Trends, cycles and convergence

Andrew Harvey
Faculty of Economics, University of Cambridge

November 13, 2001

Abstract

This article first discusses ways of decomposing a time series into trend and cyclical components, paying particular attention to a new class of model for cycles. It is shown how using an auxiliary series can help to achieve a more satisfactory decomposition. A discussion of balanced growth then leads on to the construction of new models for converging economies. The preferred models combine unobserved components with an error correction mechanism and allow a decomposition into trend, cycle and convergence components. This provides insight into what has happened in the past, enables the current state of an economy to be more accurately assessed and gives a procedure for the prediction of future observations. The methods are applied to data on the US, Japan and Chile.

KEYWORDS: Band pass filter, error correction, Kalman filter, state space, turning point, unobserved components.

1 Introduction

Determining turning points in the business cycle is a difficult problem. Making sensible predictions concerning the growth path of an economy in the medium or long term is even harder. The aim of this article is to explore what can be achieved by analysing and modeling time series observations on GDP and other macroeconomic time series.

Separating out trends and cycles is fundamental to a good deal of economic analysis. It is often done by applying filters in a rather arbitrary fashion. For example, the low-pass filter introduced by Hodrick and Prescott (1997) is frequently used to remove trends in situations where it can create considerable distortions; see Harvey and Jaeger (1993) and Cogley and Nason (1995). Rather than simply defining the cycle as the detrended series, a band pass filter may be used to extract it, the argument being that high frequency as well as low frequency components need to be removed. Baxter and King (1999) consider the design of band pass filters and their implementation in finite samples. Their prime concern is to approximate the 'ideal' filter, a perfectly sharp band pass filter which removes all frequencies outside a certain range. However, as with the Hodrick-Prescott (HP) filter, considerable distortions can arise as shown by Murray(2001).

The view expressed in this paper is trends and cycles are best constructed using unobserved component, or structural, time series models. The parameters in such models are typically estimated by maximum likelihood and, once this has been done, optimal estimates of the components are obtained by smoothing algorithms. The calculations are most easily performed by putting the model in state space form.

Section 2 begins by discussing the basic ideas of structural time series models and reviewing the link with the HP filter. An extended class of cyclical models is then introduced. Harvey and Trimbur (2001) produce argue that these models enable smoother cycles to be extracted and that they lead to a more satisfactory decomposition into trend and cycle at the end of the series. The extraction of these generalised cycles is closely linked to the application of Butterworth band pass filters. These filters are widely used in engineering but have only recently been introduced into economic statistics; see Gomez (2001). The analysis of such filters reveals that a model yielding the equivalent of an ideal band pass filter can be obtained as a limiting case. Fitting models with the generalised cyclical component to US macroeconomic series illustrates the point about their yielding clearer and smoother cycles than are normally obtained.

Structural models can also be extended so as to include more than one cycle. A model with two cycles turns out to work well on quarterly Chilean GDP data.

Multivariate models are discussed in section 3. A related series with a more pronounced cycle, such as investment, may help in extracting a 'better' cycle from GDP. Multivariate models can also be set up so as to handle economies which have converged and so have a stable relationship. These are called balanced growth models. However, the more relevant question for developing economies, such as Chile, is whether convergence is actually taking place. Section 4 examines ways of assessing and modeling convergence between two economies. A dynamic error correction model is proposed and then extended so as to incorporate a mechanism which allows convergence to take place smoothly. Unobserved component and autoregressive versions of these models are fitted to per capita data on GDP in the US and Japan.

Section 5 brings together the material from the earlier sections to set out bivariate models for the levels of two converging economies. The preferred models combine unobserved components with an error correction mechanism and allow a decomposition into trend, cycle and convergence components. This provides insight into what has happened in the past, enables the current state of an economy to be more accurately assessed and gives a procedure for the prediction of future observations. The properties of these models are explored and they are fitted to the Japanese and US series. Finally the scope for using these models for making medium term predictions for Chile is assessed.

2 Trends, cycles and balanced growth

2.1 Univariate models

The *local linear trend* model for a set of observations, $y_t, t = 1, .., T$, consists of stochastic trend and irregular components, that is

$$y_t = \mu_t + \varepsilon_t, \qquad t = 1, ..., T. \tag{1}$$

The trend, μ_t, receives shocks to both its level and slope so

$$\begin{array}{lll} \mu_t & = & \mu_{t-1} + \beta_{t-1} + \eta_t, \qquad \eta_t \sim NID(0, \sigma_\eta^2), \\ \beta_t & = & \beta_{t-1} + \zeta_t, \qquad \zeta_t \sim NID(0, \sigma_\zeta^2), \end{array} \tag{2}$$

where the irregular, level and slope disturbances, ε_t,η_t and ζ_t, respectively, are mutually independent and the notation $NID\left(0, \sigma^2\right)$ denotes normally and independently distributed with mean zero and variance σ^2. If both variances σ_η^2 and σ_ζ^2 are zero, the trend is deterministic. When only σ_ζ^2 is zero, the slope is fixed and the trend reduces to a random walk with drift

$$\mu_t = \mu_{t-1} + \beta + \eta_t. \tag{3}$$

Allowing σ_ζ^2 to be positive, but setting σ_η^2 to zero gives an integrated random walk (IRW) trend, which when estimated tends to be relatively smooth. This model is equivalent to a cubic spine and is often referred to as the '*smooth trend*' model.

The statistical treatment of unobserved component models is based on the state space form (SSF). Once a model has been put in SSF, the Kalman filter yields estimators of the components based on current and past observations. Signal extraction refers to estimation of components based on all the information in the sample. Signal extraction is based on smoothing recursions which run backwards from the last observation. Predictions are made by extending the Kalman filter forward. Root mean square errors (RMSEs) can be computed for all estimators and prediction intervals constructed.

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 maximized by an iterative procedure. The calculations can be done with the STAMP package of Koopman *et al* (2000). Once estimated, the fit of the model can be checked using standard time series diagnostics such as tests for residual serial correlation.

HP filtering can be carried by applying a signal extraction algorithm to a special case of the smooth trend model in which the signal-noise ratio, $q = \sigma_\zeta^2/\sigma_\varepsilon^2$, is set to 1/1600 for quarterly data. Figure 1 shows cycle obtained from HP detrending of quarterly, seasonally adjusted[1] data on GDP for Chile. The

[1] Seasonal adjustment was carried out using the basic X-12-ARIMA option in PcGive

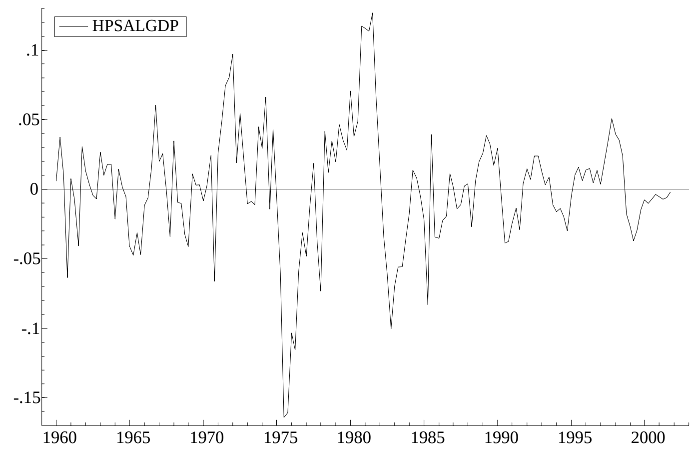

Figure 1: HP filter for Chile GDP (seasonally adjusted)

result is a rather noisy series from which no clear message emerges, particularly towards the end. The HP filter applied to US GDP is more satisfactory in that the business cycle emerges clearly, but again it is not clear what is happening at the end of the series; the HP cycle is very similar to one shown in figure 2

