Multivariate Fractional Cointegration and GDP Convergence*

Jonathan R. L. Halket†
NYU

August 8, 2005

Abstract

Tests of convergence in a time series framework are historically univariate and/or in an I(1)-versus-I(0) context. However, convergence hypotheses can and probably should be generalized to include multivariate, fractionally integrated processes. We propose several potential convergence definitions and use semi-parametric tests for bivariate and multivariate fractional cointegration in the G-7 countries. Our results indicate possible convergence as well as highlight why other procedures may fail to find convergence.
Keywords: output convergence, fractional integration, fractional cointegration, economic growth
JEL Classification: C32, F43, O49.

*This paper is submitted as part of the requirements for Professor Chen's Topics in Econometrics and Professor Sargent's Advanced Macroeconomics I. The author would like to thank C. Hurvich for much guidance and J. Bai and participants at the NYU Student Macro Lunch Seminar for helpful comments.
†Dept of Economics, 269 Mercer St, New York, NY 10003; jrhalket@nyu.edu

1 Introduction

There exists a substantial number of papers concerned with the empirical testing of convergence of per capita output across countries. The typical neoclassical growth model implies that per capita real incomes in different countries should converge, or at least should only differ by a mean-reverting amount. Endogenous growth models, on the other hand, permit a much wider variety of long-run output paths, including deterministically and/or stochastically diverging differences. The economic and policy implications of these models and therefore of convergence/divergence are large and varied.

Many of the initial time series tests for convergence are concerned only with univariate output differnces and consider them only in an I(1) vs. I(0) framework. A finding of converegence usually hinges on the rejection of the null of an I(1) output differential as according to a Dickey-Fuller-type statistic. Evidence in favor of convergence is infrequently found, particularly if some sort of structural break is not permitted.[1] Ericsson & Halket (2002) note that univariate tests typically have low power and show that a multivariate Johansen procedure may be able to detect convergence among some countries where a univariate (or even bivariate cointegration) approach had failed. They also note that multivariate tests are more likely to find convergence if convergence is a group phenomenon, as in Quah (1997)'s convergence clubs.

More recent research has noted that convergence is not an I(1) vs. I(0) phenomenon and, therefore, that tests for convergence should allow for fractional integration (denoted $I(d)$, with the order of integration, d, not necessarily an integer). Dickey-Fuller and Phillips-Perron type tests use as null and alternative hypothesies $d = 1$ and $d = 0$, respectively and thus may spuriously reject mean-reversion if the output differences are $I(d)$, with $0 < d < 1$, while multivariate Johansen procedures may find too much spurious cointegration.[2] Silverberg & Verspagen (1998), Juncal Cunado & de Gracia (2003*a*), and Beyaert (2003) test for mean-reversion in output differentials in a fractional integration framework and Michelacci & Zaffaroni (2000) proposes a thoretical, neoclassical growth model that could generate fractionally integrated and fractionally cointegrated per-capita outputs.

[1] For example, see Evans (1998) and Li & Papell (1999).

[2] For the theoretical and Monte Carlo evidence on univariate tests see Diebold & Rudebusch (1991) and Gonzalo & Lee (2000), respectively. For more on I(1) vs. I(0) multivariate cointegration tests, also see Gonzalo & Lee (2000).

We propose to test for fractional integration in per capita output and per capita output differentials using the procedure developed by Hurvich & Chen (2000). We then propose to test for bivariate cointegration without imposing a $(1, -1)$ cointegrating vector using a method outlined by Chen & Hurvich (2003a). Finally and more generally, we test for multivariate cointegration, following the procedures of Chen & Hurvich (2003c) and Chen & Hurvich (2003a).

This paper is organized as follows. Section 2 briefly rehashes the notion of fractional integration and cointegration. Section 3 discusses typical definitions of growth convergence. In section 4, we describe our data. In section 5, we describe our estimation procedures. We present our results in section 6. Section 7 concludes.

2 Fractional Integration and Cointegration

A process Y_t is said to be fractionally integrated of order d (hereafter $I(d)$) if its pth difference has spectral density

$$f(\lambda) \sim C|\lambda|^{-2(d-p)}, \lambda \to 0^+,$$

where $C > 0$ and p is a nonnegative integer such that $d - p < 0.5$. Alternatively, and perhaps more familiarily, Y_t is $I(d)$ if

$$[(1-L)^d Y_t = u_t,$$

where u_t is $I(0)$. A deterministic-trend-free $I(d)$ process where:

- $d \geq 1$ is non-mean-reverting and non-stationary,
- $1 > d \geq .5$ is mean-reverting and non-stationary,
- $.5 > d > -.5$ is stationary,
- $d \leq -.5$ is stationary but non-invertable.

Typically, an $I(0)$ process is called a short-memory process whereas an $I(d), d > 0$, process is said to have long-memory (but transitory if $d < 1$ and permanent otherwise).

Following Chen & Hurvich (2003c), suppose a process is a $(q \times 1)$ series such that its $(p-1)$th difference, $\{y_t\}$, is weakly stationary with common memory parameter $d \in (-p + 0.5, 0.5)$ and the integer $p \geq 1$. We say the original process is cointegrated with rank r, $(1 \leq r < q)$, and s cointegrating subspaces, $(1 \leq s \leq r)$, if:

$$y_t = A_0 x_t + A_1 u_t^1 + ... + A_s u_t^s, \tag{1}$$

where A_k $(0 \leq k \leq s)$ are $(q \times a_k)$ full-rank matrices with $a_0 = q - r$ and $a_1 + ... + a_s = r$, all columns of $A_0, ..., A_s$ are linearly independent, $\{x_t\}$ is an $(a_0 \times 1)$, $I(d)$ process, and $\{u_t^{(k)}\}$, $k = 1, ..., s$, are $(a_k \times 1)$, $I(d_{u_k})$ processes, with $-p - 0.5 < d_{u_s} < ... < d_{u_1} < d < 0.5$. It is worth noting that differencing the original series $p - 1$ times reduces any additive polynomial trend of $p - 1$ order or less to a constant in the differenced series $\{y_t\}$.

