (2.66)	-0.001 (-1.93)	0.02 (-0.77)	0.001 (0.90)	-0.07 (-1.70)	2
Denmark	1938	-0.74 (-6.27)*	-0.01 (-0.51)	0.003 (2.69)	-0.02 (-0.58)	-0.003 (-2.64)	0.15 (2.73)	8
Finland	1924	-0.41 (-4.85)	-0.22 (-4.31)	-0.003 (-1.65)	-0.06 (-1.55)	0.005 (2.52)	-0.02 (-0.48)	3
France	1938	-0.57 (-6.14)**	-0.09 (-2.32)	0.002 (1.62)	-0.34 (-4.37)	0.003 (1.65)	0.24 (2.95)	8
Germany	1945	-0.29 (-6.47)*	-0.21 (-5.54)	9e-004 (1.39)	0.20 (3.63)	-0.001 (-1.07)	-0.69 (-11.1)	1
Italy	1939	-0.29 (-4.93)	-0.10 (-3.30)	-3e-005 (-0.04)	-0.28 (-4.02)	0.004 (3.18)	0.08 (1.18)	1
Japan	1943	-0.60 (-5.82)**	-0.61 (-5.32)	0.004 (2.94)	-0.97 (-5.51)	0.014 (5.22)	0.23 (2.41)	8
Netherlands	1937	-0.42 (-5.08)	0.01 (0.35)	0.002 (1.01)	-0.16 (-2.38)	-0.001 (-0.10)	0.12 (1.32)	3
Norway	1938	-0.32 (-5.13)	-0.17 (-4.60)	0.004 (3.80)	-0.02 (-0.66)	-0.001 (-1.53)	0.05 (0.99)	1
Sweden	1926	-0.30 (-4.69)	-0.04 (-1.93)	-0.002 (-2.17)	-0.00 (0.20)	0.003 (2.31)	-0.06 (-1.52)	1
Switzerland	1943	-0.30 (-5.83)**	0.02 (1.56)	1e-004 (0.24)	0.20 (4.49)	-0.003 (-3.18)	-0.12 (-2.44)	1
U.K.	1928	-0.33 (-4.46)	0.15 (4.04)	-0.004 (-3.51)	0.01 (0.45)	0.001 (0.50)	-0.04 (-1.11)	3
U.S.	1938	-0.36 (-5.96)**	0.21 (6.02)	-0.002 (-3.04)	0.16 (4.65)	-0.002 (-2.82)	-0.06 (-1.96)	8
Critical values for t_α								
		1%			5%			10%
		-6.21			-5.55			-5.25

* denotes statistical significance at the 1% level.
** denotes statistical significance at the 5% level.
*** denotes statistical significance at the 10% level.
Critical values are from Perron (1994). t -statistics are in parentheses.

country, Austria.⁷ Some of the nonrejections appear to be caused by the decreased power of the sequential Dickey-Fuller test in the absence of a trend break. Of the remaining eight countries, the unit root null can be rejected at the 5% level by the ADF test in six cases.

The ADF and sequential Dickey-Fuller tests together provide strong evidence in support of stochastic convergence for the 16 OECD countries. The only two countries for which we are unable to reject the unit root null at the 5% level in either test are Italy and Sweden. Therefore, for 14 out of 16 OECD countries, we are able to reject the unit root hypothesis, a result that provides us with evidence of stochastic convergence.⁸

Testing for a unit root in relative per capita income is only an interesting question if at least some of the individual series are nonstationary. If all of the individual series were stationary, the relative series would necessarily be stationary.⁹ We conducted, but do not report, ADF and sequential Dickey-Fuller tests on the individual series. In the absence of breaks, the unit root null cannot be rejected by the ADF test for any country. With breaks, the null can be rejected in half of the cases, but with a wide variety of break dates.¹⁰

The trend breaks and dummy variables contain economic implications. By examining the coefficients of the dummies, most of the 16 countries can be put into two categories: those which are characterized by an initial fall in output (indicated by a negative sum of the one-time and intercept dummies), and a faster growth rate afterwards (indicated by a positive slope dummy), and those which are characterized by an initial rise in output (a positive sum of the one-time and intercept dummies), followed by a slower growth (a negative slope dummy).

