COMMON TRENDS AND CYCLES IN EUROPEAN INDUSTRIAL PRODUCTION: EXCHANGE RATE REGIMES AND ECONOMIC CONVERGENCE*

by

TERENCE C. MILLS

and

MARK J. HOLMES†

Loughborough University

We analyse monthly data on six European industrial production series to ascertain the presence of common cycles and trends. Particular attention is paid to the exchange rate regime, with the Bretton Woods and Exchange Rate Mechanism (ERM) regimes being analysed separately. We employ recently developed techniques in vector autoregressive (VAR) modelling (a) to test for and estimate both common trends and common cycles, (b) to estimate VAR models subject to these ‘common feature’ restrictions, (c) to compare and contrast these models with those obtained from an alternative approach, that of estimating unrestricted levels VAR models, and (d) to present a permanent–transitory decomposition of the series that is based on the common trends found in the systems. We find limited evidence of convergence during the ERM, and this is of a long-run nature. In the short run, asymmetric shocks seem to have produced a divergence compared with the earlier Bretton Woods regime. There is also some evidence of German ‘leadership’ over Belgium, France and the Netherlands.

1 INTRODUCTION

In recent decades, the economies of the European Union (EU) have participated in a number of exchange rate regimes aimed at promoting the stability of their nominal exchange rates. The Bretton Woods period, which lasted from 1946 to 1971, saw European rates pegged against the US dollar, while the Exchange Rate Mechanism (ERM), which has lasted since 1979, has aimed more specifically at stabilizing nominal exchange rates within the EU. Unlike the Bretton Woods regime, the ERM period has witnessed the relaxation of capital controls and of restrictions on the mobility of goods and labour inside the EU, and is seen as a precursor to eventual European Monetary Union (EMU). The extent to which national economies already exhibit some degree of convergence at the onset of EMU will obviously have implications for the extent of economic

* Manuscript received 23.1.97; final version received 21.1.98.
† We would like to thank two referees for helpful comments on an earlier draft. The usual disclaimer applies, of course.
adjustment that will be required as EMU proceeds. An interesting question, therefore, is whether, and to what extent, the ERM has produced an increase in economic convergence across the EU.

We address this question by examining indices of industrial production (IIP) for a sample of major EU members during the ERM period and comparing their evolution and interaction with a similar examination of the earlier Bretton Woods era. In particular, we analyse the common features in European IIP to assess the evidence on convergence and we investigate the causal relationships between indices during both regimes. Following Engle and Kozicki (1993), Engle and Issler (1995) and Vahid and Engle (1993), common features are taken to include both common trends (co-movements among non-stationary variables) and common cycles (co-movements among stationary variables), and we use the models so developed to investigate permanent–transitory decompositions of IIP using the approach of Gonzalo and Granger (1995). As the concept of convergence can be rather loose, we adopt the formal definition provided by Bernard and Durlauf (1995), which is a natural one in our modelling framework as it implies the presence of a single common trend. Causality is investigated both within the now standard vector error correction model (VECM) framework, in which we explicitly test for the number of common trends, common cycles and unit roots that exist in our IIP systems, and within a restricted levels vector autoregressive (VAR) framework. These tests enable us to focus on the nature of the interrelationships between the European economies in the two periods: in particular, we consider whether there is evidence of increased German economic leadership during the ERM period.

Such a study is of interest for a variety of reasons. First, our search for common cycles throws light on the literature on international business cycles (see, inter alia, Ahmed et al., 1993). Studies by Baxter and Stockman (1989) and Gerlach (1988) examine the international synchronization of business cycles and find, in comparisons between Bretton Woods and the subsequent float with the US dollar, that synchronization weakened during the regime of floating rates. The methodologies employed by these studies concentrate on looking at cross-country correlations, and do not allow for the persistence of shocks or for co-movement. Common features analysis allows us to do this. Second, unlike earlier studies of common trends and cycles, our investigation is conducted within the context of clearly defined exchange rate regimes. Serletis and Krichel (1992), Engle and Kozicki (1993) and Calcagnini (1995) examine common trends and cycles in national outputs, but do not explicitly account for the changes in the exchange rate regime that occurred during their sample periods. If output relationships are sensitive to the prevailing exchange rate regime, these studies may be offering findings which are too general to be meaningful. Third, although our empirical investigation is, in part, based
on the integration of common trends and common cycles analysis, we also provide complementary evidence obtained from a general VAR framework that avoids pre-testing and the imposition of, perhaps contentious, restrictions. Fourth, by analysing real convergence, our study augments earlier work on nominal convergence and enables us to assess whether earlier findings on interest rates and inflation (see, for example, Koedijk and Kool, 1992; Hafer and Kutan, 1994) have any relevance for real output.

