CENTRE FOR ECONOMETRIC ANALYSIS
CEA@Cass

http://www.cass.city.ac.uk/cea/index.html

Cass Business School
Faculty of Finance
106 Bunhill Row
London EC1Y 8TZ

A Note on Common Cycles, Common Trends and Convergence

Vasco Carvalho, Andrew Harvey and Thomas Trimbur

CEA@Cass Working Paper Series

WP–CEA–04-2006

A Note on Common Cycles, Common Trends and Convergence

Vasco Carvalho
Department of Economics, University of Chicago, Chicago, IL 60637
Andrew Harvey
Faculty of Economics, University of Cambridge,
Cambridge CB3 9DD, England
Thomas Trimbur
Division of Research and Statistics, Federal Reserve System,
Washington DC 20551

April 21, 2006

Abstract

The aim of this article is to compare and contrast structural time series models and the common features methodology. The way in which trends are handled is highlighted by describing a recent structural time series model that allows convergence to a common growth path. Post-sample data is used to test its forecasting performance for income per head in US regions. A test for common cycles is proposed, its asymptotic distribution is given and small-sample properties are studied by Monte Carlo experiments. Applications are presented, with special attention being paid to the implications of using higher-order cycles.

KEYWORDS: Common features; balanced growth; error correction mechanism; stochastic trend; unobserved components.

1 INTRODUCTION

In their investigation of relationships between stationary components, referring to such co-movements in economic series as ‘common feature cycles’, Engle and Kozicki (1993) motivate further investigation of such relationships. In particular, they make the following remark (p. 376): ‘Often, however, the interesting forms of comovement are stationary; common shocks that are less persistent than unit roots may be the most important in understanding business cycles’. They define a common feature as being present if ‘...there exists a nonzero linear combination of the series that does not have this feature’ (p 370). Vahid and Engle (1993) adapt the approach to nonstationary data by making use of multivariate Beveridge-Nelson decompositions in a preliminary step. This amounts to decomposing a series into two parts, trend and cycle, where the latter refers to ‘the stationary remainder after subtracting the random walk trend’ (footnote 1, p 341). The view of the world is therefore one in which series are integrated of order one, denoted $I(1)$, and components or features are extracted by one-sided filters. This contrasts with the approach based on structural time series models (STMs) in which trends may be I(2), filters are two-sided (except at the ends) and components satisfying the definition of common features can be constructed as special cases. As is well known, the cycles obtained in real business cycle economics by using the Hodrick-Prescott filter can be interpreted in terms of an STM with I(2) trend components.

The aim of this article is to compare and contrast the structural time series and common features methodologies as they apply to trends and cycles. Common trends are a prominent feature of many multivariate time series and their relationship to co-integration is well-known. However, the series of interest are sometimes in the process of converging. They do not then display co-integration at the outset and co-integration tests are misleading. Carvalho and Harvey (2005) recently proposed a model that facilitates the correct interpretation of long-run movements and allows eventual convergence to a balanced growth path. Here we take the opportunity to test the forecasting performance of the model fitted by Carvalho and Harvey to per capita income in US regions by using genuine post-sample data. A comparison is made with forecasts obtained with vector autoregressions (VARs) and vector error correction models (VECMs). In investigating forecasting performance we pay particular attention to whether the second-order error correction mechanism that lies at the heart of the multivariate convergence model has been able to capture the medium term behavior.

The similar cycle model used in STMs to capture correlations between cycles in different series contains common cycles as a special case. In a bivariate model this observation enables us to develop a test for the null hypothesis of a common cycle against the alternative of similar cycles. The asymptotic distribution is non-standard because a parameter lies on a boundary when cycles are in common. The small sample properties of the test are studied by Monte Carlo experiments and it is applied when a bivariate STM, with generalized higher-order cycles of the kind introduced by Harvey and Trimbur (2003), is fitted to data on GDP in Canada and the US. We also examine the evidence for common cycles in real per capita income in some of the US regions and make some comparisons with the studies by Vahid and Engle (1993) and Carlino and Sill (2001).

The plan of the article is as follows. Section 2 reviews multivariate STMs and shows how they can provide a description of trends and cycles. Section 3 sets out the second-order error correction convergence mechanism and looks at its forecasting performance. The multivariate model is described and the way in which balanced growth can be incorporated into a VECM is reviewed. Forecasting comparisons are then given. Section 4 sets out the common cycle test and gives applications. Section 5 concludes.

Estimation for the higher-order cycles and convergence models was carried out with programs written in the Ox language of Doornik (1999), with use being made of the SsfPack library of functions of Koopman *et. al* (1999). Some of the standard models with first order cycles were estimated using the STAMP package of Koopman *et al* (2000).

2 MULTIVARIATE STRUCTURAL TIME SERIES MODELS

Suppose we have N time series. Define the vector $\mathbf{y}_t = (y_{1t}, .., y_{Nt})'$ and similarly for the unobserved components representing the trend, $\boldsymbol{\mu}_t$, cycle, $\boldsymbol{\psi}_t$, and irregular, $\boldsymbol{\varepsilon}_t$. Then a multivariate structural time series model may be set up as

$$\mathbf{y}_t = \boldsymbol{\mu}_t + \boldsymbol{\psi}_t + \boldsymbol{\varepsilon}_t, \qquad \boldsymbol{\varepsilon}_t \sim NID(\mathbf{0}, \boldsymbol{\Sigma}_\varepsilon), \quad t = 1, ..., T, \tag{1}$$

where $NID(\mathbf{0}, \mathbf{\Sigma}_\varepsilon)$ denotes normally and (serially) independently distributed with zero mean vector and $N \times N$ positive semi-definite matrix, $\mathbf{\Sigma}_\varepsilon$. Seasonal components may also be added but are not considered here; see Harvey (1989).

The trend is

$$\begin{aligned} \boldsymbol{\mu}_t &= \boldsymbol{\mu}_{t-1} + \boldsymbol{\beta}_{t-1} + \boldsymbol{\eta}_t, \quad \boldsymbol{\eta}_t \sim NID(\mathbf{0}, \mathbf{\Sigma}_\eta) \\ \boldsymbol{\beta}_t &= \boldsymbol{\beta}_{t-1} + \boldsymbol{\zeta}_t, \qquad \boldsymbol{\zeta}_t \sim NID(\mathbf{0}, \mathbf{\Sigma}_\zeta), \end{aligned} \tag{2}$$

With $\mathbf{\Sigma}_\zeta = \mathbf{0}$, and $\mathbf{\Sigma}_\eta$ positive definite, each trend is a random walk plus drift. On the other hand, setting $\mathbf{\Sigma}_\eta = \mathbf{0}$ when $\mathbf{\Sigma}_\zeta$ is positive definite yields a vector of integrated random walks (IRWs) and the extracted trend is typically much smoother than is obtained with a random walk plus drift.