The trend breaks for per capita relative output mostly take place around World War II. Austria, Germany, Japan and Switzerland experience breaks around the end of the war, while Belgium, Canada, Denmark, France, Italy, the Netherlands, Norway and the United States have breaks at the beginning of the war. In their study of per capita output for the same countries, Ben-David and Papell (1995) find that the Great Depression instead of World War II is the cause of the breaks for the United States and Canada. The different impact of the two events on these countries may explain the discrepancy. Unlike World War II, the Great Depression affected almost every industrialized country, it therefore fails to induce breaks in *relative* per capita output.

Twelve of the 16 countries experience breaks around World War II. Most display an initial fall in output, followed by a higher growth rate. The United States, which was relatively unaffected by the war, is an exception. Per capita output for the United States (relative to the group) initially rises with the onset of the war, followed by slower (relative) growth. Breaks occur in the 1920s for Australia, Finland, Sweden, and the United Kingdom. In Australia, World War I brought shortages of goods which led to postwar industrial expansion. Finland's break can be explained by its independence from the Soviet Union in 1920 and the subsequent civil war. The United Kingdom's break occurs in 1928, shortly after a wave of strikes culminated in the general strike of 1926.

Sweden is one of two countries where the unit root cannot be rejected. What

separates Sweden from the rest of the world is its longtime political neutrality. During World War II, Sweden was the only Scandinavian country to succeed in remaining neutral, and emerged from both wars virtually undamaged. Its break date of 1926 can be explained by widespread unemployment during the recession in the 1920s.

We now consider deterministic convergence, the type of time series convergence defined and tested by Bernard and Durlauf (1995). Deterministic convergence requires that the log of relative output be level stationary, and is hence considered to be a stronger notion of time series convergence than stochastic convergence. We test the unit root hypothesis in relative per capita income allowing a structural change in the mean. Based on Perron and Vogelsang (1992), we estimate the following equation:

$$\Delta RI_t = \mu + \delta D(T_B)_t + \theta DU_t + \alpha RI_{t-1} + \sum_{j=1}^k c_j \Delta RI_{t-j} + \epsilon_t, \quad (3)$$

Eq. (3) differs from Eq. (2) by excluding the time trend and the trend dummy variable. Perron and Vogelsang (1992) compute finite sample critical values for $T = 100$, with the lag length chosen as described before. The results are shown in Table 3. The unit root null can be rejected in favor of the level stationary alternative at the 5% level for six out of 16 countries: Austria, Belgium, Denmark, France, Germany, and the Netherlands; and at the 10% level for two additional countries, Sweden and Switzerland.¹¹ In the absence of a break, the unit root null can also be rejected at the 5% level by the ADF test for Canada. With the unit root null being rejected for over half of the countries, this result provides considerable, but by no means universal, evidence of deterministic convergence.

The concept of deterministic convergence requires the elimination of the deterministic, as well as the stochastic, trend. An alternative test of deterministic convergence consists of testing for a unit root in Eq. (2) under the restriction that the trend break eliminates the deterministic trend. This involves estimating sequential Dickey-Fuller tests for Eq. (2) with $\beta + \gamma = 0$. If the null hypothesis of a unit root can be rejected, convergence will be to a zero postbreak trend, and hence the conditions for deterministic convergence will be satisfied.

The critical values for this test are different from the two tests described above and have not, to our knowledge, been tabulated. Following Perron and Vogelsang (1992) and Perron (1997), we consider the random walk data generating process for our sample size of 90 observations as given in Eq. (4).

$$y_t = y_{t-1} + \epsilon_t, \quad (4)$$

with $y_0 = 0$ and ϵ_t to be i.i.d. $N(0,1)$. Our test statistic is the t -statistic on α in Eq. (2), under the restriction that $\beta + \gamma = 0$, with the lag length k for each series chosen as described above. Repeating this process 5000 times, the critical values for the finite sample distributions are obtained from the sorted vector of the replicated statistics.

