Department of Economics

Business Cycle Volatility and Economic Growth
Research Paper No. 00-6

CONVERGENCE IN INTERNATIONAL OUTPUT: EVIDENCE FROM PANEL DATA UNIT ROOT TESTS

Mark J. Holmes

October 2000

This paper forms part of the ESRC funded project (Award No. L1382511013) "Business Cycle Volatility and Economic Growth: A Comparative Time Series Study", which itself is part of the *Understanding the Evolving Macroeconomy* Research programme.

Loughborough University

CONVERGENCE IN INTERNATIONAL OUTPUT: EVIDENCE FROM PANEL DATA UNIT ROOT TESTS

Abstract

This paper investigates international output convergence using methods of panel data unit root testing advocated by Im et al. (1997) and Breuer et al. (1999). Using quarterly data for a sample of OECD economies for the period 1960-98 on GDP differentials, the evidence suggests that power deficiency may be an issue where univariate ADF unit root tests find against convergence with respect to the US or Germany. However, while the Im et al. t-bar test offers strong evidence in favor of convergence, the Breuer et al. SURADF test suggests that this finding may in fact be driven by the rejection of non-stationarity in a small number of cases.

Keywords: unit root testing, panel data, convergence.
JEL Codes: C2, C3, F0.

1. Introduction

A variety of studies [see, *inter alia*, Serletis and Krichel (1992), Bernard and Durlauf (1995), Greasley and Oxley (1997) and Mills and Holmes (1999)] have measured the extent of international output convergence. Using GDP or index of industrial production data for OECD economies over varying study periods, evidence of long-run convergence is mixed. This paper examines output convergence from a new angle through the application of panel data unit root testing.[1] The focus here is on long-run bivariate convergence between a sample of countries and the US and then Germany. Panel data unit root testing offer a means of overcoming problems of low test power associated with univariate ADF tests. We employ the t-bar panel data unit root tests advocated by Im *et al*. (1997) on panels of output deviations from a base country. In addition, this study addresses two shortcomings associated with many existing panel

[1] The most common application of panel data unit root testing is the search for purchasing power parity via the stationarity of real exchange rates. See, for example, Wu (1996), Coakley and Fuertes (1997), O'Connell (1998), Taylor and Sarno (1998).

data unit root tests. First, rejection of the null of joint non-stationarity in a panel may be due to a single series within the panel being stationarity. Second, while the Im *et al.* procedure has key advantages over early panel data tests, the demeaning procedure does not exploit information in error covariances in an entirely satisfactory manner. For these reasons, this study also employs a technique advocated by Breuer *et al.* (1999) which involves the estimation of ADF regressions within a seemingly unrelated regression (SUR) framework. The SURADF results are used to qualify the results obtained through the Im *et al*. procedure. While there is little evidence of convergence using univariate ADF tests, the t-bar panel data unit root tests identify strong convergence using either the US or Germany as the base country. Application of the SURADF technique, however, suggests that convergence is present in a small number of cases where convergence with the US is stronger than with Germany.

2. Data, Methodology and Results

Quarterly GDP data are employed for eleven countries- Austria, Belgium, Canada, Denmark, France, Germany, Italy, Japan, Netherlands, UK and US- for the period 1960Q1-98Q4. The data are obtained from the *OECD Main Economic Indicators*.[2] Output differentials are defined with respect to the US and then Germany to compare the extent of convergence allowing for the possibility that convergence with Germany may be the more relevant for the EU economies. Whether or not long-run output convergence prevails between the natural logarithms of domestic (non-US or German) and base (US or German) real output (respectively denoted as y_{it} and y_{jt}) depends on the time series properties of the x_{it} which is computed as

[2] Real GDP data for Belgium are only available on an annual basis. Quarterly values have been interpolated using the index of industrial production.

$$x_{it} = y_{it} - y_{jt} \tag{1}$$

where $i = 1, 2, \ldots, N$ non-base countries, $j = US, Germany$ and $t = 1, 2, \ldots, T$ time periods. Table 1 reports the results from univariate ADF tests. At the 5% significance level, non-stationarity is rejected in only four cases: Germany-US, UK-US, Japan-Germany and UK-Germany. The univariate ADF tests may suffer from low power and are therefore unable to reject the null. Panel data unit root testing, on the other hand, utilizes more observations where the cross-country variations of the data in estimation are exploited.