The *similar cycle* model, introduced by Harvey and Koopman (1997), is

$$\begin{bmatrix} \boldsymbol{\psi}_t \\ \boldsymbol{\psi}_t^* \end{bmatrix} = \left[\rho \begin{pmatrix} \cos\lambda_c & \sin\lambda_c \\ -\sin\lambda_c & \cos\lambda_c \end{pmatrix} \otimes \mathbf{I}_N \right] \begin{bmatrix} \boldsymbol{\psi}_{t-1} \\ \boldsymbol{\psi}_{t-1}^* \end{bmatrix} + \begin{bmatrix} \boldsymbol{\kappa}_t \\ \boldsymbol{\kappa}_t^* \end{bmatrix}, \quad t = 1, ..., T, \tag{3}$$

where $\boldsymbol{\psi}_t$ and $\boldsymbol{\psi}_t^*$ are $N \times 1$ vectors and $\boldsymbol{\kappa}_t$ and $\boldsymbol{\kappa}_t^*$ are $N \times 1$ vectors of Gaussian disturbances such that

$$E(\boldsymbol{\kappa}_t \boldsymbol{\kappa}_t') = E(\boldsymbol{\kappa}_t^* \boldsymbol{\kappa}_t^{*'}) = \mathbf{\Sigma}_\kappa, \quad E(\boldsymbol{\kappa}_t \boldsymbol{\kappa}_t^{*'}) = \mathbf{0}, \tag{4}$$

where $\mathbf{\Sigma}_\kappa$ is an $N \times N$ covariance matrix. The parameter ρ is the damping factor; it satisfies $0 < \rho \leq 1$, and $0 \leq \lambda_c \leq \pi$. Because ρ and λ_c are the same in all series, the cycles in the different series have similar properties; in particular their movements are centred around the same period. The model allows the disturbances to be correlated across the series. The same pattern of correlations carries over to the cycles themselves since the covariance matrix of $\boldsymbol{\psi}_t$ is

$$\mathbf{\Sigma}_\psi = (1 - \rho^2)^{-1} \mathbf{\Sigma}_\kappa. \tag{5}$$

Rünstler (2004) extends the model to allow for leads and lags.

Harvey and Trimbur (2003) generalize the model in such a way that it can produce smoother extracted cycles. A univariate *nth-order stochastic cycle*, $\psi_{n,t}$, is defined for $i = 2, ..., n$, by

$$\begin{bmatrix} \psi_{1,t} \\ \psi_{1,t}^* \end{bmatrix} = \rho \begin{bmatrix} \cos\lambda_c & \sin\lambda_c \\ -\sin\lambda_c & \cos\lambda_c \end{bmatrix} \begin{bmatrix} \psi_{1,t-1} \\ \psi_{1,t-1}^* \end{bmatrix} + \begin{bmatrix} \kappa_t \\ \kappa_t^* \end{bmatrix}, \quad t = 1, 2, ..., T$$

$$\begin{bmatrix} \psi_{i,t} \\ \psi_{i,t}^* \end{bmatrix} = \rho \begin{bmatrix} \cos\lambda_c & \sin\lambda_c \\ -\sin\lambda_c & \cos\lambda_c \end{bmatrix} \begin{bmatrix} \psi_{i,t-1} \\ \psi_{i,t-1}^* \end{bmatrix} + \begin{bmatrix} \psi_{i-1,t-1} \\ \psi_{i-1,t-1}^* \end{bmatrix}, \tag{6}$$

where, as in (4), κ_t and κ_t^* are uncorrelated white noise disturbances with mean zero and variance σ_κ^2. Further details on the properties of higher order cycles can be found in Trimbur (2006). The extension to multivariate similar cycles is straightforward.

The disturbance vectors driving the various components are assumed to be mutually uncorrelated in all time periods. One of the implications of this assumption is that the weights in the filters for extracting components are symmetric in the middle of the series; see Harvey and Koopman (2000).

The statistical treatment of unobserved component models is based on the state space form. Once a model has been cast in state space form, the Kalman filter yields estimators of the components based on current and past observations while the associated smoother estimates the components using all the information in the sample. Predictions are made by extending the Kalman filter forward. Root mean square errors (RMSEs) can be computed for all estimators and prediction or confidence intervals constructed. In a Gaussian model the unknown variance parameters are estimated by constructing a likelihood function from the one-step ahead prediction errors, or innovations, produced by the Kalman filter. The likelihood function is then maximized by an iterative procedure.

3 COMMON TRENDS AND CONVERGENCE

The first sub-section below begins by setting out the special case of a single common trend and the restrictions that yield balanced growth. The is done as a prelude to the introduction of the convergence model, the principal feature of which is the modification of a multivariate trend to allow for convergence to a balanced growth path. Thus it is able to capture a common feature that, at best, only becomes apparent towards the end of the sample. The second sub-section investigates the performance of the model using post-sample data.

3.1 Balanced growth and the structural time series error correction model

Let $\mathbf{\Sigma}_\zeta = \mathbf{0}$ in (1). The model has common trends, or equivalently, displays co-integration, if $\mathbf{\Sigma}_\eta$ is not of full rank. If the rank of $\mathbf{\Sigma}_\eta$ is one, there is a single common trend and

$$\mathbf{y}_t = \boldsymbol{\theta}\mu_t + \boldsymbol{\alpha} + \boldsymbol{\psi}_t + \boldsymbol{\varepsilon}_t, \qquad t = 1, ..., T, \tag{7}$$

where μ_t is a random walk with drift

$$\mu_t = \mu_{t-1} + \beta + \eta_t, \quad \eta_t \sim NID(0, \sigma_\eta^2) \tag{8}$$

$\boldsymbol{\theta}$ is an $N \times 1$ vector and $\boldsymbol{\alpha}$ is an $N \times 1$ vector of constants. [Unless μ_0 is set to zero, $\boldsymbol{\alpha}$ must be constrained so as to contain only $N - 1$ free parameters]. In the IRW trend model the existence of common trends depends on the rank of $\mathbf{\Sigma}_\zeta$, and when this is less than N the series, which need to be differenced twice to become stationary, yield linear combinations that are stationary. With a single common trend, the model is again as in (7), but with

$$\mu_t = \mu_{t-1} + \beta_{t-1}, \qquad \beta_t = \beta_{t-1} + \zeta_t, \qquad \zeta_t \sim NID(0, \sigma_\zeta^2) \tag{9}$$

The *balanced growth* STM is a special case of (7) with $\boldsymbol{\theta} = \mathbf{i}$, where $\mathbf{i}$ is a vector of ones. Thus the difference between any pair of series in $\mathbf{y}_t$ is stationary. The trend component may be as in (8) or (9).

If the series are stationary in first differences, balanced growth may be incorporated in a vector error correction model (VECM) by writing

$$\Delta\mathbf{y}_t = \boldsymbol{\delta} + \mathbf{\Gamma}\mathbf{D}\mathbf{y}_{t-1} + \sum_{r=1}^{p} \mathbf{\Phi}_r^* \Delta\mathbf{y}_{t-r} + \boldsymbol{\xi}_t, \qquad \boldsymbol{\xi}_t \sim NID(\mathbf{0}, \mathbf{\Sigma}_\xi) \tag{10}$$

where the $\mathbf{\Phi}_r^{*\prime}s$ are $N \times N$ matrices, $\mathbf{D}$ is a matrix of co-integrating vectors defined such that $\mathbf{Di} = \mathbf{0}$ and the matrix $\mathbf{\Gamma}$ is $N \times (N - 1)$. The *balanced growth* VECM has a single unit root, guaranteed by the fact that $\mathbf{Di} = \mathbf{0}$. The constants in $\boldsymbol{\delta}$ contain information on the common slope, β, and on the differences in the levels of the series, as contained in the vector $\boldsymbol{\alpha}$. Specifically, $\boldsymbol{\delta} = \beta(\mathbf{I} - \sum_{j=1}^{p} \mathbf{\Phi}_j^*)\mathbf{i} - \mathbf{\Gamma}\boldsymbol{\alpha}$. Estimation of $\boldsymbol{\delta}, \mathbf{\Gamma}$ and $\mathbf{\Phi}_r^*, r = 1, .., p$ by OLS applied to each equation in turn is fully efficient since each equation contains the same explanatory variables.