3 Searching for Growth Convergence

Across the many empirical searches for "convergence", a variety of definitions of convergence has been explicitly or implicitly used. For the purposes of this paper, we sketch three general types of convergence definitions using the $I(1)$ vs. $I(0)$ framework[3]. Let $\{Y_t^i\}$ and $\{Y_t^j\}$ be the log of real per capita output of country i and country j, respectively, and are assumed $I(1)$.

1. *Stochastic Convergence* (e.g. (Carlino & Mills 1993)): $\{Y_t^i\}$ and $\{Y_t^j\}$ are said to be stochastically converging if

$$Y_t^i - Y_t^j = c + \delta t + u_{ij,t}, \quad u_{ij,t} \sim I(0). \tag{2}$$

If δ is such that the absolute per capita output difference tends to get smaller over time, then stochastic convergence has also been called *catching-up convergence* (Bernard & Durlauf 1996). If δ is of the opposite sign, then the countries are deterministically diverging.

2. *Deterministic Convergence* (e.g. (Li & Papell 1999)): $\{Y_t^i\}$ and $\{Y_t^j\}$ are said to be deterministically converging if

$$Y_t^i - Y_t^j = c + u_{ij,t}, \quad u_{ij,t} \sim I(0). \tag{3}$$

Deterministic convergence implies stochastic convergence, but not vice versa. For both deterministic and stochastic convergence, if $|c| > 0$, then the countries are *conditionally converging.*

[3]These definitions follow Ericsson & Halket (2002)'s

3. *Zero-mean Convergence* (e.g. (Bernard & Durlauf 1996)): $\{Y_t^i\}$ and $\{Y_t^j\}$ are said to be unconditionally or absolutly converging or converging to a zero-mean if

$$Y_t^i - Y_t^j = u_{ij,t}, \quad u_{ij,t} \sim I(0). \tag{4}$$

Many of the studies that test for convergence using one or more of the above definitions do so using univariate or bivariate tests. Univariate tests are usually an augmented Dickey-Fuller (ADF) test or some variant thereof (see Li & Papell (1999) for an example of the former and Evans (1998) and Cheung & Pascual (2004) for panel variation examples). ADF tests assume $I(1)$ residuals as a null hypothesis and $I(0)$ as an alternative, and thus have reduced power to find convergence in a fractional integration framework. Bivariate tests for convergence like Bernard & Durlauf (1995) and Ericsson & Halket (2002) test for cointegration with a $(1, -1)$ cointegrating vector and typically (for example, if using the Johansen procedure) take $I(0)$ and $I(1)$ as null and alternative hypotheses, respectively.

A pairwise procedure, in practice, usually assumes a base-country (that is, a fixed j for Y_t^j) either implicity or explicitly. An assumption of this sort typically rules out the possibility of detecting convergence clubs. Furthermore, studies like Bernard & Durlauf (1995), Oxley & Greaseley (1995) and Cellini & Scorcu (2000) find few cases of convergence if a structual break (in the trend or mean) is not allowed. However, using the above sorts of definitions of convergence, it is not immediately clear what the implications of a stuctural break are: How many breaks can "reasonably" be permitted between countries that are otherwise considered converging? Additionally, it is not clear to us that a trend in the cointegrating residual provides any evidence of convergence, even if it has the "proper" sign. Cointegrating relationships are long-run relationships so deterministic convergence today means deterministic divergence in the future, if we are to take a cointegrating process with a trend seriously.[4]

Some recent studies (e.g. (Beyaert 2003)) have relaxed the definition of convergence to account for the possibility of fractionally cointegrated pairs of

[4]Empirical convergence studies like this paper and many of those cited herein lack a formal theoretical model and may suffer from the "measurement without theory" critique of Koopmans (1947) and Sargent & Sims (1977). Interestingly, like Sargent & Sims (1977) do in their investigation of business cycles, we end up using cross-spectral techniques so as to avoid using econometric models that we do not take seriously.

countries (or fractionally integrated differences in log per capita output). By allowing for fractional integration and cointegration, we are able to seperately identify mean-reversion and stationarity. We use the following defintions of convergence for a bivariate setting: Assume Y_t^i and Y_t^j are both $I(d), d \geq 1$. If Y_t^i and Y_t^j are cointegrated with a $(1, -1)$ cointegrating vector and the cointegrating residual u_t^{ij} is $I(d_{ij}), d_{ij} < 1$ then we say countries i and j are converging. Otherwise we say they are diverging. If the countries are converging and $\mathrm{E}[u_t^{ij}] = 0$, we say there is *zero-mean convergence*. Moreover, the lower d_{ij} is the stronger (or perhaps faster) the convergence. A fractional cointegrating approach provides a framework where many of the above issues are avoided or at least mitigated. Our procedure for estimating the order of integration d is robust to the possible presence of low-order polynomial trends and structural breaks in the mean and trend parameters (c and δ in the above notation). As noted above, however, a bivariate approach using one base country is not appropriate if the countries are in fact in convergence clubs. In practice, one can test all possible pairings of countries, but then a contradictory conclusion is possible[5]. A multivariate approach can detect convergence clubs. A seemingly natural extension of the bivariate fractional integration convergence defintions to a multivariate framework is to examine the cointegrating residuals u_t^k from (1) for mean reversion. However, it is possible to have a mulitvariate cointegrating system which presents possibly contradictory conclusions. In practice then, determinations of convergence and divergence should be made carefully.