Following Im *et al*. (1997), suppose x_{it} is generated by a first order autoregressive process

$$\Delta x_{it} = a_i + f_i x_{i,t-1} + e_{it} \tag{2}$$

where e_{it} is a disturbance term. The null hypothesis is H_0: $f_i = 0 \ \forall \ i$ and the alternative is H_1: $f_i < 0, i = 1, 2, ..., N_1, f_i = 0, i = N_1 + 1, N_1 + 2, ..., N$. The disturbances across the panel may be correlated. Indeed, wrongly assuming identically and independently distributed disturbances can have dramatic implications for statistical size and power to the extent that the null may not be correctly accepted or rejected. Let $e_{it} = q_t + u_{it}$ where q_t is a time-specific common effect that allows for a degree of dependency across the series and u_{it} is an idiosyncratic random effect that is independently distributed across groups. To remove the effect of q_t subtract the cross-section means from both sides of (2) to obtain the following demeaned regression

$$\Delta \tilde{x}_{it} = \tilde{a}_i + \tilde{f}_i \tilde{x}_{i,t-1} + \tilde{x}_{it} \tag{3}$$

For a heterogeneous panel with serially correlated errors (3) may be rewritten as

$$\Delta \tilde{x}_{it} = \tilde{a}_i + \tilde{f}_i \tilde{x}_{i,t-1} + \sum_{k=1}^{q_i} r_{ik} \Delta \tilde{x}_{i,t-k} + \tilde{x}_{it} \tag{4}$$

Equation (4) forms the basis of the t-bar test for inflation convergence. Using data for $x_{it} = y_{it} - y_{jt}$ it is estimated for a sample involving the major EU economies where the t-bar statistic is calculated using the average value of the individual ADF statistics based on each $\tilde{f}_i$.[3]

The t-bar test results for both the US and Germany differentials are reported in Table 2. In the both cases, the null of joint non-stationarity is strongly rejected at the 1% significance level for the full panel of 10 countries.[4] It is possible that such a strong rejection is being driven by the inclusion of the univariate-stationary series. These differentials are therefore excluded from their respective panels to form the groups of '8 countries' and '7 countries'. In each case, we have panels that comprise univariate non-stationary series only. Again, the null of joint non-stationarity is strongly rejected at the 1% significance level indicating real convergence.

It might, however, be excessive to conclude that convergence holds for all members of each panel. The t-bar test does not allow for how many and which members contain a unit root. Also, a demeaning procedure has been employed to deal with contemporaneous correlation.[5] To address these problems, Breuer *et al.* (1999) advocate a panel data unit root test which involves estimating ADF regressions in a seemingly unrelated regression (SUR) framework and then testing for individual unit roots within the panel. The SURADF test is more powerful than independently estimated single equation ADF tests. Earlier SUR-based tests of Abuaf and Jorion

---

[3] The t-bar test requires that the t-bar statistic follows an asymptotic normal distribution as both $N \to \infty$ and $T \to \infty$ with $(N/T) \to k$ where *k* is a finite positive constant.

[4] Using the average estimates of $\tilde{f}_i$, the half life of a random shock in the case of the US (German) differentials is computed as 11.696 (12.995) quarters.

[5] While O'Connell (1998) argues that a demeaning approach is preferable to the use of time dummies, the relative merits of controlling for contemporaneous correlation through the estimation of the covariance matrix for e_{it} are stressed.