Carvalho and Harvey (2005) modified the multivariate STM by incorporating in the trend a mechanism for capturing convergence to a common growth path. Thus

$$\mathbf{y}_t = \boldsymbol{\alpha} + \boldsymbol{\mu}_t + \boldsymbol{\psi}_t + \boldsymbol{\varepsilon}_t, \quad t = 1, ..., T \tag{11}$$

where, in the preferred specification,

$$\boldsymbol{\mu}_t = \boldsymbol{\Phi}\boldsymbol{\mu}_{t-1} + \boldsymbol{\beta}_{t-1}, \qquad \boldsymbol{\beta}_t = \boldsymbol{\Phi}\boldsymbol{\beta}_{t-1} + \boldsymbol{\zeta}_t, \qquad Var(\boldsymbol{\zeta}_t) = \boldsymbol{\Sigma}_\zeta,$$

with $\boldsymbol{\Phi} = \phi\mathbf{I} + (1-\phi)\mathbf{i}\overline{\boldsymbol{\phi}}'$, where $\overline{\boldsymbol{\phi}} = (\overline{\phi}_1, ..., \overline{\phi}_N)$ is a vector of weights summing to one and each weight lying between zero and one. Using scalar notation to write the model in terms of the common trend, $\overline{\mu}_{\phi,t} = \sum \overline{\phi}_i \mu_{it}$, and convergence components, $\mu_{it}^\dagger = \mu_{it} - \overline{\mu}_{\phi,t}, i = 1, ..., N$, yields

$$y_{it} = \alpha_i + \overline{\mu}_{\phi,t} + \mu_{it}^\dagger + \psi_{it} + \varepsilon_{it}, \qquad t = 1, ..., T, \quad i = 1, ..., N, \tag{12}$$

where the common trend is

$$\overline{\mu}_{\phi t} = \overline{\mu}_{\phi,t-1} + \overline{\beta}_{\phi,t-1}, \qquad \overline{\beta}_{\phi t} = \overline{\beta}_{\phi,t-1} + \overline{\zeta}_{\phi t},$$

and the convergence components are

$$\mu_{it}^\dagger = \phi\mu_{i,t-1}^\dagger + \beta_{it}^\dagger, \quad \beta_{it}^\dagger = \phi\beta_{i,t-1}^\dagger + \zeta_{it}^\dagger, \quad i = 1, ..., N, \tag{13}$$

with $\zeta_{it}^\dagger = \overline{\boldsymbol{\phi}}'\boldsymbol{\zeta}_t^\dagger$, $\beta_{it}^\dagger = \overline{\boldsymbol{\phi}}'\boldsymbol{\beta}_t^\dagger$ and $0 < \phi < 1$. If we write a convergence component in what might be termed *second-order error correction* form,

$$\begin{aligned} \Delta\mu_t^\dagger &= (\phi - 1)\mu_{t-1}^\dagger + \beta_{t-1}^\dagger, \quad t = 1, ..., T, \\ \Delta\beta_t^\dagger &= (\phi - 1)\beta_{t-1}^\dagger + \zeta_t^\dagger, \end{aligned} \tag{14}$$

it can be seen that there is a convergence mechanism operating on both the gap in the level and the gap in the growth rate. One of the interesting features of the second-order mechanism is that the predicted gap between the series can actually widen in the short run. When ϕ is close to one, the extracted convergence components tend to be quite smooth and, as with an IRW trend, there is a clear separation of long-run movements and cycles. The forecasts for each series converge to a common growth path, but in doing so they may exhibit temporary divergence. On the other hand, when ϕ is equal to one, the STM of section 2 is obtained and there is no convergence.

The above convergence model will be referred to as a structural time series error correction model (STECM). In what follows all disturbances are assumed to be Gaussian and estimation is by maximum likelihood (ML).

3.2 Post-sample predictive testing for US Regions

Carvalho and Harvey (2005) investigated convergence for the logarithms of real per capita incomes in US census regions: New England (NE), Mid-East (ME), Great Lakes (GL), Plains (PL), South East (SE), South West (SW), Rocky Mountains (RM) and Far West (FW). Annual data from 1950 to 1999 was used. The preliminary investigation of stylized facts reported indicated that the two richest regions, NE and ME, follow growth paths which, especially for the last two decades, seem to be diverging from the growth paths of the other regions. Hence a STECM was only fitted to the six poorer regions. The model is as in (12) but with absolute convergence, that is $\alpha_i = 0$ for all $i = 1, .., N$. The convergence parameter, ϕ, was estimated as 0.889, while the estimates of the common trend weights, $\overline{\phi}_i$, were dominated by those for the Great Lakes and Plains at 0.64 and 0.30 respectively. In 1950 the standard deviation of the convergence components was 0.180, while in 1999 it was only 0.051. Figure 1 plots the smoothed estimates of these components from 1990 to 1999 together with their forecasts. The second-order ECM shows up in the tendency for the forecasts to move apart in the short-run with the standard deviations in 2002, 2004 and 2020 being 0.054, 0.052 and 0.021 respectively.

Since the model in Carvalho and Harvey (2005) was estimated, new observations, from 2000 to 2003, have become available. (The Bureau of Economic Analysis had altered the base year, so we deflated at 1996 prices to ensure compatibility with the earlier series). We can therefore test its forecasting performance using post-sample data. For each region the univariate post-sample predictive test statistic, $\xi^*(4)$, is calculated as the sum of squares of the standardized one-step ahead prediction errors. If the specification of the model remains the same, $\xi^*(4)$ is asymptotically distributed as χ^2_4 ; see Harvey (1989, p 271). The values of these statistics are shown in table 1 for the STECM and the STM, (1). Both models are satisfactory, but the crucial test comes with their unconditional predictive performance.

There are three components in the STECM forecast: the common trend, the cycle and the convergence components. In order to gain some insight into the contribution of the convergence components to the overall predictive performance we created a composite series by weighting the observations in the same way as for the common trend, that is $\overline{y}_{\phi t} = \sum \overline{\phi}_i y_{it}$. We then subtracted $\overline{y}_{\phi t}$ from each series to give

$$y_{it}^{\dagger} = y_{it} - \overline{y}_{\phi t} = \mu_{it}^{\dagger} + \psi_{it} - \overline{\psi}_{\phi t} + \varepsilon_{it} - \overline{\varepsilon}_{\phi t}, \qquad t = 1, ..., T, \quad i = 1, ..., N,$$

where $\overline{\psi}_{\phi t}$ and $\overline{\varepsilon}_{\phi t}$ are defined analogously to $\overline{y}_{\phi t}$. The effect of this operation is to give series very close to the convergence components since the common trend is removed and the cycles are - as we will see in the next section - very highly correlated. Hence it is possible to gauge the role of the convergence components by comparing the $y_{it}^{\dagger\prime}s$ in the post-sample period, as shown in figure 2, with the predictions in figure 1. As can be seen, the initial divergence predicted for FW does indeed take place, SE and SW remain roughly the same, while GL and PL continue to move closer to zero.