Estimating the parameters of a smooth trend model for GDP will not usually result in a HP cycle as there is nothing in the model to distinguish long-term from short-term movements. Short-term may be captured by including a serially correlated stationary component, ψ_t, in the model. Thus

$$y_t = \mu_t + \psi_t + \varepsilon_t, \quad t = 1, ..., T \tag{4}$$

An autoregressive process is often used for ψ_t, as in Kitagawa and Gersch (1996). Another possibility is the stochastic cycle

$$\begin{bmatrix} \psi_t \\ \psi_t^* \end{bmatrix} = \rho \begin{bmatrix} \cos\lambda_c & \sin\lambda_c \\ -\sin\lambda_c & \cos\lambda_c \end{bmatrix} \begin{bmatrix} \psi_{t-1} \\ \psi_{t-1}^* \end{bmatrix} + \begin{bmatrix} \kappa_t \\ \kappa_t^* \end{bmatrix}, \quad t = 1, ..., T, \tag{5}$$

where λ_c is frequency in radians and κ_t and κ_t^* are two mutually independent white noise disturbances with zero means and common variance σ_κ^2. Given the initial conditions that the vector $(\psi_0, \psi_0^*)'$ has zero mean and covariance matrix $\sigma_\psi^2 \mathbf{I}$, it can be shown that for $0 \leq \rho < 1$, the process ψ_t is stationary and indeterministic with zero mean, variance $\sigma_\psi^2 = \sigma_\kappa^2/(1-\rho^2)$ and autocorrelation function

$$\rho(\tau) = \rho^\tau \cos\lambda_c\tau, \quad \tau = 0, 1, 2, ... \tag{6}$$

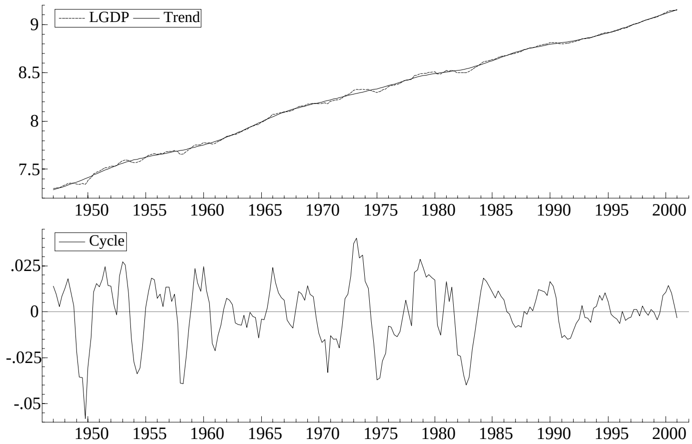

Figure 2: Trend and cycle in US GDP from a structural time series model

For $0 < \lambda_c < \pi$, the spectrum of ψ_t displays a peak, centered around λ_c, which becomes sharper as ρ moves closer to one. The period corresponding to λ_c is $2\pi/\lambda_c$. In the limiting cases when $\lambda_c = 0$ or π, ψ_t collapses to first-order autoregressive processes with coefficients ρ and minus ρ respectively. More generally the reduced form is an $ARMA(2, 1)$ process in which the autoregressive part has complex roots. The complex root restriction can be very helpful in fitting a model, particularly if there is reason to include more than one cycle.

Harvey and Jaeger (1993) showed that extracting a cycle from US GDP using a smooth trend plus cycle model gave a very similar result to the HP filter. This correspondence continues to hold with series shown in figure 2 which is from 1947/1 to 2001/3. The problem at the end of the series is apparent and so the challenge is to devise models which are capable of giving a clearer breakdown into trend and cycle.

2.2 Extracting smoother cycles

The cycle extracted for US GDP in figure 2 comes from a model in which the irregular variance was estimated to be zero. Thus, as with the HP filter, the cycle is the same as the detrended series. A clearer indication of the business cycle might be obtained by a model which manages to force some of the stationary part of the series into the irregular component. The same idea is inherent in the notion of a band-pass filter centred on the business cycle frequencies; see Baxter and King (1999).

The smoothness of a trend depends on the shape of the weighting function - the kernel - for extracting it and the signal-noise ratio. In the local linear trend model of (1), the weighting pattern for a random walk plus drift is a double exponential. For the integrated random walk trend the kernel decays more slowly; some examples can be found in Harvey and Trimbur (2001). Furthermore if the IRW trend model is fitted, the signal-noise ratio is usually smaller than it is for a random walk with drift: the result is a wider bandwidth and a smoother trend. A similar device may be employed for the cycle. To this end we consider a double, or *second-order*, stochastic cycle :

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

$$\begin{bmatrix} \psi_{\beta,t} \\ \psi_{\beta,t}^* \end{bmatrix} = \rho \begin{bmatrix} \cos\lambda_c & \sin\lambda_c \\ -\sin\lambda_c & \cos\lambda_c \end{bmatrix} \begin{bmatrix} \psi_{\beta,t-1} \\ \psi_{\beta,t-1}^* \end{bmatrix} + \begin{bmatrix} \kappa_t \\ \kappa_t^* \end{bmatrix} \tag{7}$$

where κ_t and κ_t^* are as in the first-order cycle, (5), and ρ and λ_c satisfy the same conditions.

General classes of higher order trends and cycles may be defined. A higher order trend will give a nonlinear forecast function and so may not be attractive. On the other hand, there may be merit in higher order cycles. Harvey and Trimbur(2001) define the *nth order stochastic cycle*, for positive integer n, as

$$\begin{bmatrix} \psi_{1,t} \\ \psi_{1,t}^* \end{bmatrix} = \rho \begin{bmatrix} \cos\lambda_c & \sin\lambda_c \\ -\sin\lambda_c & \cos\lambda_c \end{bmatrix} \begin{bmatrix} \psi_{1,t-1} \\ \psi_{1,t-1}^* \end{bmatrix} + \begin{bmatrix} \kappa_t \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} \psi_{i,t} \\ \psi_{i,t}^* \end{bmatrix} = \rho \begin{bmatrix} \cos\lambda_c & \sin\lambda_c \\ -\sin\lambda_c & \cos\lambda_c \end{bmatrix} \begin{bmatrix} \psi_{i,t-1} \\ \psi_{i,t-1}^* \end{bmatrix} + \begin{bmatrix} \psi_{i-1,t} \\ 0 \end{bmatrix}, \quad \text{i} = 2, ..., \text{n} \tag{8}$$

The fact that there is no κ_t^* and $\psi_{i-1,t}^*$ is a matter of convenience in working out properties. It enables us to write

$$\psi_{i,t} = C(L)\psi_{i-1,t}, \quad i = 2, ..., n$$

with $\psi_{1,t} = C(L)\kappa_t$,where

$$C(L) = \frac{1 - \rho\cos\lambda_c L}{1 - 2\rho\cos\lambda_c L + \rho^2 L^2}$$

Repeated substitution yields

$$\psi_{n,t} = [C(L)]^n \kappa_t. \tag{9}$$

The properties of the cycle are most easily expressed in the frequency domain. The power spectrum, for $\rho < 1$, is given directly from the spectral generating function as:

$$\begin{aligned} f_\psi(\lambda; \rho, \lambda_c, n) &= \left|C(e^{-i\lambda})\right|^n \sigma_\kappa^2 / 2\pi \qquad (10) \\ &= \frac{\sigma_\kappa^2}{2\pi} \left[\frac{1 + \rho^2\cos^2\lambda_c - 2\rho\cos\lambda_c\cos\lambda}{1 + \rho^4 + 4\rho^2\cos^2\lambda_c - 4(\rho + \rho^3)\cos\lambda_c\cos\lambda + 2\rho^2\cos 2\lambda} \right]^n \end{aligned}$$

As n increases the shape of the spectrum becomes such that there is relatively less power at high frequencies. If the cycle is embedded in white noise, that is

$$y_t = \psi_{n,t} + \varepsilon_t, \quad \varepsilon_t \sim WN(0, \sigma_\varepsilon^2), \tag{11}$$

the gain function is found to be

$$G(\lambda; \rho, \lambda_c) = \frac{q_\kappa \left[\frac{1+\rho^2\cos^2\lambda_c - 2\rho\cos\lambda_c\cos\lambda}{1+\rho^4+4\rho^2\cos^2\lambda_c - 4(\rho+\rho^3)\cos\lambda_c\cos\lambda + 2\rho^2\cos 2\lambda}\right]^n}{1 + q_\kappa \left[\frac{1+\rho^2\cos^2\lambda_c - 2\rho\cos\lambda_c\cos\lambda}{1+\rho^4+4\rho^2\cos^2\lambda_c - 4(\rho+\rho^3)\cos\lambda_c\cos\lambda + 2\rho^2\cos 2\lambda}\right]^n} \tag{12}$$

where $q_\kappa = \sigma_\kappa^2/\sigma_\varepsilon^2$. The higher is n, the more a block of frequencies around λ_c is passed by the filter. When a model of the form (11) is fitted, the irregular component tends to become bigger as n increases, the signal-noise ratio, q_κ, becomes smaller and the estimated cycle tends to become smoother. Similar conclusions hold if a trend is in the model as in (4).