4 Data

Our data is annual data from the G-7 countries[6] from 1889-1994 from Maddison (1995) and Maddison (2001)[7]. Figure 1 presents the logarithm of real per capita GDP for all seven countries.

As can be seen from figure 1, there is a substantial and different "disruption" to per capita real GDP for each of the countries during the WWII years. We choose to view these disruptions as part of the underlying process

[5]For instance, it is possible that, in sample, testing concludes that countries i and j and countries j and k are pairwise converging but countries i and k are not.

[6]Canada (CA), Germany (DE), France (FR), Italy (IT), Japan (JP), the United Kingdom (UK), and the United States (US)

[7]with some corrections for minor and obvious typographic errors

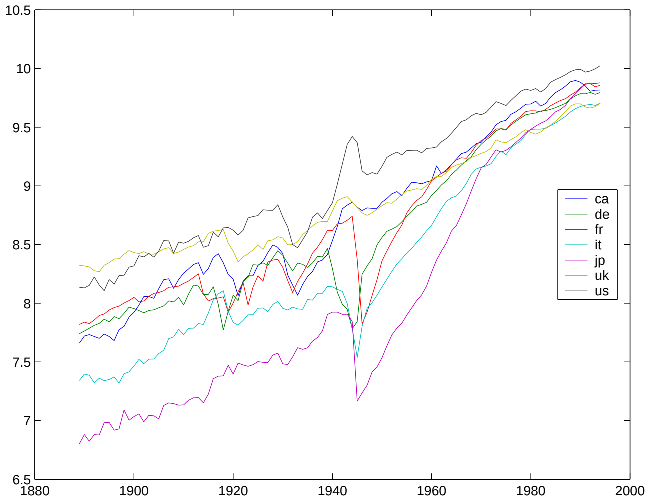

Figure 1: Log real gdp per capita

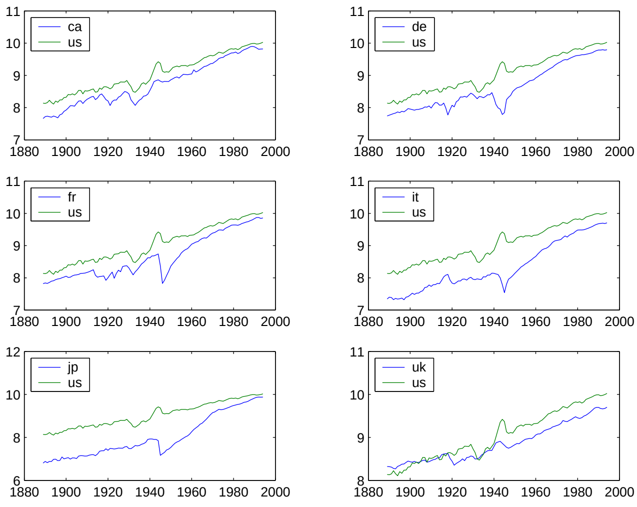

Figure 2: Log real gdp per capita of each country vs. US

both because our methodology is built to account well for these kinds of disturbances and because our sample would be unsuitably short if we only used post-WWII data. Since most work, including our own, uses, at least in part, the US as a base country for pairwise convergence studies, we also include figure 2.

5 Methodology

In this section, we present our methodology for determining the order of integration of a series, for finding fractional cointegration in a bivariate setting and then in a multivariate setting, and, finally, for finding the presence of a deterministic trend in a fractionally integrated series.

5.1 Fractional Integration

Following Hurvich & Chen (2000) and Chen & Hurvich (2003*b*), the discrete Fourier transform of a vector sequence of data $\{\xi_t\}_{t=1}^{n}$ is:

$$w_{\xi,j} = \frac{1}{(2\pi n)^{\frac{1}{2}}} \sum_{t=1}^{n} \xi_t e^{i\lambda_j t}, \quad j = 1, ..., \tilde{n}, \tag{5}$$

where $\tilde{n} = [\frac{n-1}{2}]$ and $\lambda_j = \frac{2\pi j}{n}$. We note that $w_{\xi,j}$ remains unchanged if we replace $\{\xi_t\}$ by $\{\xi_t + C\}$ because $\sum_{t=1}^{n} e^{i\lambda_j t} = 0$. Then, the cross-periodogram $I_{\zeta\xi,j} = w_{\zeta,j}\overline{w_{\xi,j}}$, where $\overline{w_{\xi,j}}$ is the complex conjugate of $w_{\xi,j}$.

Next, we define the taper[8] $h_t = \frac{1}{2}(1 - e^{\frac{i2\pi(t-\frac{1}{2})}{n}})$ and the tapered transform of $\{\xi_t\}_{t=1}^{n}$

$$w_j^T = \frac{1}{(2\pi \sum (|h_t|^2))^{\frac{1}{2}}} \sum_{t=1}^{n} h_t y_t e^{i\lambda_j t}. \tag{6}$$

Tapering the data helps reduce the periodogram bias (called leakage) caused by strong peaks and troughs in the spectral density. A taper also mitigates the potential loss of information from overdifferencing and is helpful for estimating the mean of a potentially overdifferenced series[9]. The cost of tapering

[8]This taper is adapted from Hurvich & Chen (2000) and the tapered transform is for series that need to be once differenced to stationarity. For the general version of this taper see their paper.