(1990), O'Connell (1998) and Taylor and Sarno (1998) have an 'all or nothing' characteristic to their tests where all series were either stationary or non-stationary.[6] The Breuer *et al.* SURADF procedure allows f_i to differ across the series under the alternative hypothesis while exploiting information in error covariances to produce efficient estimators with potentially powerful test statistics. The test statistics from the SUR model feature nonstandard distributions with critical values that must be derived through simulations.

Table 3 reports the SUR estimates for the '8 countries' and '7 countries' panels along with 1, 5 and 10% critical values tailored to each ADF statistic that have been generated using Monte Carlo simulations. With regard to US-based output deviations, the null of non-stationarity is rejected at the 5% significance level in the case of Italy. In the previous t-bar tests it is likely that a single series accounts for rejecting the null of joint non-stationarity. However, at the 10% significance level it is also possible to include Austria, Belgium and Japan as stationary series within the panel. In the case of German-based output differentials, the null of stationarity can only be rejected in the cases of Italy and the Netherlands at the 10% significance level. This is weaker evidence of convergence in this particular panel.

3. Summary and Conclusion

Using quarterly data for 1960-98, univariate ADF tests offer limited evidence that output differentials defined against the US or Germany are stationary. It is possible that such a finding is attributable to low test power as the outcome is dramatically

[6] The Taylor and Sarno SUR-based test is accompanied by a Johansen likelihood test of the null that the long-run matrix of the series is less than full rank. This can consume considerable degrees of freedom as the panel expands and can still leave the researcher unable to infer the breakdown between stationary and non-stationary series.

modified to one of strong convergence if the univariate non-stationary data are used to create a panel and the Im *et al.* technique is applied. Further seemingly unrelated regression ADF tests based on the Breuer *et al.* suggest that this rejection of the joint null of non-stationarity is in fact driven by a small number of countries and that convergence with the US is greater than with Germany.

Table 1. Univariate ADF Unit Root Tests

$y_{it} - y_{USt}$	*No Trend*	*Trend*
Austria	-2.251	-2.295
Belgium	-1.674	-2.149
Canada	-1.800	-0.329
Denmark	-0.232	-2.022$^{\#}$
France	-0.846	-0.957
Germany	-2.908**	-3.258*
Italy	-1.695	-0.629
Japan	-2.483	-0.169
Netherlands	-1.459	-1.826
UK	-1.744	-3.510$^{**,\#}$
$y_{it} - y_{GERt}$		
Austria	-2.048	-2.089
Belgium	-0.918	-2.018
Canada	-2.208	-2.298
Denmark	-0.956	-2.038
France	-1.927	-1.623
Italy	-2.126	-1.158
Japan	-3.027**	-0.737
Netherlands	-2.578*	-2.493
UK	-2.211	-3.470$^{**,\#}$

For each ADF regression, the lag length was chosen using Said and Dickey's (1984) $T^{1/3}$ rule. In all cases, the residuals were free of serial correlation. The conclusions were qualitatively unaffected by the employment of alternative procedures for lag length selection. ** and * indicate rejection of the null of non-stationarity at the 5 and 10% levels of significance respectively, # denotes the significance of the time trend in the ADF regression at the 5% level. For regressions excluding a trend, relevant critical values taken from Fuller (1976) are -2.89 and -2.58, while for regressions including a trend, these are -3.45 and -3.15 respectively.

Table 2. IPS Panel Data Unit Root Tests

		Critical Values		
$y_{it} - y_{USt}$	t-bar	1%	5%	10%
10 Countries	-2.266***	-2.145	-1.967	-1.869
8 Countries	-2.418***	-2.224	-2.203	-1.913
$y_{it} - y_{GERt}$				
10 Countries	-2.161***	-2.145	-1.967	-1.869
7 Countries	-2.305***	-2.315	-2.078	-1.951

These are t-bar tests based on equation (4). The lag length, q_i, is chosen using Said and Dickey's (1984) $T^{1/3}$ rule. In all cases, the residuals were free of serial correlation. The conclusions were qualitatively unaffected by the employment of alternative procedures for lag length selection. *** denotes rejection of the null of non-stationarity at the 1% level of significance. Critical values are simulated with 10,000 replications.