If $\rho = 1$, Gomez (2001) shows that the signal extraction filter for the cycle is a member of Butterworth class. More generally, Harvey and Trimbur(2001) refer to a filter obtained with $0 < \rho \leq 1$ as a *generalised Butterworth band-pass* filter of order n.

With ρ equal to one the gain becomes more rectangular as n increases and an *ideal band-pass filter* is obtained as a limiting case. Baxter and King (1999) argue for the desirability of ideal band pass filters and suggest how they may be approximated in the time domain by truncating weights beyond a certain lag and then modifying them so they sum to zero. A model containing a higher order cycle can also approximate an ideal band pass filter, but without sacrificing observations at the beginning and end of the series. However, the model suggests that this may be unappealing, one reason being that the cycle is nonstationary. Business cycles are normally thought of as being stationary, so the additional flexibility resulting from the inclusion of the damping factor is an important generalisation.

Fitting a fully specified model consisting of trend, cycle and irregular components, together with any other necessary components, such as a seasonal, yields a filter which is optimal for extracting a cycle with clearly defined properties and which is consistent with the data. The calculations may be programmed in Ox using the Ssfpack set of subroutines documented in Koopman, Shephard and Doornik (1999). Figure 3 shows the cycle extracted from US GDP when $n = 2$. This cycle is smoother than the one shown in figure 2 and, even more importantly, it gives a much clearer indication of the state of the economy at the end of the series. It appears that the US is at the top of a boom and there is a strong indication of a turning point.

2.3 Several cycles: the case of Chile

Fitting the trend plus cycle model to the logarithms of annual data on Chilean GDP, in 1995 pesos, from 1870 to 1995 gives the trend shown in figure 4. The period of the cycle is $P = 1205$ with $\rho = 0.75$ and the signal-noise ratio, q_ζ, being

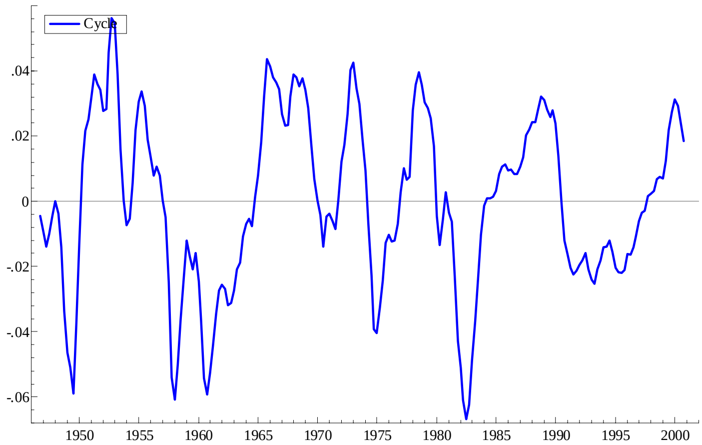

Figure 3: Second-order cycle for US GDP

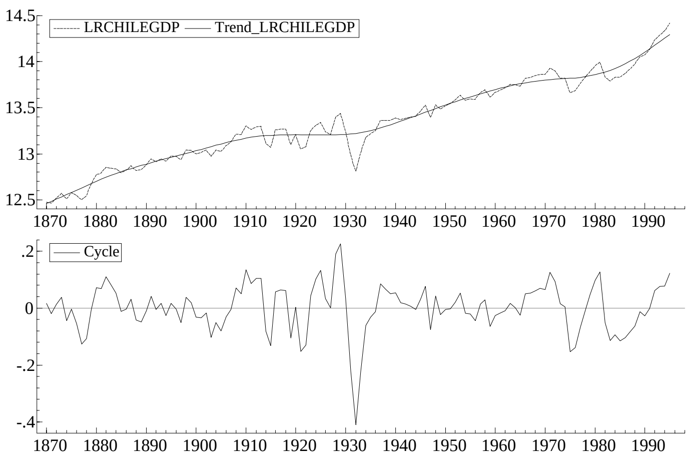

Figure 4:

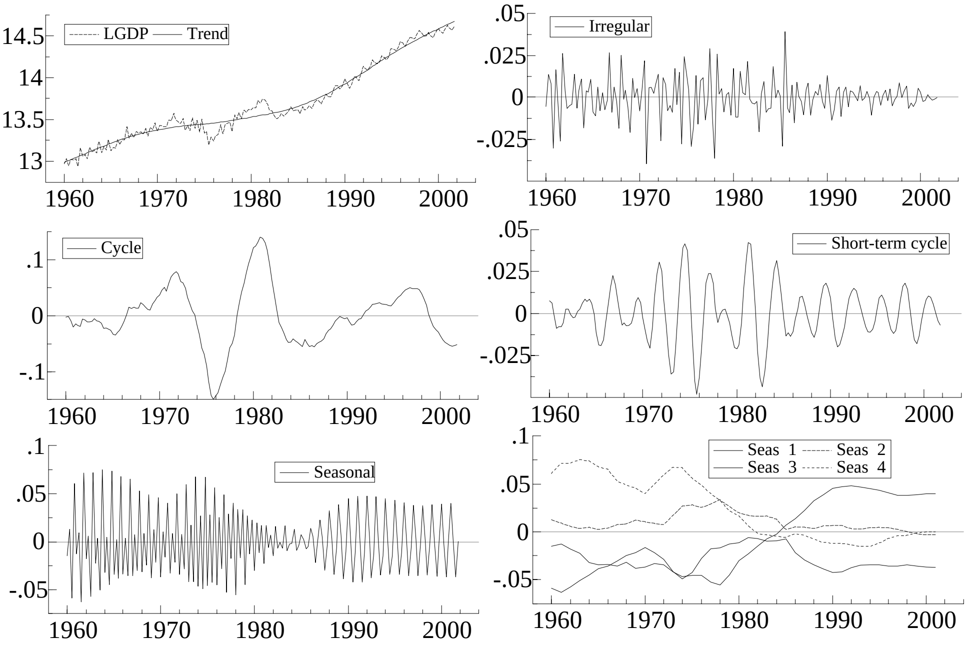

Figure 5: Decomposition of real Chile GDP into trend, two cycles, seasonal and irregular

0.0056. Fitting the same model to the quarterly data set, 1960/1 to 2001/4[2], in real 1986 Chilean pesos, is less successful: the recessions in the 1970s and 1980s are very pronounced and because they so dominate the sample period they become incorporated into the trend, leaving only very short term movements in the cycle. However, estimating a model with two cyclical components solves the problem. The first cycle, which picks up the major recessions, has a period of 10.66 years with $\rho = 0.97$, while the second has $\rho = 0.92$ and a period of just under three years. If one uses the monthly series, from 1982/01 to 2001/07, only the same short term cycle can be extracted; Caputo (2001) argues that this cycle has a meaningful interpretation. As can be seen from figure 1, the HP filter (applied to seasonally adjusted data) is unsatisfactory as it yields a confusing mixture of short and long term cycles together with the noise from the irregular. The Baxter-King filter would be of little help as it normally focuses on frequencies between six and thirty-two quarters.

Figure 5 shows the five components into which the series is decomposed. Of particular note is the fact that the economy is near the trough of the longer term cycle. Figure 6 shows forecasts of the series, with one RMSE on either

[2] The last two quarters are estimates

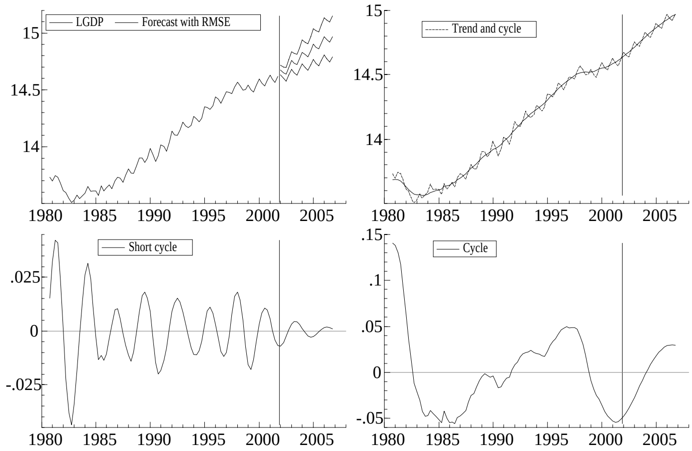

Figure 6: Forecasts for Chile real GDP

side, together with extrapolations of the two cycles.