[9]see Deo & Hurvich (1998) for further discussion

is a loss of efficiency in the estimation of the order of integration[10], d. Hurvich & Chen (2000) find the above defined taper to be most efficent among a class of tapers and safely usable on series with a true order of integration $d \in (-1.5, 0.5)$. If we note that $w^T_{\xi,j} = \sqrt{2}\{0.5w_{\xi,j} - 0.5w_{\xi,j+1}e^{-i\frac{\pi}{n}}\}$, it follows that the tapered cross-periodogram $I^T_{\zeta\xi,j} = w^T_{\zeta,j}\overline{w^T_{\xi,j}}$ is an estimator of the spectrum $f(\lambda_{\zeta\xi,\tilde{j}})$, where $\tilde{j} = j + \frac{1}{2}$.

Beyaert (2003) uses the same data but does not use a taper. Tapers are usually bell-shaped and, in the time domain, they accentuate the middle part of the series at the expense of the two ends. For our data, this middle includes the WWII years. Beyaert does not want to emphasize this part of the data and uses the less robust method of Marmol & Velasco (2002)[11]. Nevertheless, we maintain that the effects of a taper are best viewed in the frequency domain, where they remain substainially ameliorative. Figures 4 and 5 help illustrate this.

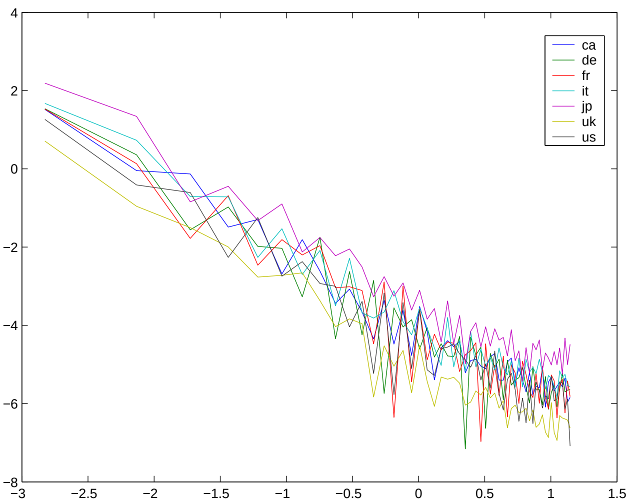

Figure 3: Log periodogram vs Log frequency

This is the log of the periodogram of log real GDP per capita y^i_t at the Fourier frequencies vs. the log of the Fourier frequencies.

---

[10] This loss of efficiency becomes more prominent if $d \approx 0$.

[11] In order to avoid using a taper, Beyaert chooses a method which does not require differencing the data. However, this method is only appropriate for $d \in (-0.5, 1.5)$. As we find below, this is somewhat too close for confort.

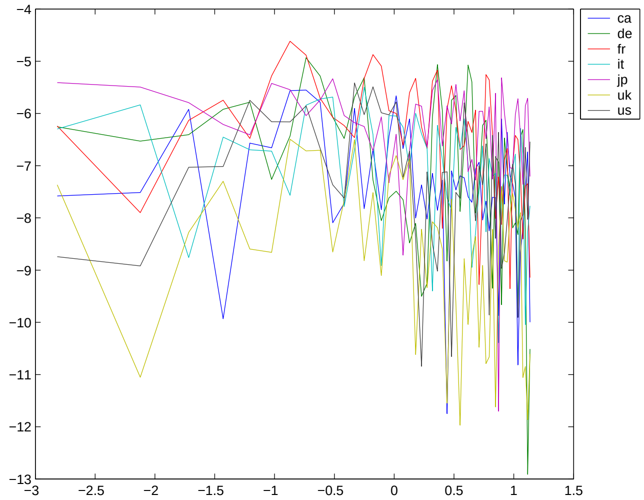

Figure 4: Log vs Log: Δy_t

Same as figure 3, except using Δy_t^i

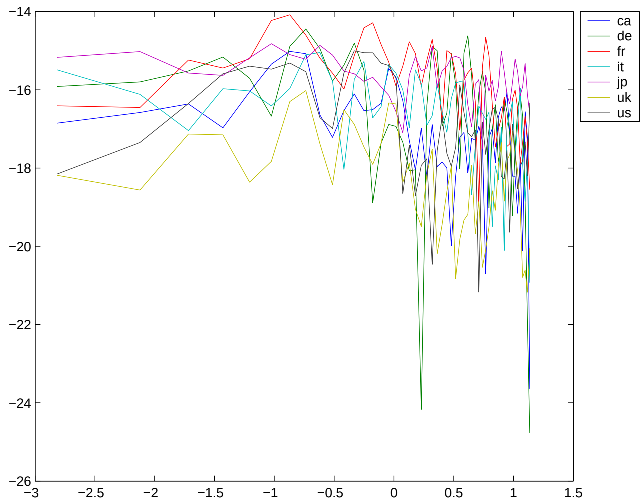

Figure 5: Log vs Log: $\Delta {y_t}^T$

Same as figure 3, except using the difference of the tapered series, $\Delta y_t^{i,T}$