Autoregressive models are the norm in econometric studies of convergence and so they provide a good yardstick against which to judge the forecasting performance of the STECM. Table 2 compares the forecasting performance of the STECM, STM, VAR(2) and balanced growth VECM for each region using the extrapolative sum of squares of the unconditional forecast errors, that is

$$ESS(T, \ell) = \sum_{j=1}^{\ell} (y_{T+j} - \widehat{y}_{T+j|T})^2, \quad \ell = 1, 2, ..,$$

where T is the last period of data used to generate the forecasts, $\widehat{y}_{T+j|T}$, and ℓ is the lead time; see Harvey (1989, p 273). The forecasting performance of the STECM is similar to that of the STM; it will take a longer period to judge the effectiveness of the convergence mechanism. The VAR does not perform well on the whole. It does particularly badly for RM and it does not deal well with the temporary divergence in FW. The balanced growth VECM fares better. However, it is not as good as the STECM for two regions and is about the same for three.

Some researchers, for example Carlino and Sill (2001), have been tempted to use co-integration tests to find the number of common trends in US regions. For our sample of six regions the Johansen trace test, reported in table 3, cannot reject the null hypothesis of two co-integrating vectors, that is four common trends. However, in our view there is little point in estimating a model with two co-integrating vectors. The application of the test again simply confirms that it is misleading if series are in the process of converging.

4 COMMON AND SIMILAR CYCLES

Fitting a multivariate STM to the eight US regions, as was done in Carvalho and Harvey (2005), provides a good illustration of the stylized facts produced by a similar (first-order) cycle model, (3), and yields an interesting comparison with the Beveridge-Nelson cycles reported in Carlino and Sill (2001). ML estimation gave a damping factor of 0.80 and a period of 5.3 years. In a plot of the smoothed cyclical components the recessions of 1954, 1961, 1970, 1975, 1980, 1982 and 1991 all show up with a high degree of coherence across regions; this is not the case for the cycles estimated by Carlino and Sill (2001). There are considerable differences in volatility, with the variance of the disturbances in PL being almost six times as great as that of ME. Since the expected value of $\psi_t^2 + \psi_t^{*2}$ is $2\sigma_\psi^2$, the amplitude of a cycle is best measured by its standard deviation times $\sqrt{2}$; these figures, multiplied by 100, are shown in the table 4. Carlino and Sill (2001, p 452) also report big differences in volatility, but our ordering differs from theirs. In particular we find that the richest regions (NE, ME and FW) are those with the least volatile cyclical components.

The extent to which similar cycles move together depends on the correlations between the disturbances driving them since, $\mathbf{\Sigma}_\psi$, the covariance matrix of the vector of cycles is $(1-\rho^2)^{-1}\mathbf{\Sigma}_\kappa$.There are high positive correlations between the cyclical disturbances in all pairs of US regions. The minimum is 0.720 and the maximum is 0.987. The first principal component of $\widetilde{\mathbf{\Sigma}}_\kappa$ accounts for 91% of the total variance while the second accounts for a further 5%.

4.1 Common cycles

If $\mathbf{\Sigma}_\psi$ is less than full rank, there are common cycles. If the rank of $\mathbf{\Sigma}_\psi$ is one, there is a single common cycle and the model can be written

$$y_{it} = \mu_{it} + \theta_i \psi_t + \varepsilon_{it}, \qquad i = 1, ..., N, \quad t = 1, ..., T \tag{15}$$

where ψ_t is a scalar cycle and the $\theta_i' s$ allow the common cycle to appear in each series with a different amplitude. One of the $\theta_i' s$ is set equal to unity for identifiability and, since the cycles have zero mean, there is no need for a vector of constants to be added as with common trends. A single common cycle is a common feature in the sense of Engle and Kozicki (1993) in that it may be removed by a linear combination, $\overline{\boldsymbol{\theta}}$, of the observations with the property that $\overline{\boldsymbol{\theta}}'\boldsymbol{\theta}$ =0, where the $N \times 1$ vector $\boldsymbol{\theta} = (\theta_1, ..., \theta_N)'$. [There is a slight difference between the Vahid and Engle (1993) common cycle definition and the one above in that if the series are $I(1)$ and not co-integrated, the former requires that there exist linear combinations whose first differences are unpredictable from their past. In a structural model with random walk trends, the presence of an irregular component means that although a common cycle is removed by $\overline{\boldsymbol{\theta}}'\mathbf{y}_t$, $\overline{\boldsymbol{\theta}}'\Delta\mathbf{y}_t$ has a first order moving average representation and so is predictable. We are grateful to a referee for pointing this out.]

The common cycle constraint is a strong one, particularly if the irregular component is relatively small. Harvey and Trimbur (2003) found that fitting higher order cycles tends to result in the irregular component becoming relatively more important while the extracted cycle is smoother. Thus higher order cycles may be better for modeling common cycles.

4.2 Testing for common cycles

We now consider how to test the null hypothesis of a single common cycle in a bivariate STM against the alternative of similar cycles. Under the alternative hypothesis $\boldsymbol{\Sigma}_\kappa$ is of full rank, while under the null hypothesis the correlation between the disturbances in the two cycles is one. Since a correlation of one is on a boundary of the admissible parameter space, the asymptotic distribution of the likelihood ratio (LR) statistic is an even mixture of χ^2_0 and χ^2_1 and the $10\%, 5\%$ and 1% critical values are 1.642, 2.706 and 5.412 respectively. This is an example of the application of a classic result of Chernoff (1954). Andrews (2001, p 712-4) provides a recent unified discussion of the relevant theory under very general conditions. Another way of seeing the result is to transform the vector $(\psi_{1t}, \psi_{2t})'$ so that there is a common cycle in each series and a specific cycle in one: we can then test the null hypothesis that the variance of the specific cycle is zero.

In order to investigate small sample properties we carried out a series of Monte Carlo experiments on a simple bivariate cycle plus irregular model

$$\begin{aligned} y_{1t} &= \psi_{1t} + \varepsilon_{1t}, \qquad t = 1, ..., T \\ y_{2t} &= \psi_{2t} + \varepsilon_{2t}. \end{aligned}$$

with $E(\boldsymbol{\varepsilon}_t\boldsymbol{\varepsilon}_t') = \sigma_\varepsilon^2\mathbf{I}$ and

$$E(\boldsymbol{\psi}_t\boldsymbol{\psi}_t') = \boldsymbol{\Sigma}_\psi = \sigma_\psi^2 \begin{bmatrix} 1 & \omega \\ \omega & 1 \end{bmatrix}.$$

Only first and second-order cycles were considered with $\rho = 0.9$ for $n = 1$ and 0.75 for $n = 2$. The test is of the null hypothesis of a common cycle, that is $\omega = 1$ against the alternative of similar cycles, $\omega < 1$. All the results reported below are based on 10,000 replications with random numbers generated by Ox subroutines.