The series has a seasonal component, which shows marked changes over the period; the graph of individual seasons is particularly informative. The reasons behind these changes in seasonality will not be investigated further here, though it is interesting to ponder on the effect of trying to tackle such movements with a non-model-based seasonal adjustment procedure such as the U.S. Census Bureau's X-12.

3 Multivariate models and balanced growth

More precise information on the target series can sometimes be obtained by bringing in information in a related series. This is done by constructing a bivariate model. For example, the cycle in investment is quite marked and so it may help in giving a better estimate of the cycle in GDP. Several auxiliary series may be used in a multivariate systems, but the principle remains the same. The ideas of co-integration, common trends and balanced growth are directly relevant to the potential gains in efficiency.

3.1 Bivariate structural time series models

The bivariate local level model is

$$\begin{aligned} y_{1t} &= \mu_{1t} + \varepsilon_{1t}, \qquad \mu_{1t} = \mu_{1t-1} + \eta_{1t}, \qquad t = 1, ..., T, \\ y_{2t} &= \mu_{2t} + \varepsilon_{2t}, \qquad \mu_{2t} = \mu_{2t-1} + \eta_{2t} \end{aligned} \tag{13}$$

The covariance matrix of $(\eta_{1t}, \eta_{2t})'$ may be written

$$\Sigma_\eta = \begin{bmatrix} \sigma_{1\eta}^2 & \rho_\eta \sigma_{1\eta} \sigma_{2\eta} \\ \rho_\eta \sigma_{1\eta} \sigma_{2\eta} & \sigma_{2\eta}^2 \end{bmatrix}$$

where ρ_η is the correlation. More generally,

$$y_{it} = \mu_{it} + \psi_{it} + \varepsilon_{it}, \; i = 1, 2, \qquad t = 1, ..., T, \tag{14}$$

where μ_{it} is a local linear trend and ψ_{it} is a cycle as defined earlier.

The *similar cycle* model, introduced by Harvey and Koopman (1997), allows the disturbances driving the cycles to be correlated across the series. However, the damping factor and the frequency, ρ and λ_c, are the same in all series, so the cycles in the different series have similar properties; in particular their movements are centred around the same period. This seems eminently reasonable if the cyclical movements all arise from a similar source such as an underlying business cycle. Furthermore, the restriction means that it is often easier to separate out trend and cycle movements when several series are jointly estimated.

3.2 Stability and balanced growth

In the *balanced growth* model, the same trend, μ_t, appears in the two series. Thus the bivariate local level model becomes

$$\begin{aligned} y_{1t} &= \mu_t + \alpha + \varepsilon_{1t}, \qquad t = 1, ..., T, \\ y_{2t} &= \mu_t + \varepsilon_{2t}, \end{aligned} \tag{15}$$

In terms of (13), $\rho_\eta = 1$ and $\sigma_{1\eta} = \sigma_{2\eta}$. A corresponding property holds for the slope disturbance in the local linear trend.

The series have a stable relationship over time in that they are evolving in such a way that their difference $y_{1t} - y_{2t}$ is stationary. In other words the series are co-integrated with a known co-integrating vector. A stability test of the null hypothesis if a stable relationship can be carried out using a stationarity test, such as the one proposed by Nyblom and Mäkeläinen (1983). Under the null hypothesis, the limiting distribution of the test statistic is Cramér-von Mises. The test can be modified so as to include a nonparametric correction for serial correlation as Kwiatkowski *et al* (1992). Parametric adjustments can also be made. If there are no constant term in (15), that is $\alpha = 0$, the series contain an *identical* common trend. The test statistic is then constructed without the mean subtracted and its asymptotic distribution under the null is then comes from a different member of the Cramér-von Mises family, see Harvey and Carvalho (2001).

The common trend restriction is a strong one, but it can lead to considerable gains in the efficiency with which components in the target series are estimated. An analysis can be found in Harvey and Chung (2000) in connection with the estimation of the underlying change in the level of unemployment. Another point of interest is that the paper demonstrates how state space methods can be used to combine information produced at different sampling intervals. Thus, in the case of the UK, quarterly survey data is combined with monthly claimant count figures to produce a better estimate of the underlying change in unemployment.

3.3 Japan and the U.S.

Models with smooth trends were fitted to the logarithms of quarterly, seasonally adjusted, data on real GDP per capita in the US and Japan over the period 1961:1 to 2000:1.The data were obtained from the OECD Main Economic Indicators and the population series were constructed as quarterly moving averages of annual figures spread over all four quarters. The series are in 1990 US dollars; the choice of conversion date of course affects the gap between the series, but is otherwise irrelevant.

Fitting a univariate model to Japan does not yield a satisfactory cycle[3]. By contrast, it becomes much more like the US cycle in the similar cycle bivariate model. Table 1 shows the estimates of the parameters, obtained using STAMP, together with the standard error (SE) for each equation and the Box-Ljung statistic, $Q(P)$, based on the first P residual autocorrelations. The correlations between the slope, cycle and irregular disturbances were -0.143, 0.274 and 1 respectively. The period of 27.07 quarters corresponds to 6.77 years.

	TABLE 1	*Bivariate*	*Model*
	Hyperparameters	*Japan*	*US*
Trend	$\sigma_\zeta(\times 10^{-3})$	1.638	0.907
Cycle	$\sigma_\kappa(\times 10^{-3})$	7.177	7.642
	$\sigma_\psi(\times 10^{-3})$	17.22	18.34
	ρ	0.91	0.91
	Period $(2\pi/\lambda_c)$	27.07	27.07
Irregular	$\sigma_\varepsilon(\times 10^{-3})$	4.380	0.174
Fit	$\log L$		
	$SE(\times 10^{-3})$	11.144	9.058
Diagnostics	$Q(11)$	11.766	14.719

Figure 7 are for the bivariate model. Their presence means that the trends are quite smooth. However, it is clear that the forecasts will diverge as there is virtually no growth in Japan at the end of the series. This issue is taken up in section 5 where a convergence model is fitted.

4 Models of Converging Economies

Two countries *have converged* if the difference between them is stable. If initial conditions are unimportant, stability implies that the difference between the series, y_t, is stationary for virtually the whole period. If the mean of y_t is zero the countries are in a state of *absolute convergence*. If the mean, α, is not zero we have *conditional* or *relative convergence*. This is a possibility if we entertain

[3]The cycle is almost nonstationary, with $\rho = 0.998$, while the period is only 2.97 years.

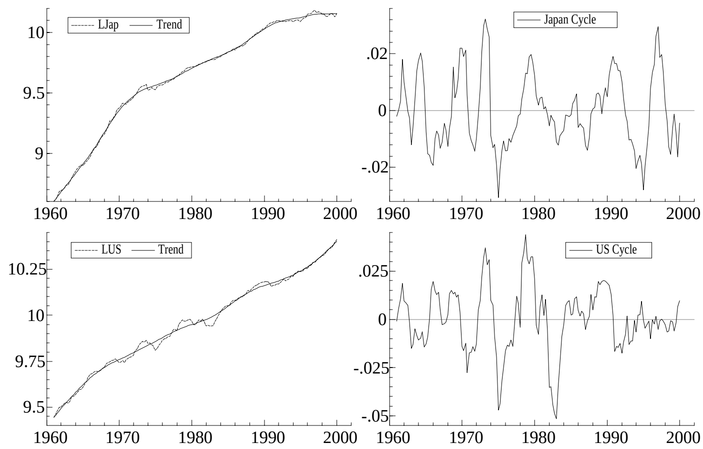

Figure 7: Trends and cycles from a bivariate structural time series model

the existence of increasing costs of convergence and possible barriers to absolute convergence; see, for example, Bernard and Durlauf (1996). The limiting growth paths for the regions are then parallel, differing by α.

Suppose now that we wish to model the process of convergence. If two economies are converging, the model for y_t will have the property that forecasts converge to α. The models set up below are able to satisfy this condition and they become stationary for economies which have converged.