To estimate d, we use the tapered Gaussian semi-parametric estimator ($\hat{d}_{GSET}$) of Hurvich & Chen (2000):

$$\hat{d}_{GSET} = \arg \min_{d \in (-1.5,.5)} \{\log\left(\frac{1}{m}\sum_{j=1}^{m} \lambda_{\tilde{j}}^{2d} I_j^T\right) - \frac{2d}{m}\sum_{j=1}^{m} \log \lambda_{\tilde{j}}\}, \tag{7}$$

where[12] $m \approx n^{0.8}$. $\hat{d}_{GSET}$ of a series with $d \in (-1.5, 0.5)$ and constant mean and using the above taper is consistent and asymptotically normal with asymptotic variance $\frac{1.5}{4m}$. The GSET estimator also provides a pseudo-estimate of d when $d \geq 0.5$. This is important in practice (see below). Comparitively, the untapered Gaussian semi-parametic estimate has an asymptotic variance of $\frac{1}{4m}$ but is biased towards zero. A small sample correction for the variance (also from Hurvich & Chen (2000)) is

$$\text{Var}(\hat{d}_{GSET}) \approx 1.5/(4\sum_{j=1}^{m} \tilde{\nu}_j^2), \tag{8}$$

where

$$\tilde{\nu}_j^2 = \log(2\sin(\frac{\lambda_{\tilde{j}}}{2})) - m^{-1}\sum_{j=1}^{m} \log(2\sin(\frac{\lambda_{\tilde{j}}}{2})).$$

It is worth noting that a good eyeballing method is to examine the slope of the plot of $\log(I_{\tilde{j}})$ vs. $\log(\lambda_{\tilde{j}})$ at low frequencies. The steeper is the downward slope, the higher is d. This also provides good intuition as to why leakage causes an untapered estimate of d to be biased towards zero: leakage leads to side lobes near sharp peaks in the data, which reduces the slope of such a plot. A flatter slope implies a $\hat{d}$ closer to zero. In practice, we:

1. Taper the series and then get estimate (or pseudo-estimate) $\hat{d}_{GSET}$.
2. If $\hat{d}_{GSET} \approx> 0.5$, difference the untapered series and then repeat step 1.
3. $\hat{d}_{GSET}$ of the original series $= \hat{d}_{GSET,p} + p$, where $\hat{d}_{GSET,p}$ is the GSET estimate of the p^{th} times differenced series.

Since $\hat{d}_{GSET}$ is invariant to the addition of a constant to the series, differencing the series p times makes $\hat{d}_{GSET}$ invariant to p^{th}-order trends. Differencing the untapered series avoids confounding the trend.

[12]For larger samples, $n^{0.7}$ or less may more appropriate. This sample is relatively small though.

5.2 Fractional Bivariate Cointegration

Following Chen & Hurvich (2003a), define the averaged tapered periodogram of two processes $\{x_t\}_{t=1}^n = \{\Delta X_t\}_{t=0}^n$ and $\{y_t\}_{t=1}^n = \{\Delta Y_t\}_{t=0}^n$, with $d_X = d_Y \in (0.5, 1.5)$, as[13]

$$\hat{F}_{xy}^T(m_o) = \frac{2\pi}{m_o} \sum_{j=1}^{m_o} Re\{I_{xy,j}^T\}, \quad 1 \leq m_o < \frac{n}{2}, \tag{9}$$

where m_o is a bandwidth parameter. Since we are using a taper, m_o can be held fixed. Otherwise, $m_o \to \infty$. Suppose the true cointegrating relationship is $(-\beta_o, 1)$ such that $y_t - \beta_o x_t = u_t$ and correspondingly $Y_t - \beta_o X_t = U_t$ with $d_u < d_x$. Then an estimate of β_o is

$$\hat{\beta}_{m_o} = \frac{\hat{F}_{xy}^T(m_o)}{\hat{F}_{xx}^T(m_o)}. \tag{10}$$

$\hat{\beta}_{m_o}$ is $n^{d_x - d_u}$ consistent; it is always at least as fast as OLS of y on x. In practice, since d_x and d_u must be estimated, standard errors for $\hat{\beta}_{m_o}$ are not computed.

Following most of the literature, we choose to estimate bivariate cointegration, first using the US as a base country (that is, all other counrtries vs. US, pairwise). We estimate the cointegration residual $\hat{d}_{US,i}$ twice: using an unrestricted (that is, estimated) $\hat{\beta}_{m_o,US,i}$, and restricting $\hat{\beta}_{m_o,US,i} = 1$, as is often done. We then repeat the exercise using DE as a base country, in order to search for potential convergence clubs.

5.3 Multivariate Fractional Cointegration

Following Chen & Hurvich (2003c) and Chen & Hurvich (2003b), let[14] $\{y_t\}_{t=1}^n = \{\Delta Y_t\}_{t=0}^n$, where Y_t is the $(q \times 1)$ vector of the q countries' log real GDP per capita at time t, each of which is $I(d_y)$. To examine the validity of this restriction, we first compute $\hat{d}_{res}$, the restricted version of $\hat{d}_{GSET}$ adapted from Lobato (1999) and Chen & Hurvich (2003b):

$$\hat{d}_{res} = \arg \min_{d \in [-1.5, 0.5]} \{\log |\frac{1}{m} \sum_{j=1}^{m} I_{yy,\tilde{j}}^T \tilde{j}^{2d}| - \frac{2qd}{m} \sum_{j=1}^{m} \log \tilde{j}\}. \tag{11}$$

[13] Hereafter we assume that the original log real GDP per capita series have at most a linear trend.

[14] Appologies for recycling notation.

Typically, we could follow with a Hausman-type test for equality of orders of integration. Lobato (1999) and Chen & Hurvich (2003*b*) have derived distributions for this test. However, these distributions do not account for the possibility of cointegration under the null and thus are not valid for our purposes. We leave this for future examination. Nevertheless, $\hat{d}_{res}$ remains a good benchmark for eyeballing.