Table 5 shows the estimated test sizes for $T = 100$, 200 and 500 for signal-noise ratios, $q = \sigma_\psi^2/\sigma_\varepsilon^2$, of 1 and 10. As can be seen, there is a slight tendency for the tests to be undersized, but there is a movement towards the nominal size as T increases. Some experiments were also run for $q = 0.1$ and these showed the test to be somewhat oversized, particularly for $n = 1$. However, such a small q is unlikely to arise in practice. The probabilities that the LR statistic is zero are somewhat greater than the 0.5 predicted by asymptotic theory, but they fall as T increases and seem to be closer to 0.5 for higher values of q. There is a slight tendency for the probabilities to be closer to 0.5 for $n = 2$, but there is no corresponding movement in the size up towards the nominal.

Table 6 shows estimated probabilities of rejection when asymptotic critical values are used. Size corrected powers would be somewhat larger for small samples. As might be anticipated, the power of the test increases with q. There appears to be little difference between first and second order cycles.

For $N > 2$, the distributional theory for the LR statistic becomes more complex and it may be necessary to resort to simulation methods to obtain critical values; see Andrews (2001), Robin and Smith (2000) and Stoel *et al* (2006). For example, when $N = 3$, Stoel *et al* (2006) show that the asymptotic distribution of the LR statistic for a test of the null hypothesis of one common cycle against the alternative that $\boldsymbol{\Sigma}_\psi$ is of full rank is a mixture of chi-squares with zero, one and three degrees of freedom with weights depending on the data.

4.3 Examples

This sub-section illustrates the common cycles test with two applications that have appeared in the common features literature.

US and Canada - Engle and Kozicki (1993) entertain the possibility of common business cycle features in US and Canadian GDP. Here we investigate whether STMs with generalized cyclical specifications support the notion of a common cycle. The data are the logarithms of quarterly, seasonally adjusted real GDP from 1961:1 to 2001:4, for the United States (Source: Bureau of Economic Analysis, U.S. Department of Commerce) and Canada (Source: Statistics Canada).

Models of the form (1), with smooth trends, were estimated with generalized cycles of order one to six. To save space only the results for cycles of order one, two and four are reported in table 7. The first and second-order models show the best diagnostics. The second-order model would be chosen on the basis of goodness of fit as measured by the standard error of each equation. The estimated period is around 7.5 years. Figures 3 and 4 show the first and second order cycles obtained by state space signal extraction. The second order cycles are smoother. The cycles appear to differ somewhat in their timing, though on the whole the differences are slight, and there is no firm evidence to suggest that the US cycle leads the Canadian one.

The estimated periods for the common cycle models are close to those for the similar cycle case and the Box-Ljung Q-statistics are not very different; see table 8. The best fit is again obtained for the second-order cycle. The load factor for Canada is 0.85 (the model is normalized by setting the US coefficient in $\boldsymbol{\theta}$ to unity) and so the amplitude of the Canadian cycle is, on average, only 0.85 of that of the US. Figure 5 shows the extracted cycles.

The correlation between the cycle disturbances is around 0.8 in the similar cycle models, with the maximum being 0.85 for $n = 2$. In all cases the null hypothesis of a common cycle is convincingly rejected by the LR test. Such evidence as there is in favor of a common cycle is strongest for $n = 4$, but even here the LR statistic of 6.10 is well in excess of the 1% critical value of 5.41. The Akaike and Bayes information criteria, AIC and BIC, point to similar conclusions.

Are US and Canadian GDP co-integrated? The Johansen trace test suggests not. The prob value of a test of no-cointegration (against one cointegrating vector) is 0.383. The STM - which has I(2) trends - has a correlation between growth rates of 0.801 for $n = 2$. The two trends are shown in figure 6 and a simple plot of their difference makes it clear why tests reject the null. The lack of balanced growth may be due to a number of factors, for example not using per capita data or converting currencies at an inappropriate exchange rate.

Perhaps more interesting is the graph of the smoothed estimates of the growth rates shown in figure 6. Both exhibit a steady decline from the sixties coupled with long swings in the eighties and nineties. As with the extracted cycles we have a near common feature.

US regions - Attempting to estimate a model for all eight US regions with one common cycle led to implausible results, even though, as noted earlier, the first principal component, accounted for over 90% of the total variance. The same exercise for quarterly data, from 1969:1 to 1999:4, produced similar conclusions; see Carvalho and Harvey (2002). However, these models used only first-order cycles.

We now turn to investigating the plausibility of common cycles between certain pairs of US regions when higher order cycles are used. We focus on the NE and ME regions with quarterly data where a bivariate model gives first order cycles with a correlation of 0.97. Figure 7 shows the trends and cycles. As can be seen there are slight differences in the cycles. We also note that the correlation between the slopes is 0.962; as with the US-Canada example, it is interesting to see the extent to which the growth rates move together.

Tables 9 and 10 show the results of estimating models with higher-order cycles. The model selection criteria favor the common cycle restriction for $n = 2$ and 4, and the likelihood ratio test fails to reject the null hypothesis of common cycles at the 10% level of significance in both cases, whereas for $n = 1$ it easily rejects at the 1% level. The BIC, which for moderately large sample sizes attaches a greater penalty to model complexity, reaches a minimum for the second-order common cycle model. The smoothed cyclical components for the two regions are shown in figure 8.

If there is a single common cycle, the cycles in each series in (15) have the same amplitude if $\theta_i = 1$ for all $i = 1, .., N$, and so are identical. This restriction may be tested by a standard LR test in which the test statistic has a χ^2_{N-1} distribution, asymptotically, under the null hypothesis. For NE and ME the LR test statistics for $n = 2$ and 4 are 1.46 and 0.28 respectively and so it seems reasonable to treat the cycles as being identical.

Although some other pairs of regions also appear to share a common higher-order cycle, there is not enough communality to allow a single common cycle for all regions even when higher order cycles are used.

Finally we note that Vahid and Engle (1993, p 355-8) analyzed the four industrial regions NE, ME, GL and FW. They found three common trends, as did Carvalho and Harvey (2005). While Vahid and Engle point out in a footnote on p356 that they '..are not convinced that cointegration is necessary for convergence', they nevertheless estimate a VECM incorporating a co-integrating vector $ME - 0.48NE - 0.45GL - 0.07FW$, the economic interpretation of which is unclear to us. On the basis of canonical correlations of the first differences, they find evidence for two common cycle features which are linear combinations of the first differences. As with the co-integrating vector, we are uncertain about the meaning of these combinations.

5 CONCLUSIONS

Common trends are undoubtedly the most important common feature of economic time series. However, data on countries or regions often shows that they are converging, have just converged or have converged some time ago but still have a large part of the series dependent on initial conditions. This is certainly the case with per capita GDP in the US regions. Such series are not co-integrated within the sample, but the fact that they are converging to a special case of a co-integrated system, namely balanced growth, needs to be taken on board if coherent medium term forecasts are to be made. This article has reviewed the unobserved components error correction model, the STECM, proposed in Harvey and Carvalho (2005) and compared its forecasting performance in a post-sample period with unrestricted VARs and balanced growth VECMs. The way in which the second-order error correction mechanism copes with temporary divergence is particularly appealing. In the case of the US regions, the STECM forecasts well and predicts the temporary divergence of the Far West at the beginning of the post-sample period. The relatively poor forecasts obtained with an unrestricted VAR illustrates the case for taking account of convergence to balanced growth.