Table 3. SURADF Analysis

		Critical Values		
$y_{it} - y_{USt}$	SURADF	1%	5%	10%
Austria	-3.036*	-3.899	-3.382	-3.027
Belgium	-2.823*	-3.716	-2.906	-2.715
Canada	-2.388	-3.560	-3.036	-2.845
Denmark	-0.878	-3.578	-3.174	-2.806
France	-2.244	-3.926	-3.378	-2.893
Italy	-4.039***	-3.719	-3.204	-2.969
Japan	-3.522*	-4.350	-3.587	-3.147
Netherlands	-2.934	-3.582	-3.161	-2.940
$y_{it} - y_{GERt}$				
Austria	-2.590	-4.227	-3.750	-3.338
Belgium	-2.416	-3.769	-3.543	-3.219
Canada	-2.335	-3.984	-3.581	-3.152
Denmark	-2.269	-3.444	-3.099	-2.982
France	-2.785	-3.544	-3.116	-2.915
Italy	-3.625*	-4.106	-3.689	-3.244
Netherlands	-3.319*	-3.913	-3.528	-3.135

The lag length, q_i, is chosen using Said and Dickey's (1984) $T^{1/3}$ rule. In all cases, the residuals were free of serial correlation. The conclusions were qualitatively unaffected by the employment of alternative procedures for lag length selection. Critical values specific to each series in the panels are simulated using 10000 replications based on the estimated covariance matrix of the system, *N* and *T*.

References

Abuaf, Niso and Phillipe Jorion (1990), ‘Purchasing Power Parity in the Long-Run’, *Journal of Finance*, 45, 157-74.

Bernard, Andrew B., and Steven N. Durlauf (1995), ‘Convergence in International Output’, *Journal of Applied Econometrics*, 10, 97-108.

Breuer, Janice Boucher, McNown, Robert, and Myles Wallace (1999), ‘Series-specific Tests for a Unit Root in a Panel Setting with an Application to Real Exchange Rates’, Discussion Paper No. 99-9, University of Colorado, June.

Coakley, Jerry, and Ana-Maria Fuertes (1997), ‘New Panel Unit Root Tests of PPP’, *Economics Letters*, 57, 17-22.

Greasley, David and Les Oxley (1997), 'Time-Series Tests of the Convergence Hypothesis: Some Positive Results', *Economics Letters*, 56, 143-7.

Fuller, Wayne. A. (1976), *Introduction to statistical time series.* New York: John Wiley.

Im, Kyong S., Pesaran, M. Hashem, and Yongcheol Shin (1997), ‘Testing for Unit Roots in Heterogeneous Panels’, Department of Applied Economics Working Paper, University of Cambridge, December.

Mills, Terry and Mark J. Holmes (1999), ‘Common Trends and Cycles in European Industrial Production: Exchange Rate Regimes and Economic Convergence’, *Manchester School*, 67, 557-87.

O’Connell, Paul (1998), ‘The Overvaluation of Purchasing Power Parity’, *Journal of International Economics*, 44, 1-19.

Said, Said E., and David A. Dickey (1984), ‘Testing for Unit Roots in Autoregressive Moving Average Models of Unknown Order’, *Biometrica*, 71, 599-608.

Serletis, A., and Thomas Krichel (1992), ‘Output Trends in the EC Countries and the Implications for Transition to Monetary Union’, *Economics Letters*, 40, 211-6.

Taylor, Mark P., and Lucio Sarno (1998), ‘The Behavior of the Real Exchange Rate During the Bretton Woods Period’, *Journal of International Economics*, 46, 281-312.

Wu, Yangru (1996), ‘Are Real Exchange Rates Non-Stationary? Evidence from a Panel Data Unit Root Test’, *Journal of Money Credit and Banking*, 28, 54-63.