Let $\hat{F}^T_{yy}(m_o)$ be the averaged tapered cross-periodogram[15]. Let λ_i and $\hat{\beta}_i$ be the eigenvalues (ordered from smallest to largest) and their corresponding eigenvectors of $\hat{F}^T_{yy}(m_o)$, respectively. Chen & Hurvich (2003*b*) note that the r smallest eigenvalues correspond to the r strongest cointegrating relationships. One way to determine r is:

$$\hat{r} = \arg\min_{u<q}\{V(n)(q-u) - \hat{\sigma}_{u+1,q}\}, \tag{12}$$

where $\{V(n)\}$ some deterministic sequence such that

$$\frac{n^{d_u+d}}{V(n)} + \frac{V(n)}{n^{2d}} \to 0, \quad \hat{\sigma}_{j,q} = \sum_{i=j}^{q} n^{2d}\lambda_i.$$

In practice, the choice of $\{V(n)\}$ often allows the user to "choose" $\hat{r}$, so it is easier to just eyeball the eigenvalues, looking for a large jump. The eigenvectors corresponding to each of the eigenvalues correspond, with high probability, to the space spanned by the cointegrating vector:

$$\sin(\Theta) = O_p(n^{d_u-d_y}), \tag{13}$$

where $\sin(\Theta)$ is the square-root of the sum of the squared lengths of the residuals from the orthogonal projection of the r 'smallest' eigenvectors onto the cointegrating vectors.

5.4 Estimation of a Trend

As stated above, we assume that there is at most a first order trend in the log real GDP per capita series. Furthermore, we would like to test the cointegrating residuals for first order trends, as the presence of a trend can muddle notions of convergence (and even point to a deterministic divergence). Three methods by which we could estimate trends in the cointegrating residuals $u_{i,t}$ are:

[15] $m_o > q + 3$

1. OLS of $u_{i,t}$ on a trend term,
2. the mean of $\{\Delta u_{i,t}\}$ and
3. the mean of $\{\Delta u_{i,t}^T\}$, the differenced, tapered residuals.

For residuals with $d_{u_i} \leq 0.5$, the mean of the untapered, differenced series is inappropriate, but for $1.5 > d_{u_i} \gtrapprox 0.7$, it is at least as efficient, if not more[16]. We thus use the mean of the untapered, differenced series.

6 Results

6.1 Integration Results

Table 1 presents results on the estimation of the trend and the order of integration for each country's GDP[17] series as well as the order of integration when all series are restricted to have the same order. No country has a significant trend. There is a noticeable downward bias in $\hat{d}_{\Delta gdp}$ and CA, UK and US are each significantly more than $I(1)$, as is $\hat{d}_{res}$. Output growth has long memory.

6.2 Bivariate Cointegration Results

Tables 2 and 3 present the unrestricted and restricted bivariate cointegration results using US as a base country, respectively. For the unrestricted results, the cointegrating vector is $(-\hat{\beta}_{11}, 1)$ for the series (GDP_{US}, GDP_i). Except for perhaps CA and almost certainly in the case of DE, the vector is different from $(-1, 1)$. Unrestricted, only the CA cointegrating residual is mean-reverting, though no residual is significantly different from 1. Restricting the cointegrating vector raises all but CA's residual; not surprising given that all other vectors were quite different from $(-1, 1)$. Given the restricted bivariate results, it is not surprising that most previous bivariate studies have rejected convergence, despite the fact that US seems to be cointegrated with both CA and UK, at least[18].

[16]For further discussion and the precise variance formulas see Deo & Hurvich (1998), though they examine a different taper then the one above.

[17]Henceforth GDP refers to the log of real per capita GDP.

[18]a reminder that cointegration need not imply mean-reversion

Table 1: Degree of Integration

	CA	DE	FR	IT	JP	UK	US
$\widehat{trend}_{\Delta gdp}$	0.0205	0.0196	0.0194	0.0225	0.0293	0.0132	0.0180
$\sigma_{\widehat{trend}_{\Delta gdp}}$	(.0961)	(.0682)	(.0814)	(.0628)	(.0733)	(.0662)	(.1057)
$\hat{d}_{gdp}$	0.9502	0.9430	0.9580	0.9388	0.9422	0.9509	0.9504
$\hat{d}_{\Delta gdp}$	1.2034	1.0888	1.1018	1.1702	1.0825	1.2867	1.1201
$\hat{d}^{T}_{\Delta gdp}$	1.3398	1.1508	1.1577	1.2124	1.1138	1.3847	1.3368

T denotes tapered series
All statistics refer to the order of integration of the orginal level series.
$\sigma_{\hat{d}^{T}_{\Delta gdp}} = 0.1296$, adjusted for small sample
$\hat{d}_{res} = 1.2765$

Tables 4 and 5 present the same using DE as a base country. There is no evidence of cointegration in the unrestricted case and slight evidence of cointegration with FR in the restricted case (though of the non-mean-reverting variety).

Table 2: Unrestricted Estimation vs US

(m=11)	CA	DE	FR	IT	JP	UK
$\hat{\beta}_{11}$	0.7928	-0.7776	0.5679	-0.2186	0.2100	0.3263
d_{u_i}	0.9043	1.0348	1.1420	1.2084	1.1124	1.1190
σ_{d_u}	0.1296					

Table 3: Restricted Estimation vs US

(m=11)	CA	DE	FR	IT	JP	UK
d_{u_i}	0.8746	1.2992	1.1568	1.2718	1.1475	1.1280
σ_{d_u}	0.1296					

Table 4: Unrestricted Estimation vs DE

(m=11)	CA	FR	IT	JP	UK
$\hat{\beta}_{11}$	-0.1729	0.0266	0.4466	0.0682	-0.1376
d_{u_i}	1.2639	1.1500	1.1423	1.1168	1.2889
σ_{d_u}	0.1296				