The structure of the paper is as follows. In Section 2 we consider the literature on economic integration and some theoretical issues. In Section 3 we discuss the data and present some preparatory statistical analysis. The common trends and cycles methodology is then outlined in Section 4. These techniques are used to analyse our two samples of IIP data in Section 5. The analysis is extended in Sections 6 and 7, where we investigate levels VAR estimation of the systems and the decomposition of the indices into permanent and transitory components, respectively. Section 8 draws together the implications of our results and presents conclusions.

2 Economic Integration in the EU

A key consideration in the issue of whether national outputs are integrated concerns the nature and stance of individual policy. Economic theory has argued that a fixed exchange rate regime combined with perfect capital mobility and asset substitutability will restrict the scope for an autonomous monetary policy (see, most notably, Fleming (1962) and Mundell (1963), who address this issue within the context of the IS–LM model). Recent empirical studies of the EU by Koedijk and Kool (1992), Hafer and Kutan (1994) and Katsimbris and Miller (1995) test the somewhat looser hypothesis that a regime of relatively fixed exchange rates (the ERM) will promote financial integration among participating members. By examining covariation among interest rates through principal component and cointegrating techniques, they conclude that EU monetary policies are fairly interdependent but that there is some limited scope for independent policy and that notions of German leadership can largely be rejected. Similar conclusions with regard to inflation convergence are reached by Hall et al. (1992), Koedijk and Kool (1992), Caporale and Pittis (1993) and Thom (1995), all of whom suggest that the ERM may have instilled a degree of inflation discipline among its members. With the loss of monetary policy as a policy instrument, Tavlas (1993) argues that the ERM has increased output variability in general, so that, in response to a given shock, more emphasis is now placed on output adjustment rather than price changes.

© Blackwell Publishers Ltd and The Victoria University of Manchester, 1999.
We explicitly examine output linkages in the form of common trends and cycles. Common trends are viewed in the context of cointegrating relationships between EU outputs. Factors which are relevant in generating common trends are associated with long-run structural and institutional characteristics of the individual EU economies. Evidence of a single common trend would suggest that the EU economies have structural and institutional characteristics of such similarity that outputs converge in the long run. Key structural differences between the two exchange rate regimes concern the mobility of goods, labour and capital, which have been greater during the ERM period. For this reason, we might expect fewer common trends to be uncovered in the ERM model. On the basis of analysing common trends, Hafer and Kutan, Caporale and Pitis and Thom all conclude that the ERM has provided partial convergence of interest rates and inflation, in the sense that they find evidence of cointegration but with more than one single common trend. Common cycles, on the other hand, are short-run phenomena linking stationary series. Shocks to an economy and their effects on output growth are of relevance here. Following Emerson et al. (1992), we may categorize shocks according to whether they are, on the one hand, common or country-specific, and, on the other, permanent or temporary. By definition, country-specific shocks are asymmetric. However, a common shock can, in effect, become asymmetric if a particular country has idiosyncrasies in terms of its structure or in its means of dealing with the shock. For example, the oil price rise in 1979 was a common shock, but the UK, being a net exporter of oil, was affected differently from the other EU members, who were net importers. Bayoumi (1992) finds that a comparison between the ERM and the 1970s floating rate regime yields no significant difference in the nature of the shocks hitting the ERM economies but, rather, key differences in the responses to given shocks. That study, however, was based on a restrictive model which does not allow demand-side shocks to exert a long-run influence on output.