4.1 Stylised facts

Suppose we wish to look at stylised facts without positing a particular mechanism for convergence. The difference, y_t,is assumed to be made up of a stochastic trend or level, μ_t, together with other components such as cycle and irregular as in (4). The smoothed estimates of the trend describe the time path reflecting the long-run difference between the two economies. Simply plotting this time path may be very informative. For example, figure 8 shows the difference in the trend of per capita GDP between the USA and Japan obtained by fitting a smooth trend, that is with σ_η^2 set to zero, plus cycle model using the STAMP package of Koopman et al (2000). We can go further and carry out tests of whether the gap between the two economies has narrowed significantly and/or whether the gap is zero, that is $\mu_T = 0$, indicating that absolute convergence has taken place. The result can be seen from the graph where a confidence interval of two RMSE's is shown. The level in the trend at the end of the sample is 0.230 with a RMSE of 0.032 giving a '$t - value$' of 7.10. Although Japan came close to catching up with the USA in the early 1990's the movement since then has been in the opposite direction.

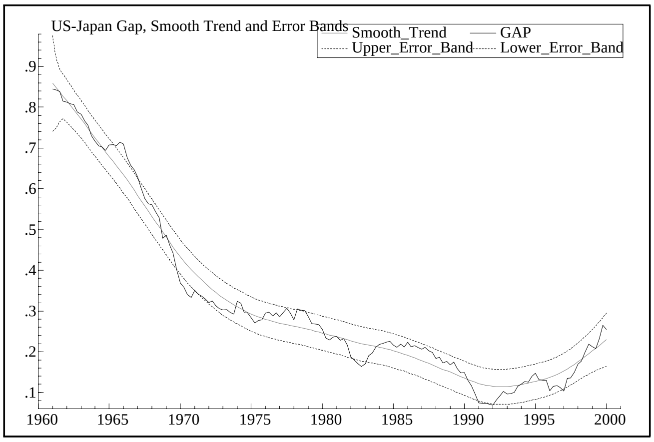

Figure 8: US-Japan gap modelled by a smooth trend

4.2 Error correction mechanisms

The use of non-stationary components to model convergence is apparently contradictory since once convergence has taken place the series are stationary. It is now shown how an error correction mechanism (ECM) can be used to capture convergence dynamics instead of approximating the process by a stochastic trend.

The simplest model is

$$y_t = \alpha + \mu_t, \qquad \mu_t = \phi\mu_{t-1} + \eta_t, \quad t = 1, ..., T, \tag{16}$$

with a fixed initial value, μ_0. The crucial point is that this is not constructed as a model of a stable contrast but rather as a model of transitional dynamics in a situation where the initial value is some way from zero. If $\phi < 1$, the gap tends to narrow over time. It makes little sense to have ϕ negative and so we may assume that $\phi \geq 0$. Of course when the initial conditions have worked themselves out, the series becomes stationary. The equivalent error correction (EC) representation for μ_t is

$$\Delta y_t = (\phi - 1)(y_{t-1} - \alpha) + \eta_t = \delta + (\phi - 1)y_{t-1} + \eta_t, \quad t = 2, ..., T, \tag{17}$$

where $\delta = \alpha(1 - \phi)$. This can be interpreted as saying that, for data in logarithms, the expected growth rate in the current period is a negative fraction of the gap between the two economies after allowing for the permanent difference, α. For example, with $\phi = 0.98$ and a ratio of 1.65 in income per head, which corresponds to a gap in logarithms of 0.5, the difference in growth rates is 1%. Some idea of what different values of ϕ imply about the closing of the gap can be obtained by noting that the $\tau-$step ahead forecast from an AR(1) model is ϕ^τ times the current value. Thus ϕ^τ is the fraction of the gap expected to remain after τ time periods.

Written in the EC form, (17), the model accords with the notion of convergence in the cross-sectional literature, as expounded by Barro and Sala-i-Martin (1992) and others, except that there the growth rate is taken to be a linear function of the initial value, giving a model which is internally inconsistent over time; see Evans and Karras (1996, p 253).

The ECM may be generalised to allow for richer dynamics. Within an autoregressive framework, (17) may be augmented with lagged values of differenced observations. Fitting such a model to the US-Japan series without the constant gave

$$\widehat{\Delta y_t} = -0.0086 y_{t-1} + 0.127\Delta y_{t-1} + 0.083\Delta y_{t-2} + 0.136\Delta y_{t-3} + 0.128\Delta y_{t-4}.$$

The equation standard error, denoted SE (equal to $\widehat{\sigma}_\eta$ here) is 0.0126 and $Q(11)$, the Box-Ljung statistic based on 11 residual autocorrelations, is 7.29; under the null hypothesis of correct specification, the asymptotic distribution of this statistic is χ^2_6. With a constant added to the right hand side

$$\begin{array}{ccccccc} \widehat{\Delta y_t} = & 0.0029 & -0.0156 y_{t-1} & +0.118\Delta y_{t-1} & +0.076\Delta y_{t-2} & +0.133\Delta y_{t-3} & +0.127\Delta y_{t-4}, \\ & (0.0019) & (0.0056) & (0.081) & (0.083) & (0.083) & (0.083) \end{array}$$

with $SE = 0.0125$ and $Q(11) = 6.84$. The estimate of ϕ has fallen from 0.991 to 0.984. The '$t-$statistic' of the constant is 1.54 and the implied value of α is 0.187. None of the lagged differences is statistically significant at the 5% level. With no lags, the estimate of ϕ was 0.979 and the implied value of α was 0.143. However, there was evidence of residual serial correlation with $Q(11) = 25.14$.

4.3 Unobserved components and smooth convergence

The UC approach is to add cycle and irregular components to the error correction mechanism. This avoids confounding the transitional dynamics of convergence with short-term steady-state dynamics. Thus

$$y_t = \alpha + \mu_t + \psi_t + \varepsilon_t, \qquad \mu_t = \phi\mu_{t-1} + \eta_t, \quad t = 1, ..., T. \tag{18}$$

Estimation is effected by using the state space form with a diffuse prior for μ_t (as though it were nonstationary). Although α is regarded as a fixed parameter, it can also be estimated by including it in the state vector with a diffuse prior. Care must be taken as α is not identified when ϕ is unity; it is advisable to carry out numerical optimisation with respect to a transformed variable, such as $-\log(1-\phi)$, which lies between 0 and ∞, thereby keeping ϕ strictly less than one. The appendix explores ML estimation for the simple model in (16). A likelihood ratio test of the null hypothesis that $\alpha = 0$ can be carried out, but in order to ensure comparability of likelihood the one for the unrestricted model must be calculated by treating α as being fixed.

Smoother transitional dynamics can be achieved by specifying μ_t in (18) as

$$\begin{aligned} \mu_t &= \phi\mu_{t-1} + \beta_{t-1}, \quad t = 1, ..., T, \\ \beta_t &= \phi\beta_{t-1} + \zeta_t, \end{aligned} \tag{19}$$

If we write the model with what might be termed a second-order ECM, that is

$$\begin{aligned} \Delta\mu_t &= (\phi - 1)\mu_{t-1} + \beta_{t-1}, \quad t = 1, ..., T, \\ \Delta\beta_t &= (\phi - 1)\beta_{t-1} + \zeta_t, \end{aligned} \tag{20}$$

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. Alternatively this second-order ECM can be expressed as

$$\Delta\mu_t = -(1 - \phi)^2\mu_{t-1} + \phi^2\Delta\mu_{t-1} + \zeta_t$$

showing that the underlying change depends not only on the gap but also on the change in the previous time period. This means that changes take place more slowly. Note that the model is a special case of the second-order cycle, (7), obtained by setting $\lambda_c = 0$.

The model is equivalent to an AR(2) process with both roots equal to ϕ. Obviously the condition for stationarity is $|\phi| < 1$. With a value of ϕ close to one, μ_t will behave in a similar way to the smooth trend shown in figure 8. On the other hand, the first-order ECM behaves rather like a random walk specification and tracks the observations closely, leaving little scope for the addition of short-term non-transitional components. The ACF of the second-order model is

$$\rho(\tau) = [1 + \{(1 - \phi)/(1 + \phi)\}\tau]\phi^\tau,$$

so the decay is slower than in an AR(1) with the same value of ϕ. The $k-$step ahead forecast function, standardised by dividing by the current value of the gap, is

$$f(k) = (1 + (1 - \phi\lambda)k)\phi^k, \qquad k = 0, 1, 2, ..$$

where $\lambda = \mu_{t-1}/\mu_t$. If $\lambda = 1/\phi$, the expected convergence path is the same as in the first order model. Some notion of the average convergence behaviour is obtained by setting λ so that $f(k)$ is the same as the ACF; this implies $\lambda = 2/(1 + \phi)$. However, the most interesting aspect of the second-order model is that if the convergence process stalls sufficiently, the gap can be expected to widen in the short run.