Table 5: Restricted Estimation vs DE

(m=11)	CA	FR	IT	JP	UK
d_{u_i}	1.3102	1.0090	1.0782	1.1678	1.2314
σ_{d_u}	0.1296				

6.3 Multivariate Cointegration Results

Table 6 presents the full multivariate cointegration results. The smallest eigenvalue is over 100 times smaller than the largest; evidence of some cointegration. The cointegrating vectors are naturally difficult to make sense of. Some elements in some of the vectors are probably insignificantly different from zero (for instance, DE and FR in the second vector). Even after some sensible simplifications, the vectors are still hard to make economic sense of: many different stories could probably be told so we will not try. Nevertheless, several of the cointegrating residuals are significantly reduced and are mean-reverting - potential evidence for convergence - though all are non-stationary. Furthermore, from table 7, we see that each residual still contains no significant trend[19].

Table 8 presents the multivariate cointegration results using only the Euro area countries. There is some evidence of cointegration in this sub-group. Though we cannot reject permanent memory for any residual, the strongest cointegrating relationship may be mean reverting. This is a potential convergence club within the larger group.

[19] Even though $\hat{d}_{u_1} > \hat{d}_{u_2}$, they are not significantly different

Table 6: Mulitvariate Cointegration Results

eigval	$4.5\text{x}10^{-9}$	$4.9\text{x}10^{-9}$	$2.0\text{x}10^{-8}$	$4.3\text{x}10^{-8}$	$1.3\text{x}10^{-7}$	$3.6\text{x}10^{-7}$	$6.2\text{x}10^{-7}$
Corresponding Eigenvectors							
CA	-0.0532	-0.6178	0.0026	0.1950	0.6642	0.3340	0.1574
DE	0.2894	0.0461	0.1925	-0.3436	0.5077	-0.7057	0.0565
FR	-0.1686	0.0026	-0.4826	-0.5348	-0.0492	0.0778	0.6664
IT	-0.5018	0.3505	-0.0884	0.5964	0.1962	-0.3291	0.3391
JP	0.2647	-0.2578	0.5537	0.1943	-0.3821	-0.0782	0.6058
UK	0.7489	0.3023	-0.4054	0.3433	0.1267	0.1532	0.1616
US	-0.0484	0.5792	0.5012	-0.2158	0.3134	0.4964	0.1404
Integration order of cointegrating residual							
$d_{\Delta u^T}$	0.6991	0.6271	0.7570	0.8687	1.1207	1.2902	1.3366
$\sigma_{d_{\Delta u^T}}$	0.1296						

Table 7: Trend Results

trend	-0.0067	-0.0030	-0.0124	-0.0066	-0.0232	0.0042	-0.0473
	(.0116)	(.0112)	(.0231)	(.0283)	(.0555)	(.1205)	(.1822)
mean	4.6993	3.7921	2.1484	1.7380	12.2810	-0.0099	18.0291

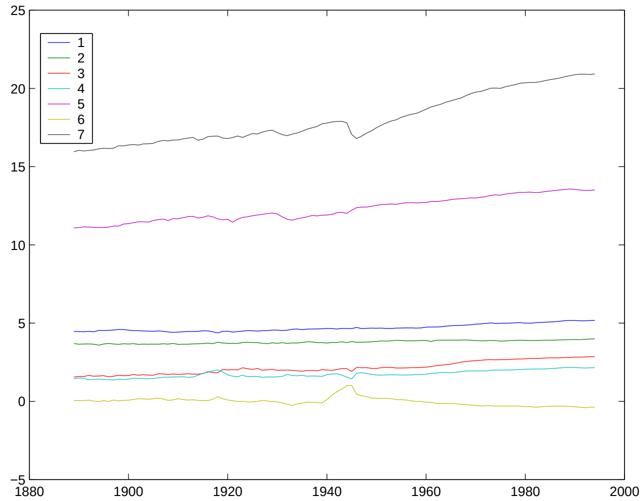

Figure 6: Multivariate Cointegrating Residuals vs. Time

Smallest eigenvalue = 1, largest eigenvalue = 7

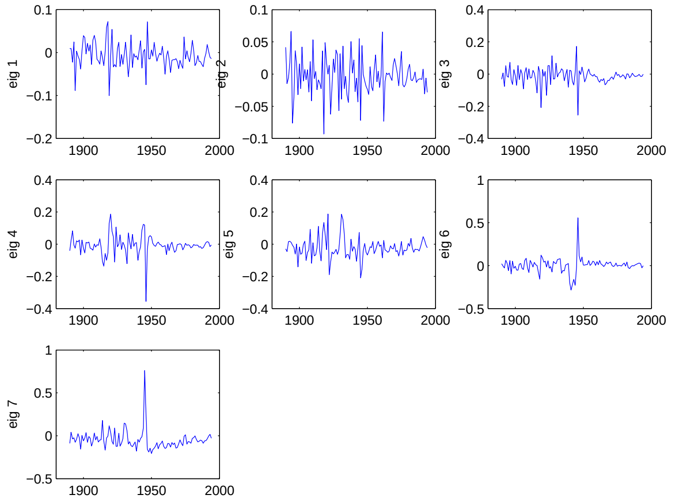

Figure 7: Differenced Cointegrating Residuals vs. Time

Same as figure 6.