Structural time series models can handle common cycles within the framework of similar cycles and the cycle model can be generalized so that when extracted it is relatively smooth. We develop a likelihood ratio test for common cycles in a bivariate series and apply it to the series on US and Canadian GDP. Although the US and Canadian cycles move closely together, the hypothesis that they share a common cycle is decisively rejected for a range of cyclical models. However, for some pairs of US regions it seems that a common second-order cycle cannot be rejected: the example we give is that of the Mid-East and New England.

In the bivariate case, the asymptotic critical values for the common cycle test statistic are obtained from tables of the χ_1^2 distribution by doubling the nominal significance level. More generally, it seems difficult to find a correspondingly simple test for a specific number of common cycles. It is perhaps worth remarking, though, that common cycle restrictions may not offer significant gains in forecasting; making allowance for the cycles being highly correlated may be all that is needed.

Finally we note that because a structural time series model can include several components, the possibility of more than two common features can be entertained. In particular, it is easy to allow for common seasonal effects and seasonal co-integration tests can be carried out as described in Busetti (2006).

ACKNOWLEDGEMENTS

The initial stages of the work on convergence was supported by the ESRC as part of a project on Dynamic Factor Models for Regional Time Series, grant number L138 25 1008. Trimbur acknowledges the support of the US Census Bureau during his time as a Post-Doctoral Researcher. Note that the analysis and conclusions set forth in the paper do not indicate concurrence by the Board of Governors or the staff of the Federal Reserve System.

We wish to thank Bill Bell, David Findlay and Richard J. Smith, the referees and participants at the ESRC Econometric Study Group meeting at IFS in May, 2005 for helpful comments. Special thanks go to Giovanni Urga for encouraging us to write the paper in the first place and seeing it through various drafts.

REFERENCES

Andrews, D.W.K. (2001), "Testing when a parameter is on the boundary of the maintained hypothesis," *Econometrica,* 69, 683-734.

Busetti, F. (2006), "Tests of seasonal integration and cointegration in multivariate unobserved component models," forthcoming *Journal of Applied Econometrics.*

Carlino, G., and Sill, K. (2001), "Regional Income Fluctuations: Common Trends and Common Cycles," *Review of Economics and Statistics*, 83, 446-56.

Carvalho, V.M., and Harvey, A.C. (2002)., "Growth, cycles and convergence in US regional time series," DAE Working paper 0221, University of Cambridge.

Carvalho, V.M., and Harvey, A.C. (2005), "Growth, cycles and convergence in US regional time series," *International Journal of Forecasting*, 21, 667-86.

Chernoff, H. (1954). "On the distribution of the likelihood ratio," *Annals of Mathematical Statistics,* 25, 573-8.

Doornik, J. A. (1999), *Ox: An Object-Oriented Matrix Language.* 3rd edition. Timberlake Consultants Press: London.

Doornik, J.A., and Hendry, D.F. (2001). *Modelling Dynamic Systems using PcGive. Volume II.* Timberlake Consultants Press: London.

Engle, R.F., and Kozicki, S. (1993), "Testing for common features," *Journal of Business and Economic Statistics*, 11, 369-80.

Harvey A.C. (1989), *Forecasting, Structural Time Series Models and the Kalman Filter*, Cambridge University Press, Cambridge.

Harvey, A.C., and Koopman, S.J. (2000), "Signal extraction and the formulation of unobserved components models," *Econometrics Journal,* 3**,** 84-107.

Harvey, A C and Trimbur, T. (2003), "General model-based filters for extracting trends and cycles in economic time series," *Review of Economics and Statistics*, 85, 244-55.

Harvey, A.C., and Koopman, S.J. (1997), "Multivariate structural time series models," in *System dynamics in economic and financial models,* C. Heij et al.(eds)*,* 269-98: Chichester: Wiley and Sons.

Koopman S.J., Harvey, A.C., Doornik, J.A and Shephard, N. (2000). *STAMP 6.0 Structural Time Series Analyser, Modeller and Predictor*, London: Timberlake Consultants Ltd.

Koopman, S.J., Shephard, N. and Doornik, J.A. (1999), "Statistical algorithms for models in state space form using SsfPack 2.2 (with discussion)," *Econometrics Journal*, 2, 107-160.

Robin, J-M., and Smith, R. J. (2000), "Tests of rank," *Econometric Theory,* 16, 151-75.

Rünstler, G. (2004), "Modelling phase shifts among stochastic cycles," *Econometrics Journal,* 7, 232-48.

Stoel, R.D., Garre, F.G., Dolan,C., and van den Wittenboer, G. (2006), "On the likelihood ratio test in structural equation modeling when parameters are subject to boundary constraints," forthcoming *Psychological Methods.*

Trimbur, T. (2006), "Properties of higher order stochastic cycles," *Journal of Time Series Analysis,* 27, 1-17.

Vahid, F. and Engle, R.F. (1993), "Common trends and common cycles," *Journal of Applied Econometrics* 8, 341-60.

Table 1 Post sample predictive test statistics, $\xi^*(4)$, for US regions

	SW	**RM**	**FW**	**SE**	**GL**	**PL**
STM	2.2091	4.2354	3.7936	2.0760	0.54808	1.9645
$STECM$	4.7282	5.1635	3.8904	2.1291	1.7598	2.9232

Table 2 $ESS(1999,4)$ for US regions

	SW	**RM**	**FW**	**SE**	**GL**	**PL**
VAR	0.0006	0.0123	0.0031	0.0012	0.0019	0.0076
$VECM$	0.0035	0.0006	0.0014	0.0003	0.0025	0.0008
STM	0.0008	0.0009	0.0013	0.0002	0.0016	0.0016
$STECM$	0.0013	0.0021	0.0015	0.0003	0.0017	0.0007

Table 3 Trace test for co-integration in US regions

R	$Trace$ test statistic
0	141.92 [0.000] **
1	78.474 [0.008] **
2	36.938 [0.355]
3	7.225 [0.632]
4	5.0749 [0.799]
5	0.94009 [0.332]

NOTE: Computed using PcGive; see Doornik and Hendry (2001)

.

Table 4 Estimated amplitudes of cycles in real per capita GDP in US regions

$$\begin{bmatrix} NE & ME & GL & PL & SE & SW & RM & FW \\ .9 & .6 & 1.1 & 1.4 & 1.0 & 0.9 & 1.3 & .8 \end{bmatrix}$$

Table 5 : Estimated sizes of tests based on asymptotic critical values and probability that the test statistic is zero

		$q=1$				$q=10$			
T	n	10%	5%	1%	Pr(0)	10%	5%	1%	Pr(0)
100	1	0.069	0.040	0.010	0.646	0.081	0.042	0.013	0.612
	2	0.080	0.039	0.008	0.612	0.076	0.037	0.008	0.600
200	1	0.078	0.040	0.009	0.596	0.080	0.040	0.009	0.585
	2	0.078	0.039	0.008	0.577	0.078	0.037	0.006	0.567
500	1	0.085	0.042	0.007	0.564	0.084	0.043	0.009	0.549
	2	0.081	0.040	0.007	0.551	0.083	0.040	0.008	0.545

Table 6 Estimated probability of rejection at the (asymptotic) 5% level of significance for the common cycle test