The results of this test, as well as the finite sample critical values, are reported in Table 4. The critical values are, as expected, in between those of Perron and Vogelsang (1992) and Perron (1997).¹² Turning to the results, the unit root null could be rejected at the 5% level for Austria, Belgium, Denmark, and Germany, and at the 10% level

Table 3

Sequential Dickey-Fuller tests for deterministic convergence

$$\Delta RI_t = \mu + \delta D(T_B)_t + \theta DU_t + \alpha RI_{t-1} + \sum c_j \Delta RI_{t-j} + \epsilon_t$$

Country	Year of break	α	μ	θ	δ	k
Australia	1920	−0.05 (−1.74)	0.01 (0.90)	−0.02 (−1.65)	0.09 (2.77)	7
Austria	1944	−0.14 (−4.80)**	−0.05 (−4.76)	0.02 (2.08)	−0.85 (−17.5)	4
Belgium	1938	−0.31 (−5.43)*	0.02 (2.34)	−0.09 (−4.64)	0.08 (1.57)	1
Canada	1945	−0.20 (−3.58)	0.02 (2.66)	−0.00 (−0.45)	0.12 (3.20)	2
Denmark	1937	−0.44 (−5.92)*	0.03 (3.29)	−0.07 (−4.49)	0.14 (2.78)	3
Finland	1917	−0.08 (−2.32)	−0.07 (−2.73)	0.05 (3.10)	−0.20 (−3.86)	7
France	1946	−0.25 (−4.92)**	−0.03 (−1.98)	−3e-004 (−0.02)	−0.41 (−3.98)	3
Germany	1945	−0.28 (−8.25)*	−0.19 (−7.81)	0.15 (8.15)	−0.68 (−11.1)	1
Italy	1944	−0.14 (−3.46)	−0.06 (−3.58)	0.03 (2.34)	−0.22 (−3.09)	1
Japan	1944	−0.04 (−2.88)	−0.04 (−2.50)	0.05 (4.49)	−0.66 (−14.23)	1
Netherlands	1946	−0.29 (−4.83)**	0.02 (1.32)	−0.05 (−2.43)	−0.45 (−3.29)	1
Norway	1943	−0.17 (−3.95)	−0.05 (−3.54)	0.04 (3.24)	−0.12 (−2.27)	1
Sweden	1928	−0.25 (−4.47)***	−0.06 (−4.00)	0.06 (3.84)	−0.04 (−1.10)	1
Switzerland	1942	−0.20 (−4.39)***	0.02 (2.06)	0.02 (1.57)	−0.10 (−2.05)	1
U.K.	1940	−0.09 (−2.60)	0.02 (1.75)	−0.03 (−2.56)	0.03 (0.68)	6
U.S.	1958	−0.13 (−3.05)	0.06 (2.92)	−0.04 (−2.87)	0.03 (1.00)	1
Critical values for t_α						
1%		5%		10%		
−5.33		−4.58		−4.27		

* denotes statistical significance at the 1% level.

** denotes statistical significance at the 5% level.

*** denotes statistical significance at the 10% level.

Critical values are from Perron and Vogelsang (1992). t -statistics are in parentheses.