Given that the presence of a single common cycle can be construed as evidence that output variability across the EU depends upon the same kind of shocks everywhere, our enquiry examines short-run integration in the context of searching for common cycles. Again, we might expect that the ERM, with relaxed capital controls, has restricted the use of autonomous monetary policy in response to given shocks, thus increasing the likelihood that there are fewer common cycles compared with the Bretton Woods period. However, De Grauwe (1994) argues that the ERM may actually have exacerbated economic cycles for some members because responses to common or country-specific exchange rate crises are typically based on asymmetric adjustment on the part of the weaker currency economies.

The final part of our enquiry is to examine causality among EU members in both regimes. As mentioned above, studies of interest rates...
and inflation have suggested that the ERM is based on interdependence rather than leadership from Germany. Our enquiry, therefore, considers whether this translates into economic interdependence: can we identify a ‘core’ grouping? A further question is whether we can identify certain countries which exhibit relatively more economic independence: the UK, for example, which has had only limited participation in the ERM.

3 The Data

The data that we use are monthly observations on the logarithms of seasonally adjusted IIP for six European countries: Belgium, France, Germany, the Netherlands, Italy and the UK. Two sample periods are analysed in detail: the ‘Bretton Woods’ period from 1960.01 to 1971.08, and the ‘ERM’ period from 1979.03 to 1994.12 (the starting date for the Bretton Woods sample is dictated by data availability). While these periods have been selected as regimes of relatively fixed exchange rates, a number of caveats should be borne in mind. First, in many countries some form of exchange controls has existed at various times. For example, the UK abolished capital controls in 1979, Germany had lifted all restrictions on capital inflows by 1981, and France and Italy, although ERM members from the outset, have used exchange controls to protect their exchange rates (see, for example, Kenen, 1995). These controls have been gradually relaxed over the period of the ERM, with the removal of all controls by January 1991. A second caveat is that neither Bretton Woods nor the ERM have been periods of rigidly fixed exchange rates. Bretton Woods allowed currencies to fluctuate within a 1 per cent par value in terms of gold, and furthermore the UK devalued in 1967. During the ERM, the permitted fluctuations in exchange rates were set at \( \pm 2.25 \) per cent around the central band. In addition, there were a number of ERM realignments, particularly up to January 1987, and, of course, September 1992 saw the exit of Italy and the UK, with the subsequent widening of the permitted bands of exchange rate fluctuation for the remaining members to \( \pm 15 \) per cent.¹

On initial examination of the data, it was found that some of the series contained occasional outliers, notably Belgium for 1960.11 and 1960.12, France for 1968.05 and 1968.06, and Germany for 1984.06, 1994.07 and 1994.08, most of which were a consequence of known external events. Outliers have been found to adversely affect unit root and cointegration tests, producing spurious evidence of stationarity (Franses and Haldrup, 1994). Since our analysis makes heavy use of such tests,

¹There were major realignments in the ERM in October 1981, February 1982, March 1983, April 1986 and January 1987. Many of these realignments resulted from asymmetric exchange rate crises: see Artis (1990).
these outlying observations were adjusted using the methodology of Gómez and Maravall (1994), and the adjusted data were used in all subsequent modelling.

Figures 1(a) and 1(b) present plots of the adjusted data for the Bretton Woods and ERM periods respectively. The much closer evolution of the IIP series in the ERM period is clearly evident: the ‘swirl’ of data effectively precludes the identification of individual series, unlike the Bretton Woods data, where clear divergences are apparent. In both periods there is evidence of non-stationarity for all six countries, and the unit root tests presented in Table 1 confirm that all series can be regarded as I(1). Interestingly, unit root tests on the unadjusted Belgium and French series in the Bretton Woods period were significant at the 5 per cent and 1 per cent levels respectively, the test statistics being $-3.50$ and $-4.66$, and thus are consistent with Franses and Haldrup’s analysis.