Estimating the first-order UC model, (18), resulted in relatively small values for the cycle and irregular variances. The same thing happened when a random walk trend was fitted in the preliminary model, instead of a smooth trend. The dominance of the transitional component over the cycle and irregular means that the model is not too far from a simple ECM as in (17). The convergence parameter, ϕ, is 0.984 for absolute convergence and 0.977 when α is estimated.

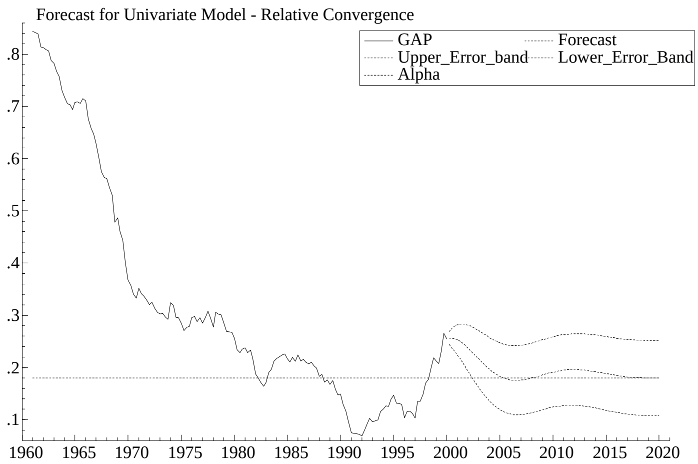

Figure 9: Forecasts for second-order convergence model

The estimate of α is 0.134, but the LR statistic is 3.33 which is not significant against a χ_1^2 distribution.

The second-order convergence model, (20), fared much better insofar as it was able to separate out a cyclical component. The results are shown in table 2. The smoothed path of μ_t is very similar to that shown in figure 8. Figure 9 shows the predictions for the series, y_t, over a twenty year horizon. These predictions show some influence from the cycle. The parameters obtained when the smooth trend was fitted to give figure 8 are shown in the last column. The estimate of α, 0.180, now has a statistically significant LR statistic of 6.46.

	$TABLE\ 2$	$US - Japan$		
	$Hyperparameters$	Absolute	Relative	Trend
$Convergence$	$\sigma_\zeta(\times 10^{-3})$	1.933	1.286	1.244
	ϕ	0.963	0.943	1 (fixed)
$Cycle$	$\sigma_\kappa(\times 10^{-3})$	11.33	11.51	11.25
	ρ_c	0.94	0.96	0.95
	$Period\ (2\pi/\lambda_c)$	50.51	50.14	50.91
$Irregular$	$\sigma_\varepsilon(\times 10^{-3})$	0.014	0.071	1.54
Gap	α	0 (fixed)	0.180	–
Fit	$\log L$	454.679	457.909	451.94
	$SE(\times 10^{-3})$	12.7	12.4	12.8
$Diagnostics$	$Q(11)$	11.54	10.85	9.37

It was argued in section 2 that the higher order cycles of (8) may be more clearly defined in that they cut out more high frequencies. Such cycles could also be used in (18), although they may be more effective in the bivariate models to be described in the next section.

5 Bivariate models for the levels of converging economies

The previous section devised a mechanism for capturing convergence between two economies. This section explores how this mechanism can be incorporated into a bivariate model for the levels of converging economies. The aim is to be able to extract trend and convergence components and to make forecasts which take convergence to a common trend into account. The extension to multivariate modeling is not covered but a discussion can be found in Harvey and Carvalho (2001).

5.1 Bivariate error correction mechanism

A bivariate model for two converging economies can be set up as

$$\begin{aligned}\Delta y_{1t} &= \phi_1(y_{2,t-1} - y_{1,t-1}) + \eta_{1t} \\ \Delta y_{2t} &= \phi_2(y_{1,t-1} - y_{2,t-1}) + \eta_{2t}\end{aligned} \qquad (21)$$

where $y_{i,t}$ denotes, for example, per capita output for economy i at time t. Absolute convergence and no growth is initially assumed for simplicity. Thus the growth rate of the first economy depends on the gap between its level and that of the second economy and vice versa.

The model corresponds to the first-order vector autoregression

$$\begin{aligned}y_{1t} &= (1-\phi_1)y_{1,t-1} + \phi_1 y_{2,t-1} + \eta_{1t} \\ y_{2t} &= \phi_2 y_{1,t-1} + (1-\phi_2)y_{2,t-1} + \eta_{2t}\end{aligned} \qquad (22)$$

The roots of the transition matrix

$$\Phi = \begin{bmatrix} 1-\phi_1 & \phi_1 \\ \phi_2 & 1-\phi_2 \end{bmatrix}$$

are one and $\phi_1 + \phi_2 - 1$. The condition for the second root to lie inside the unit circle is $0 < \phi_1 + \phi_2 < 2$. This being the case, the long-run forecasts converge to the same value since

$$\lim_{k\to\infty} \Phi^k = \begin{bmatrix} \overline{\phi} & 1-\overline{\phi} \\ \overline{\phi} & 1-\overline{\phi} \end{bmatrix} \qquad (23)$$

where $\overline{\phi} = \phi_2/(\phi_1 + \phi_2)$. This is a standard result from the theory of Markov chains.

The model (21) can be premultiplied by a matrix with unit Jacobian thereby transforming it to

$$y_{1t} - y_{2t} = \phi(y_{1,t-1} - y_{2,t-1}) + \eta_{1t} - \eta_{2t}$$

$$\overline{y}_{\phi t} = \overline{y}_{\phi,t-1} + \overline{\eta}_{\phi t}$$

where $\phi = 1 - (\phi_1 + \phi_2)$ and

$$\overline{y}_{\phi t} = \overline{\phi} y_{1t} + (1 - \overline{\phi}) y_{2t}; \tag{24}$$

the disturbance $\overline{\eta}_{\phi t}$ is defined similarly. The first equation corresponds to the univariate convergence equation of (17) since it is an ECM for the difference $y_{1t} - y_{2t}$. In the second equation the weighted sum follows a random walk and, as is clear from (23), this is the growth path to which the two economies are converging.

Parameterising the model in terms of $\overline{\phi}$ and ϕ has some attractions. The stability condition is $|\phi| < 1$, though it makes little sense to have ϕ negative. It seems desirable (though not essential for stability) to have $0 \leq \overline{\phi} \leq 1$. This condition implies that ϕ_1 and ϕ_2 are both greater than or equal to zero. Note that if $\phi = 1$, then $\overline{\phi}$ is not identified.

Benchmark model Setting $\phi_2 = 0$ (or $\phi_1 = 0$) implies that country one (two) converges to country two (one), the benchmark country. Provided ϕ_1 is positive, $\phi_2 = 0$ does not imply a second unit root and so a test of this hypothesis can be based on standard distribution theory. Note that $y_{1,t-1} - y_{2,t-1}$ is stationary (the variables are co-integrated) in (21).

Trend and constant The model may be extended so as to include a common deterministic trend and a constant α to allow for relative convergence. Thus

$$\begin{aligned} y_{1t} &= \alpha + \beta t + \mu_{1t} \\ y_{2t} &= \beta t + \mu_{2t} \end{aligned} \tag{25}$$

where

$$\begin{aligned} \Delta\mu_{1t} &= \phi_1(\mu_{2,t-1} - \mu_{1,t-1}) + \eta_{1t} \\ \Delta\mu_{2t} &= \phi_2(\mu_{1,t-1} - \mu_{2,t-1}) + \eta_{2t} \end{aligned} \tag{26}$$

The gap, $y_{2,t} - y_{1,t}$, is as in (16), except that the sign of α is different (this is more convenient for what follows). Substituting for μ_{1t} and μ_{2t} gives

$$\begin{aligned} \Delta y_{1t} &= \beta - \phi_1\alpha + \phi_1(y_{2,t-1} - y_{1,t-1}) + \eta_{1t} \\ \Delta y_{2t} &= \beta + \phi_2\alpha + \phi_2(y_{1,t-1} - y_{2,t-1}) + \eta_{2t} \end{aligned} \tag{27}$$

Note that the weighted average, (24), is a random walk with a drift of β and that the gap, $y_{2,t} - y_{1,t}$, is as in (17).