Table 4

Sequential Dickey-Fuller tests with restrictions

$$\Delta RI_t = \mu + \beta_t + \delta D(T_B)_t + \theta DU_t + \gamma DT_t + \alpha RI_{t-1} + \sum c_j \Delta RI_{t-j} + \epsilon_t$$

Restriction: $\beta + \gamma = 0$							
Country	Year of break	α	μ	$\beta (-\gamma)$	θ	δ	k
Australia	1974	-0.49 (-4.94)***	0.23 (4.73)	-0.004 (-4.80)	-0.29 (-4.84)	0.04 (1.26)	4
Austria	1944	-0.16 (-5.20)**	-0.03 (-1.47)	-0.001 (-1.81)	-0.01 (-0.53)	-0.85 (-17.8)	4
Belgium	1938	-0.33 (-5.58)**	0.04 (2.18)	-9e-004 (-1.21)	-0.12 (-4.20)	0.08 (1.63)	1
Canada	1916	-0.24 (-3.98)	-0.00 (-0.12)	0.004 (1.52)	0.02 (0.78)	0.07 (2.04)	2
Denmark	1938	-0.73 (-6.46)*	-0.01 (-0.56)	0.003 (2.71)	-0.04 (-1.38)	0.15 (2.80)	8
Finland	1977	-0.22 (-3.65)	-0.15 (-3.45)	0.001 (3.01)	0.12 (3.35)	-0.03 (-0.64)	1
France	1946	-0.26 (-4.93)***	-0.01 (-0.26)	-8e-004 (-0.77)	-0.02 (-0.65)	-0.41 (-3.93)	3
Germany	1945	-0.29 (-8.41)*	-0.22 (-6.74)	9e-004 (1.44)	0.19 (6.70)	-0.69 (-11.2)	1
Italy	1955	-0.19 (-3.71)	-0.05 (-2.06)	-0.001 (-1.98)	0.00 (0.07)	-0.03 (-0.50)	1
Japan	1958	-0.12 (-2.72)	-0.10 (-2.32)	-2e-004 (-0.34)	0.09 (2.55)	-0.06 (-0.68)	0
Netherlands	1946	-0.32 (-5.05)***	0.06 (1.85)	-0.002 (-1.41)	-0.09 (-2.56)	-0.44 (-3.24)	1
Norway	1977	-0.23 (-4.03)	-0.09 (-3.47)	9e-004 (2.67)	0.10 (3.28)	-0.02 (-0.37)	1
Sweden	1928	-0.31 (-4.98)***	-0.04 (-2.40)	-0.002 (-2.01)	0.04 (2.09)	-0.06 (-1.42)	1
Switzerland	1973	-0.27 (-5.02)***	0.01 (0.99)	7e-004 (2.42)	-0.00 (-0.28)	0.06 (1.33)	1
U.K.	1964	-0.23 (-3.54)	0.09 (3.16)	-0.002 (-3.15)	-0.13 (-3.35)	0.02 (0.48)	3
U.S.	1958	-0.13 (-3.02)	0.06 (2.83)	-4e-005 (-0.16)	-0.04 (-2.59)	0.03 (0.99)	1
Critical values for t_α							
1%		5%		10%			
-5.67		-5.12		-4.80			

* denotes statistical significance at the 1% level.

** denotes statistical significance at the 5% level.

*** denotes statistical significance at the 10% level.

Critical values are determined as described in the text for 90 observations. t -statistics are in parentheses.

for Australia, France, Netherlands, Sweden, and Switzerland. These generally accord with the previously reported results for deterministic convergence, although the rejections for several of the countries are at lower significance levels. The rejection of the unit root null for Australia was not found above, and provides additional evidence of deterministic convergence.

3. Conclusions

The neoclassical growth theory originated by Solow (1956) has sparked many empirical efforts to determine the long run behavior of different countries. A large literature on convergence across economies has produced a diverse body of testing methodologies as well as definitions. While most convergence studies focus on testing the cross-section version of convergence, referred to as conditional convergence, Bernard and Durlauf (1996) find that the time series version is a stronger notion of convergence.

Bernard and Durlauf (1995) use time and frequency domain approaches to examine persistence in per capita output deviations, and fail to find evidence of deterministic convergence for 15 OECD countries. We employ time series techniques which incorporate structural breaks to explore both deterministic and stochastic convergence among 16 OECD countries. In particular, we test the unit root hypothesis on the log relative per capita output (to that of the group). If we find evidence against the unit root null for its relative per capita output, a country's output is converging to the aggregate output of the whole group.