Table 8: Euro Area Mulitvariate Cointegration Results

eigval	$3.2\text{x}10^{-8}$	$2.4\text{x}10^{-7}$	$3.8\text{x}10^{-7}$
Corresponding Eigenvectors			
DE	-0.4219	0.8293	0.3664
FR	-0.3604	-0.5242	0.7716
IT	0.8319	0.1935	0.5200
Integration order of cointegrating residual			
$d_{\Delta u^T}$	0.8663	1.0057	1.4560
$\sigma_{d_{\Delta u^T}}$	0.1296		

7 Conclusions

This paper has examined convergence among the G-7 countries using a bivariate and then multivariate framework. Traditional $I(1)$ vs. $I(0)$ time series tests for convergence are biased towards convergence or non-convergence, depending on the framework. A fractional integration framework seperates long memory and non-stationarity properties from mean-reversion. In some respects, a bivariate cointegration analysis is more natural and easier to interpret, but may fail to pick up convergence in the presence of convergence clubs. Multivariate cointegrating frameworks are robust to convergence clubs but results can be hard to interpret. In the case of the G-7, we find little evidence for convergence in the bivariate framework. Multiple multivariate cointegrating relationships exist, including some with non-stationary, mean-reverting residuals. However, the cointegrating vectors do not lend a clear convergence interpretation.

Convergence clubs have traditionally been found in data which include both developing and developed nations, so the multivariate framework may be more enlightening if more countries are included in the analysis. Our multivariate procedure is data intensive, though, and, to our knowledge, there currently exists insufficient data sets to pursue this sort of analysis. Furthermore, more work on the various asymptotic distributions of parameters used in fractional cointegration needs to be done.

References

Bernard, A. B. & Durlauf, S. N. (1995), 'Convergence in 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**, 161–173.

Beyaert, A. (2003), 'Fractional integration and output convergence: An application to the G-7 countries', *mimeo* .

Carlino, G. A. & Mills, L. O. (1993), 'Are u.s. regional incomes converging?', *Journal of Monetary Economics* **32**, 335–346.

Cellini, R. & Scorcu, A. E. (2000), 'Segmented stochastic convergence across the G-7', *Empirical Economics* **25**, 463–474.

Chen, W. W. & Hurvich, C. M. (2003*a*), 'Estimating fractional cointegration in the presence of polynomial trends', *Journal of Econometrics* **117**, 95–121.

Chen, W. W. & Hurvich, C. M. (2003*b*), 'Semiparametric estimate of fractional cointegration subspaces', *mimeo* .

Chen, W. W. & Hurvich, C. M. (2003*c*), 'Semiparametric estimate of multivariate fractional cointegration', *Journal of the American Statistical Association* **98**(463), 629–642.

Cheung, Y.-W. & Pascual, A. G. (2004), 'Testing for output convergence: A re-examination', *Oxford Economic Papers* **56**, 45–63.

Deo, R. S. & Hurvich, C. M. (1998), 'Linear trend with fractionally integrated errors', *Journal of Time Series Analysis* **19**(4), 379–397.

Diebold, F. X. & Rudebusch, G. D. (1991), 'On the power of Dickey-Fuller tests against fractional alternatives', *Economics Letters* **35**, 155–160.

Ericsson, N. R. & Halket, J. R. (2002), 'Convergence of ouptut in the G-7 countries', *mimeo* .

Evans, P. (1998), 'Using panel data to evaluate growth theories', *International Economic Review* **39**(2), 295–306.

Gonzalo, J. & Lee, T.-H. (2000), 'On the robustness of cointegration tests when series are fractionally integrated', *Journal of Applied Statistics* **27**(7), 821–827.

Hurvich, C. M. & Chen, W. W. (2000), 'An efficient taper for potentially overdifferenced long-memory time series', *Journal of Time Series Analysis* **21**(2), 155–180.

Juncal Cunado, L. A. G.-A. & de Gracia, F. P. (2003*a*), 'Additional empirical evidence on real convergence: A fractionally integrated approach', *working paper* .

Juncal Cunado, L. A. G.-A. & de Gracia, F. P. (2003*b*), 'Empirical evidence on real convergence in some OECD countries', *Applied Economics Letters* **10**, 173–176.

Koopmans, T. C. (1947), 'Measurement without theory', *Review of Economics and Statistics* .

Li, Q. & Papell, D. (1999), 'Convergence of international output time series evidence for 16 OECD countries', *International Review of Economics and Finance* **8**(3), 267–280.

Lobato, I. N. (1999), 'A semiparametric two-step estimator in a multivariate long memory model', *Journal of Econometrics* **90**, 129–153.

Maddison, A. (1995), *Monitoring the World Economy: 1820-1992*, Organisation for Economic Co-operation and Development, Paris.

Maddison, A. (2001), *The World Economy: A Millenial Perspective*, Organisation for Economic Co-operation and Development, Paris.

Marmol, F. & Velasco, C. (2002), 'Trend stationarity versus long-range dependence in time series analysis', *Journal of Econometrics* **108**, 25–42.

Michelacci, C. & Zaffaroni, P. (2000), '(fractional) beta convergence', *Journal of Monetary Economics* **45**, 129–153.

Oxley, L. & Greaseley, D. (1995), 'A time-series perspective on convergence: Australia, uk and usa since 1870', *The Economic Record* **71**(214), 259–270.

Quah, D. T. (1997), 'Empirics for growth and distribution: Stratification, polarization, and convergence clubs', *Journal of Economic Growth* **2**, 27–59.

Sargent, T. J. & Sims, C. A. (1977), Business cycle modeling without pretending to have too much a priori economic theory, Federal Reserve Bank of Minneapolis.

Silverberg, G. & Verspagen, B. (1998), Long memory in time series of economic growth and convergence. paper prepared for Session II-3: Applied Dynamic Modeling, of the Austrian Focus Sessions of the 10th SASE Conference, Vienna, Austria, July 13-16, 1998.