5.2 Autoregressive models

The dynamics in (27) may be extended by adding lagged differences to the right hand side of the equations and re-arranging to give

$$\begin{aligned}\Delta y_{1t} &= \delta_1 - \phi_1(y_{1,t-1} - y_{2,t-1}) + \sum_{r=1}^{p} \phi_{11r}^{\dagger} \Delta y_{1,t-r} + \sum_{r=1}^{p} \phi_{12r}^{\dagger} \Delta y_{2,t-r} + \eta_{1t} \\ \Delta y_{2t} &= \delta_2 + \phi_2(y_{1,t-1} - y_{2,t-1}) + \sum_{r=1}^{p} \phi_{21r}^{\dagger} \Delta y_{1,t-r} + \sum_{r=1}^{p} \phi_{22r}^{\dagger} \Delta y_{2,t-r} + \eta_{2t}\end{aligned}$$

where $\delta_i = \beta(1 - \sum_{j=1}^{p}(\phi_{i1j}^{\dagger} + \phi_{i2j}^{\dagger})) + (-1)^i \phi_i \alpha$, $i = 1, 2$. The parameters α and β can be identified from the estimated constants once estimates of ϕ_1 and ϕ_2 have been obtained. The model belongs to the vector error correction mechanism (VECM) class. The co-integrating vector is known and ML estimation can be carried out by OLS since the regressors are the same in each equation. If we were to set α to zero then the restriction that the slopes are the same would need to be enforced.

In the benchmark model, ϕ_i is set to zero in one equation and so β is identified from that equation. Using the estimate of β, an estimate of α parameter can extracted from the estimated constant in the other equation. There should, in theory, be gains from SURE estimation, although in practice it seems to make little difference here.

A bivariate model was estimated for the US and Japan with $p = 4$. For Japan we find $\widetilde{\phi}_1 = 0.0184$ while for the US, $\widetilde{\phi}_2 = -0.0046$. The model is stable, but the negative sign for $\widetilde{\phi}_2$ suggests that it should be set to zero, as in a benchmark model. Indeed the '$t-$statistic' is only 0.937; recall that this is asymptotically standard normal provided ϕ_1 is positive. The benchmark model gave an estimate of ϕ_1 equal to 0.0176, corresponding to $\phi = 0.9824$. From the estimates of the constants, α and β are estimated as 0.140 and 0.0062 respectively. The estimate of β corresponds to an annual growth rate of 2.6%. Recall that the univariate estimate of α from modelling the difference as an autoregression was 0.143.

5.3 Unobserved components

Embedding the ECM within a UC model by adding a cycle and an irregular to (25) gives

$$\begin{aligned}y_{1t} &= \alpha + \beta t + \mu_{1t} + \psi_{1t} + \varepsilon_{1t} \\ y_{2t} &= \beta t + \mu_{2t} + \psi_{2t} + \varepsilon_{2t}\end{aligned} \tag{28}$$

If ψ_{1t} and ψ_{2t} are modelled as similar cycles, subtracting y_{1t} from y_{2t} in (28) gives a univariate model of the form (18).

The vector $(\mu_{1t}, \mu_{2t})'$ may be initialised with a diffuse prior in the SSF. The parameters α and β may also be included in the state and initialised with a diffuse prior, though it order to compare likelihoods they should be treated as fixed. Note that if $\phi_1 = \phi_2 = 0$, then there is no convergence. The pure trend model of section 2 is then obtained provided α is set to zero. However, a balanced growth model is obtained if $\eta_{1,t}$ and $\eta_{2,t}$ are perfectly correlated with the same variance.

A smooth stochastic trend can replace the random walk with common drift. This is most natural if a second-order model for the convergence dynamics, generalising (20), is adopted. We then have

$$\begin{aligned} y_{1t} &= \alpha + \mu_{1t} + \psi_{1t} + \varepsilon_{1t}, \\ y_{2t} &= \mu_{2t} + \psi_{2t} + \varepsilon_{2t}, \end{aligned} \quad (29)$$

$$\begin{aligned} \mu_{1t} &= (1-\phi_1)\mu_{1,t-1} + \phi_1\mu_{2,t-1} + \beta_{1,t-1}, \\ \beta_{1t} &= (1-\phi_1)\beta_{1,t-1} + \phi_1\beta_{2,t-1} + \zeta_{1,t}, \\ \mu_{2t} &= (1-\phi_2)\mu_{2,t-1} + \phi_2\mu_{1,t-1} + \beta_{2,t-1}, \\ \beta_{2t} &= (1-\phi_2)\beta_{2,t-1} + \phi_2\beta_{1,t-1} + \zeta_{2,t}. \end{aligned} \quad (30)$$

Again, if $\phi_1 = \phi_2 = 0$, then there is no convergence but a balanced growth model is obtained if $\zeta_{1,t}$ and $\zeta_{2,t}$ are perfectly correlated with the same variance.

The model can be re-arranged so as to have two convergence components defined in terms of deviations from the common trend, that is $\mu_{it}^{\dagger} = \mu_{it} - \overline{\mu}_{\phi t}$ and $\beta_{i,t}^{\dagger} = \beta_{it} - \overline{\beta}_{\phi t}$, $i = 1, 2$, where $\overline{\mu}_{\phi t}$ is

$$\overline{\mu}_{\phi t} = \overline{\phi}\mu_{1t} + (1-\overline{\phi})\mu_{2t}, \quad (31)$$

and similarly for $\overline{\beta}_{\phi t}$. Then

$$\begin{aligned} y_{1t} &= \alpha(1-\overline{\phi}) + \mu_{1t}^{\dagger} + \mu_{\phi t} + \psi_{1t} + \varepsilon_{1t}, \\ y_{2t} &= -\overline{\phi}\alpha + [-\overline{\phi}/(1-\overline{\phi})]\mu_{1t}^{\dagger} + \mu_{\phi t} + \psi_{2t} + \varepsilon_{2t}, \end{aligned} \quad (32)$$

with

$$\begin{aligned} \mu_{1t}^{\dagger} &= \phi\mu_{1t-1}^{\dagger} + \beta_{1,t-1}^{\dagger}, \\ \beta_{1t}^{\dagger} &= \phi\beta_{1,t-1}^{\dagger} + \zeta_{1,t}^{\dagger}, \\ \mu_{\phi t} &= \mu_{\phi,t-1} + \beta_{\phi,t-1}, \\ \beta_{\phi t} &= \beta_{\phi,t-1} + \zeta_{\phi,t}, \end{aligned}$$

The second convergence component can be obtained from the first since $\overline{\phi}\mu_{1t}^{\dagger} + (1-\overline{\phi})\mu_{2t}^{\dagger} = 0$. Both economies converge to the growth path of the common trend, except insofar as the first economy is at a constant level, α, above (or below) the second one. Convergence is at the same rate, ϕ.

If the second economy is taken to be a benchmark then $\phi_2 = 0$ in the last two equations of (30). In this case μ_{2t} is a smooth trend. Setting up the model as in (32) with $\overline{\phi} = 0$, focuses attention on the transitional gap between the two economies as $\mu_{1t}^{\dagger} = \mu_{1t} - \mu_{2t}$ and $\beta_{1,t}^{\dagger} = \beta_{1,t} - \beta_{2,t}$. The implied model for $y_{1t} - y_{2t}$ is as in (18) with μ_t replaced by $\mu_{1t}^{\dagger}$.

5.4 UC Model for Japan and US

The smooth stochastic trends model fitted in sub-section 3.3 gives an indication of the kind of results which might be expected from a convergence model and can provide starting values for some of the parameters. As already noted, the model is a limiting case which results when $\phi_1 = \phi_2 = 0$ and $\alpha = 0$.