We find considerable evidence of convergence among the 16 OECD countries. Combining tests with and without structural breaks, we can reject the unit root null against an alternative of trend stationarity, and thus provide evidence for stochastic convergence, for 14 of the 16 countries. We can also reject the unit root null in favor of a level stationary alternative, and provide evidence of deterministic convergence, for 10 of the 16 countries. The results of the sequential unit root tests also reveal that World War II is the major cause for the structural shifts of relative per capita outputs.

Acknowledgments

We are grateful to Andrew Bernard, Dan Ben-David, Michael Loewy, and Kei-Mu Yi for helpful comments and discussions.

Notes

1. Ben-David (1994) finds some evidence in support of Baumol's finding of a convergence club among both the world's wealthiest countries and among the world's poorest countries.
2. Li (1994) conducts bivariate and multivariate cointegration tests using techniques where cointegration, as well as no cointegration, is the null hypothesis.

The combined results provide some evidence of deterministic cointegration among most OECD countries.

3. This terminology corresponds to Ogaki and Park's (1998) definition of stochastic cointegration when the stochastic trend components of the (two or more) series are cointegrated, and deterministic cointegration when both the stochastic and deterministic trends are eliminated by the cointegrating vector.
4. Bernard and Durlauf (1991) report that allowing for a constant difference in logs did not markedly change their results.
5. These tests all restrict breaks to one-time changes. Unit root tests which allow for two breaks have been developed by Lumsdaine and Papell (1997) and applied to long-term international output by Ben-David, Lumsdaine, and Papell (1997).
6. Ng and Perron (1995) discuss the advantages of this recursive t -statistic method over an alternative procedure where k is chosen to minimize the Akaike Information Criterion.
7. Since we have 90 observations, using critical values with $T = 100$ could potentially cause us to inappropriately reject the unit root null. Perron (1997) also provides critical values for $T = 70$. The only difference is that the null for Denmark is rejected at the 5%, rather than the 1%, level. The critical values incorporate the lag length selection methods. They are considerably lower, and hence more conservative, than either asymptotic critical values or finite sample critical values which fix the value of k .
8. For any given version of the tests, however, we can reject the unit root null for only about one-half of the countries.
9. If the output series are trend stationary, convergence requires that the time trends for each country's output must be the same.
10. In the absence of breaks, Ben-David and Papell (1995), using the Maddison data for individual countries extended back to 1870, can only reject the unit root null at the 5% level for the United States. With breaks, they can reject the null for 11 countries, but only nine of these remain significant with finite sample critical values.
11. Perron and Vogelsang (1992) also tabulate critical values for $T = 50$, a much smaller number than our 90 observations. The only difference from using these critical values is that the null is no longer rejected for Switzerland. This makes no difference for the convergence results because, using the ADF test, the null for Switzerland can be rejected at the 1% level.
12. In order to provide a comparison with previous work, we also simulated critical values with $T = 100$. The results were very close to those with $T = 90$.

References

- Banerjee, A., Lumsdaine, R. L., & Stock, J. (1992). Recursive and sequential tests of the unit root and trend break hypothesis: theory and international evidence. *Journal of Business and Economic Statistics* 10, 271–287.