		$q=1$		$q=10$	
n	T	$\omega=0.95$	$\omega=0.8$	$\omega=0.95$	$\omega=0.8$
1	100	0.202	0.518	0.869	0.997
2		0.193	0.521	0.886	1
1	200	0.263	0.760	0.988	1
2		0.242	0.757	0.994	1

Table 7 Similar cycle model fitted to US and Canadian GDP: (a) Estimated variance parameters ($\times 10^5$), equation standard errors ($\times 10^5$) and Box-Ljung statistics; (b) Correlations, cycle parameters and information criteria

		US					Canada				
	n	σ^2_ζ	σ^2_ψ	σ^2_ε	$Q(12)$	SE	σ^2_ζ	σ^2_ψ	σ^2_ε	$Q(12)$	SE
(a)	1	0.08	28.1	0.78	15.48	805	0.13	25.8	1.04	13.13	834
	2	0.04	43.9	1.54	12.97	789	0.09	37.5	1.83	15.78	816
	4	0.11	25.7	1.59	18.82	823	0.14	23.6	1.81	20.49	845

	n	$corr(\zeta)$	$corr(\kappa)$	$corr(\varepsilon)$	ρ	$2\pi/\lambda_c$	AIC	BIC
(b)	1	0.869	0.796	-0.763	0.941	24.21	-2185.81	-2139.31
	2	0.801	0.855	-0.137	0.794	30.5	-2186.89	-2140.39
	4	0.704	0.824	-0.088	0.505	31.75	-2174.94	-2128.45

Table 8 Common cycle model for US and Canadian GDP : equation standard errors ($\times 10^5$), goodness of fit, cycle parameter estimates, information criteria and common cycle LR test statistic

	US		Canada		Cycle parameters			Test	Information criteria	
n	SE	$Q(12)$	SE	$Q(12)$	ρ	$2\pi/\lambda_c$	θ	LR	AIC	BIC
1	813	13.90	854	14.92	0.941	21.99	0.805	14.94	-2172.86	-2129.46
2	806	12.61	837	16.41	0.767	28.80	0.851	11.81	-2177.08	-2133.68
4	822	15.06	838	20.19	0.506	28.57	0.814	6.10	-2170.85	-2127.45

Table 9 Similar cycle model for per capita GDP in New England and Mid-East: correlations, cycle parameters and model selection criteria

n	$corr(\zeta)$	$corr(\kappa)$	$corr(\varepsilon)$	ρ	$2\pi/\lambda_c$	AIC	BIC
1	0.962	0.970	0.797	0.940	29.24	-1760.94	-1718.63
2	0.948	0.975	0.829	0.729	35.03	-1760.72	-1718.42
4	0.937	0.986	0.828	0.377	28.70	-1756.74	-1714.43

Table 10 Common cycle model for New England and Mid-East: cycle parameters, LR test statistic and model selection criteria

n	ρ	$2\pi/\lambda_c$	θ	LR	AIC	BIC
1	0.951	21.96	0.852	9.95	-1752.99	-1713.50
2	0.731	31.51	0.867	1.50	-1761.23	-1721.74
4	0.466	25.81	0.870	1.46	-1757.27	-1717.79

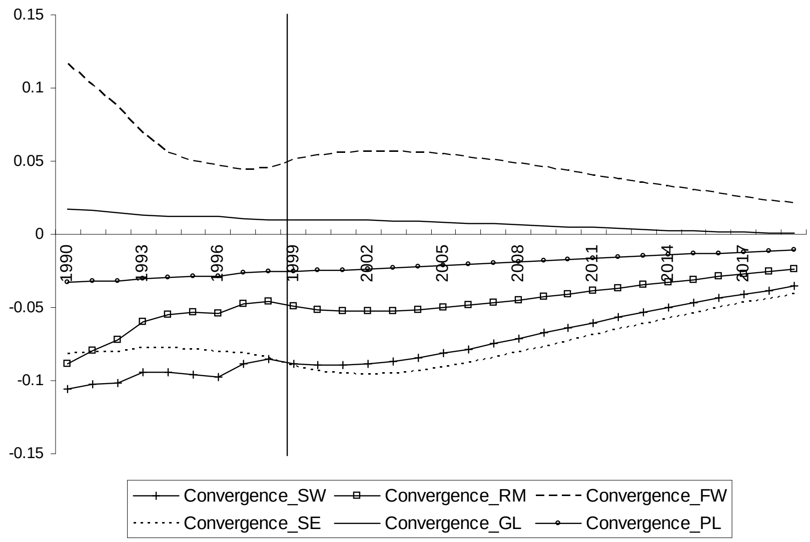

Figure 1: Forecasts for convergence components for US regions together with smoothed estimates up to 1999.

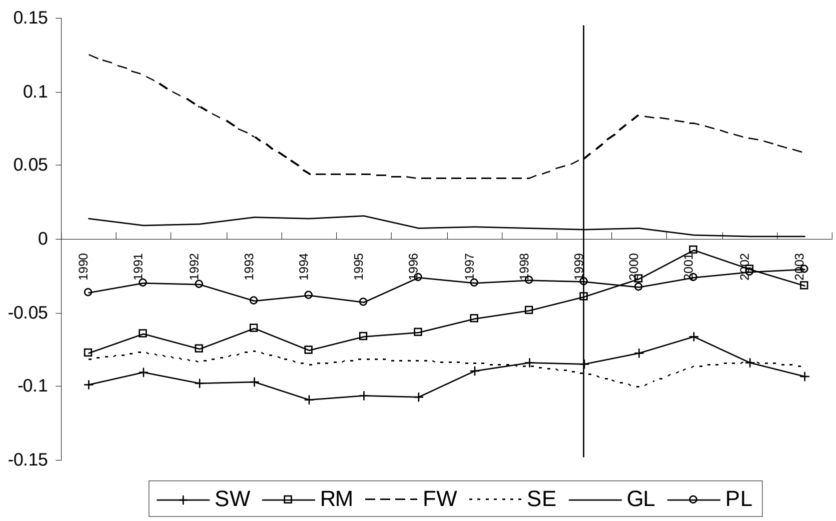

Figure 2: US regional series after removing a composite series contructed using the common trend weights