TABLE 3	Hyperparameters	Japan Absolute	US	Japan Relative	US
Convergence	$\sigma_\zeta(\times 10^{-3})$	1.466	0.989	1.399	1.007
	ϕ	0.969		0.958	
Cycle	$\sigma_\kappa(\times 10^{-3})$	6.932	7.464	9.413	7.681
	ρ	0.903		0.892	
	$Period(2\pi/\lambda_c)$	24.77		28.67	
Irregular	$\sigma_\varepsilon(\times 10^{-3})$	4.482	0.521	0	0
Gap	α	$0(fixed)$		-0.174 (0.047)	
Fit	$\log L$	988.050		988.766	
	$SE(\times 10^{-3})$	8.9	1.06	8.9	10.6
Diagnostics	$Q(11)$	12.37	14.77	12.37	14.77

The results of fitting the bivariate convergence model are shown in table 3. The model was estimated with the US taken as the benchmark, with α set to zero and α unrestricted. When the more general model with no restrictions on ϕ_1 and ϕ_2 was estimated it collapsed to the benchmark model. This is consistent with what was found when the bivariate autoregressive model was fitted.

The main features are:

i) The cycle parameters are similar to those obtained with the bivariate pure trend model reported in table 1 and the fitted cycles seem to provide a more satisfactory decomposition than was obtained for the univariate model for the difference.

ii) The estimate of α is only slightly smaller than the one obtained in the univariate gap model. Again there is clear evidence of relative convergence, though the LR statistic is only 1.432.

iii) The estimated convergence component, $\mu_{1t}^{\dagger}$, assigned to Japan, is very similar to the smoothed gap shown in figure 8.

iv) Figure 10 shows the forecasts for the two countries. It can be seen that they converge to the same growth path, μ_t, but at a constant distance , α, apart. A value of $\alpha = -0.174$ implies that the level of Japanese per capita GDP is about 16% below that of the US.

5.5 Chile and US

The difference between Chilean and US GDP is characterised by an enormous swing in favour of the US during the 60s and 70s followed by an equally strong movement in favour of Chile. This makes modelling any kind of convergence

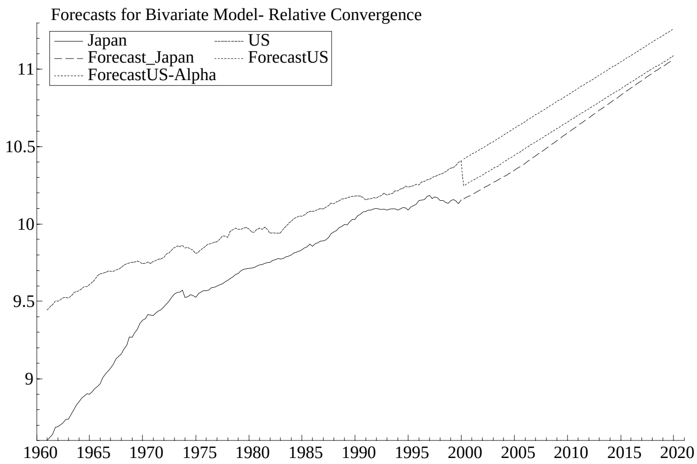

Figure 10:

process extremely difficult. The difficulty is compounded by the fact that the cyclical processes in the two countries have little common. In the case of Chile, the structural time series model for quarterly per capita GDP is virtually the same as the one fitted to GDP in section 2. Extracting a trend and then subtracting from the US trend yields the pattern shown in figure 11. Both series are in 1986 US dollars. The forecasts are simply extrapolations made using the smooth trend model[4]

6 Conclusion

This article has described an extension to the class of structural time series models which allows more clearly defined cycles to be extracted from economic time series. This was illustrated with US GDP. The attraction of this model-based approach is that the filters implicitly defined by the model are consistent with each other and with the data. Furthermore they automatically adapt to the ends of the sample and, if desired, root mean square errors can be calculated. The models can also be used to gain insight into the more *ad hoc* filters used in business cycle analysis, indicating when it might be appropriate to use them and when they can lead to serious distortions of the kind which can arise for the HP filter and band pass filters. The preferred model for Chilean GDP has two cycles, both of which have a direct and meaningful interpretation in terms of economic activity. This decomposition could not have been achieved by an *ad hoc* filter.

[4]Fitting a smooth trend model results in only the slope disturbance being non-zero. This is not surprising since the series is constructed from two estimated smooth trends.

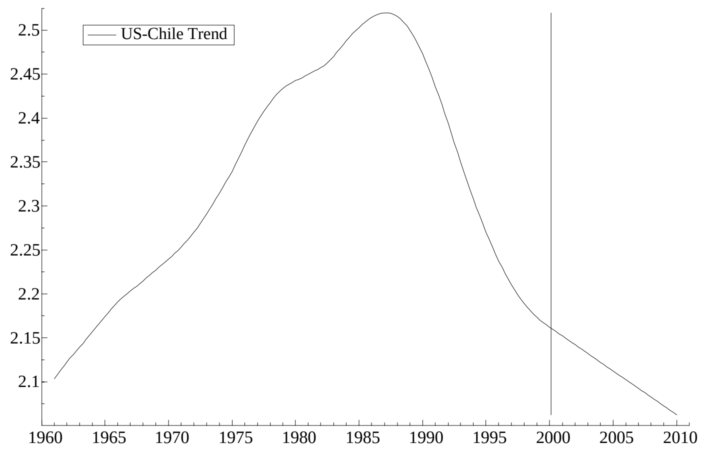

Figure 11: Difference in underlying levels of US and Chilean real per capita GDP (logarithms)

Bivariate structural time series models allow the information on another series to be taken into account in order to extract better information from a target series. Joint modelling of different countries may also be useful. A bivariate model of Japanese and US GDP was shown to give a more informative decomposition of Japanese GDP. The model used was subsequently developed to include a convergence mechanism. This yielded more coherent forecasts for the levels of GDP in the two countries.

ACKNOWLEDGEMENTS

I am grateful to Vasco Carvalho for research assistance and to Rodrigo Caputo and Thomas Trimbur for helpful discussions. Much of the work reported was supported by the ESRC as part of a project on Dynamic Factor Analysis, grant number L138 25 1008.

References

Barro, R.J. and Sala-i-Martin (1992). Convergence. *Journal of Political Economy,* ***100,*** 223-51.

Baxter, M. and R.G.King (1999). Measuring business cycles: approximate band-pass filters for economic time series. *Review of Economics and Statistics,* 81: 575-93.

Bernard, A.B. and S. Durlauf. (1996). Interpreting tests of the convergence hypothesis. *Journal of Econometrics,* ***71,*** 161-73

Caputo, R. (2001). Inflation Targeting in small open economies: the Chilean experience. Mimeo, Cambridge

Cogley, T. and J.M.Nason (1995). Effects of the Hodrick-Prescott filter on trend and difference stationary time series: implications for business cycle research. *Journal of Economic Dynamics and Control,* 19, 253-78.

Evans, P and G. Karras (1996). Convergence revisited. *Journal of Monetary Economics,* ***27,*** 249-65

Gomez, V. (2001). The use of Butterworth filters for trend and cycle estimation in economic time series. *Journal of Business and Economic Statistics, 19, 365-73.*

Harvey, A.C. and V.Carvalho (2001). Models for converging economies, Mimeo, Cambridge.

Harvey, A.C., and C-H. Chung (2000). Estimating the underlying change in unemployment in the UK (with discussion), *Journal of the Royal Statistical Society, Series A,* 163: 303-39.

Harvey, A.C., and A. Jaeger (1993). Detrending, stylised facts and the business cycle. *Journal of Applied Econometrics* 8: 231-47.

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

Harvey, A.C. and T. Trimbur (2001). General model-based filters for extracting cycles and trends in economic time series. *DAE discussion paper, 0113. Cambridge.*

Hodrick, R.J. and E.C.Prescott (1997), Postwar US business cycles: an empirical investigation, *Journal of Money, Credit and Banking,* **24,** 1-16.

Kitagawa, G., and W. Gersch (1996). *Smoothness priors analysis of time series*. Berlin: Springer-Verlag.

Koopman, S.J., N. Shephard and J. Doornik (1999). Statistical algorithms for models in state space using SsfPack 2.2. *Econometrics Journal,* **2,** 113-66.

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

Kwiatkowski, D., Phillips, P.C.B, Schmidt, P. and Y.Shin (1992), Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series have a unit root ? *Journal of Econometrics* 44, 159-78.

Murray, C.J. (2001). Cyclical properties of Baxter-King filtered time series. Mimeo

Nyblom, J. and T. Mäkeläinen (1983), Comparison of tests for the presence of random walk coefficients in a simple linear model, *Journal of the American Statistical Association,* ***78****, 856-64.*