- Barro, R. J. (1991). Economic growth in a cross-section of countries. *Quarterly Journal of Economics* 106, 407–443.
- Baumol, W. J. (1986). Productivity growth, convergence, and welfare. *American Economic Review* 76, 1072–1085.
- Ben-David, D. (1994). Convergence Clubs and Diverging Economies. CEPR Discussion Paper, 922.
- Ben-David, D., Lumsdaine, R. L., & Papell, D. H. (1997). The Unit Root Hypothesis in Long-Term Output: Evidence from Two Structural Breaks for 16 Countries. Working Paper, University of Houston.
- Ben-David, D., & Papell, D. H. (1995). The Great Wars, the Great Crash and steady state growth: some new evidence about an old stylized fact. *Journal of Monetary Economics* 36, 453–475.
- Bernard, A. B., & Durlauf, S. N. (1991). Convergence of International Output Movements. NBER Working Paper, 3717.
- Bernard, A. B., & Durlauf, S. N. (1995). Convergence of international output. *Journal of Applied Econometrics* 10, 97–108.
- Bernard, A. B., & Durlauf, S. N. (1996). Interpreting tests of the convergence hypothesis. *Journal of Econometrics* 71, 161–174.
- Campbell, J. Y., & Mankiw, N. G. (1989). International evidence on the persistence of economic fluctuations. *Journal of Monetary Economics* 23, 319–333.
- Campbell, J. Y., & Perron, P. (1991). What macroeconomists should know about unit roots. In O. Blanchard & S. Fischer (Eds.), *NBER Macroeconomics Annual* (pp. 141–201). Cambridge, MA: MIT Press.
- Carlino, G. A., & Mills, L. O. (1993). Are U.S. regional incomes converging? A time series analysis. *Journal of Monetary Economics* 32, 335–346.
- Christiano, L. (1992). Searching for a break in GNP. *Journal of Business and Economic Statistics* 10, 237–250.
- DeLong, J. B. (1988). Productivity growth, convergence, and welfare: comment. *American Economic Review* 78, 1138–1154.
- Li, Q. (1994). Cointegration of International Output Movements: the Case of OECD Countries. Working Paper, Wichita State University.
- Loewy, M. B., & Papell D. H. (1996). Are U.S. regional incomes converging? Some further evidence. *Journal of Monetary Economics* 38, 587–598.
- Lumsdaine, R. L., & Papell D. H. (1997). Multiple trend breaks and the unit root hypothesis. *Review of Economics and Statistics* 79, 212–218.
- MacKinnon, J. (1991). Critical values for cointegration tests. In R. F. Engle & C. W. J. Granger (Eds.), *Readings in Cointegration* (pp. 267–276). New York: Oxford University Press.
- Maddison, A. (1991). *Dynamic Forces in Capitalist Development: A Long-Run Comparative View*. Oxford: Oxford University Press.
- Mankiw, N. G., Romer, D., & Weil, D. N. (1992). A contribution to the empirics of economic growth. *Quarterly Journal of Economics* 107, 407–438.
- Ng, S., & Perron, P. (1995). Unit root tests in ARMA models with data dependent methods for the selection of the truncation lag. *Journal of the American Statistical Association* 90, 268–281.
- Ogaki, M., & Park, J. Y. (1998). A cointegration approach to estimating preference parameters. *Journal of Econometrics* 82 (1), 107–134.
- Perron, P. (1989). The Great Crash, the oil price shock, and the unit root hypothesis. *Econometrica* 57, 1361–1401.
- Perron, P. (1994). Trend, unit root and structural change in macroeconomic time series. In B. B. Rao (Ed.), *Cointegration for the Applied Economist* (pp. 113–140). New York: St. Martin's.
- Perron, P. (1997). Further evidence on breaking trend functions in macroeconomic variables. *Journal of Econometrics* 80, 355–385.
- Perron, P., & Vogelsang, T. (1992). Nonstationarity and level shifts with an application to purchasing power parity. *Journal of Business and Economic Statistics* 10, 301–320.
- Quah, D. (1990). International Patterns of Growth I: Persistence in Cross-Country Disparities. Working Paper, MIT.

- Romer, P. (1986). Increasing returns and long run growth. *Journal of Political Economy* 94, 1002–1037.
- Solow, R. M. (1956). A contribution to the theory of economic growth. *Quarterly Journal of Economics* 70, 65–94.
- Summers, R., & Heston, A. (1988). A new set of international comparisons of real product and price levels estimates for 130 countries, 1950–1985. *Review of Income and Wealth* 34, 1–25.
- Zivot, E., & Andrews, D. (1992). Further evidence on the Great Crash, the oil price shock, and the unit root hypothesis. *Journal of Business and Economic Statistics* 10, 251–270.