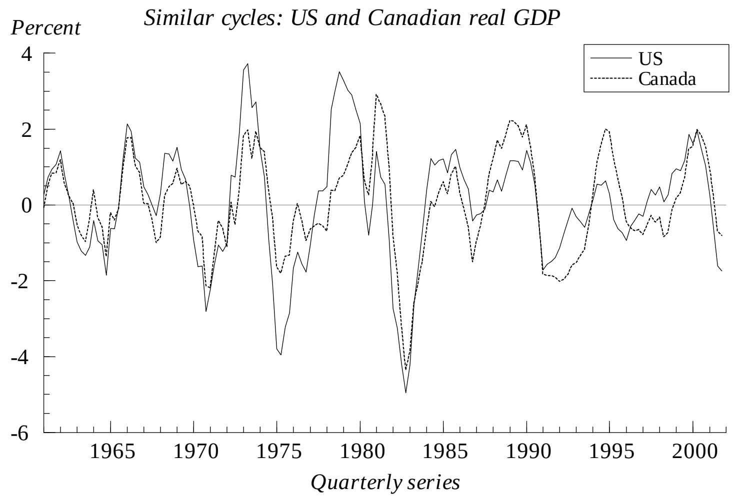

Figure 3: Similar first-order cycles in US and Canadian GDP

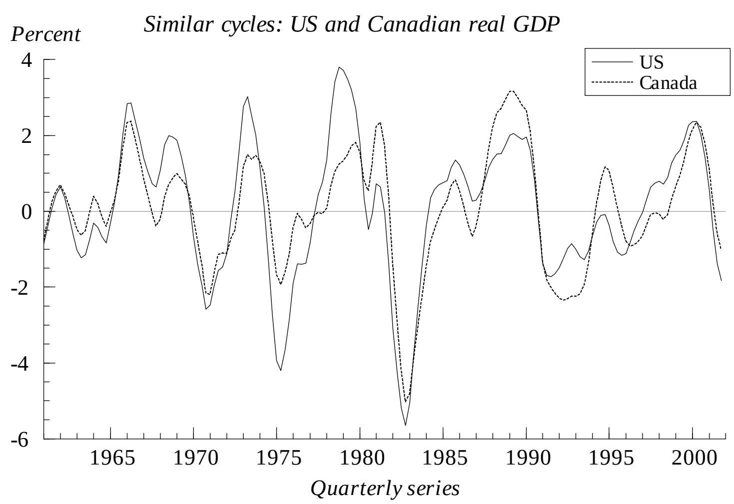

Figure 4: Similar second-order cycles for US and Canada

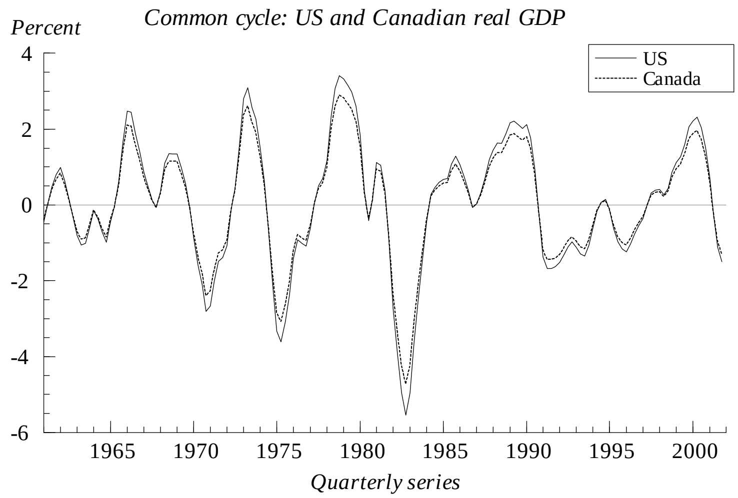

Figure 5: Common second-order cycle for US and Canada

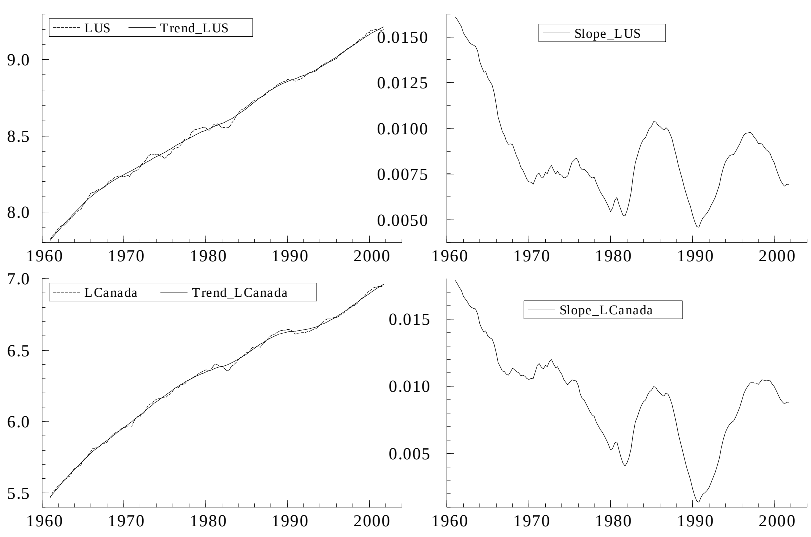

Figure 6: Trends and growth rates for US and Canadian GDP

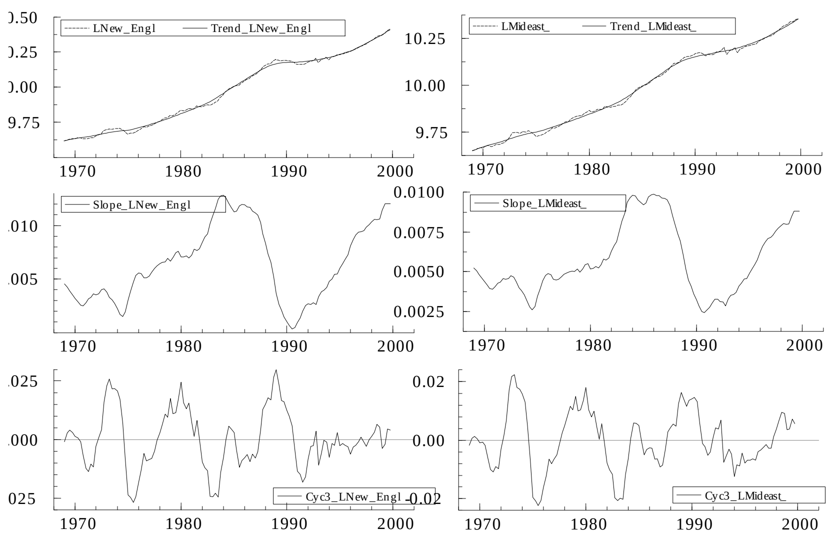

Figure 7: Trends, growth rates and (first-order) cycles in the New England and Mid-East regions of the US.

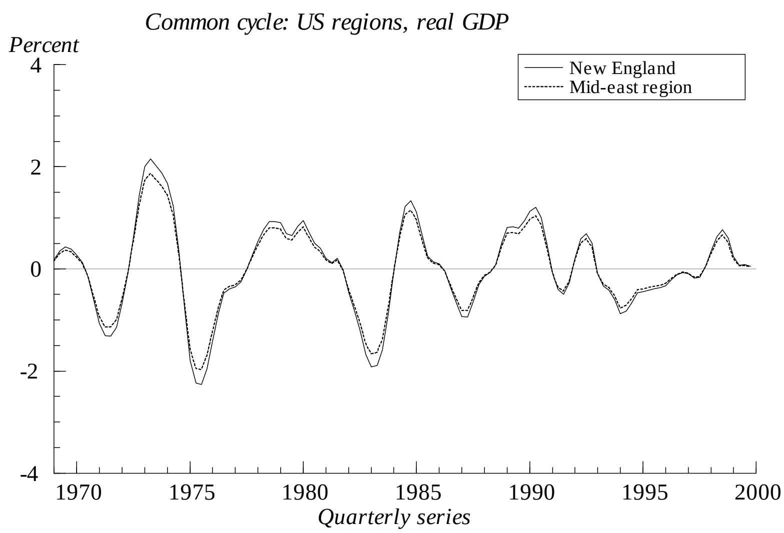

Figure 8: Second order common cycle in real GDP in the New England and Mid-East regions of the US.