4 Common Trends and Common Cycles Analysis

We begin our formal modelling by denoting as $y_t = (y_{1t}, y_{2t}, \ldots, y_{nt})$, $t = 1, \ldots, T$, an $n$-vector of I(1) time series, which we assume has a $p$th order VAR representation:

$$A(L)y_t = \mu + \epsilon_t \quad (1)$$

where $L$ is the lag operator

$$A(L) = I + A_1 L + A_2 L^2 + \ldots + A_p L^p$$

and $\epsilon_t$ is iid$(0, \Sigma_p)$ with $\Sigma_p$ positive definite. The VAR($p$) can be written using first differences

$$\Delta y_t = B(L)\Delta y_{t-1} + \Pi y_{t-1} + \mu + \epsilon_t \quad (2)$$

where

$$B(L) = B_1 + B_2 L + \ldots + B_{p-1} L^{p-2}$$

$$B_i = - \sum_{j=i+1}^{p} A_j \quad i = 1, \ldots, p - 1$$

$$\Pi = A(1)$$

If there is no cointegration, $\Pi$ will be a zero matrix and (2) becomes a VAR($p-1$) in the differences $\Delta y_t$. When the series are cointegrated, $\Pi$ will have rank $r < n$ and can be decomposed as $\Pi = \beta z'$, where both $z$ and $\beta$ are of rank $r$. (2) can then be written as the VECM

$$\Delta y_t = B(L)\Delta y_{t-1} + \beta z_{t-1} + \mu + \epsilon_t \quad (3)$$

where $z_t = z'y_t$, $z'$ being the $r \times n$ matrix containing the $r$ linearly independent cointegrating vectors (Granger representation theorem: Engle

© Blackwell Publishers Ltd and The Victoria University of Manchester, 1999.
Fig. 1 European Industrial Production: (a) Bretton Woods; (b) ERM
and Granger, 1987). The matrix \( \mathbf{A} \) is not unique, and one way in which it can be identified is as \( \mathbf{A} = \mathbf{I} - \mathbf{L} \), which is known as Phillips's (1991) triangular representation. Since \( z_t \) is I(0), the I(1) components of \( y_t \) can be expressed as linear functions of \( k = n - r \) common stochastic trends gathered together in the \( k \)-vector \( \tau_t \):

\[
y_t = \gamma \tau_t + c_t \\
\tau_t = \tau_{t-1} + \delta e_t
\]

where

\[
\tau_t = \mu^* + \delta \sum_{s=0}^{\infty} e_{t-s} \\
c_t = C^*(L)e_t
\]

This ‘common trends’ representation, due originally to Stock and Watson (1988), is obtained from the Wold representation of the stationary \( \Delta y_t \)

\[
\Delta y_t = C(L)e_t
\]

where

\[
C(L) = \sum_{i=0}^{\infty} C_i L^i = \Delta A^{-1}(L) \\
C_0 = I
\]

Since we can always express \( C(L) \) as (see Banerjee et al., 1993, pp. 140–145)

\[
C(L) = C(1) + \Delta C^*(L)
\]

where

\[ \text{Table 1}

<table>
<thead>
<tr>
<th>Series</th>
<th>Bretton Woods</th>
<th>ERM</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( p )</td>
<td>( \text{ADF}(p) )</td>
</tr>
<tr>
<td>Belgium</td>
<td>2</td>
<td>-2.36</td>
</tr>
<tr>
<td>France</td>
<td>2</td>
<td>-2.76</td>
</tr>
<tr>
<td>Germany</td>
<td>3</td>
<td>-1.96</td>
</tr>
<tr>
<td>Netherlands</td>
<td>1</td>
<td>-1.76</td>
</tr>
<tr>
<td>Italy</td>
<td>0</td>
<td>-3.38</td>
</tr>
<tr>
<td>UK</td>
<td>2</td>
<td>-1.43</td>
</tr>
</tbody>
</table>

Notes: \( p \) is the order of the autoregression in the model \( \Delta y_t = z_t + z_t \mathbf{A} + \phi_1 \Delta y_{t-1} + \ldots + \phi_k \Delta y_{t-k} + \epsilon_t \) and \( \text{ADF}(p) \) is the augmented Dickey-Fuller statistic (the \( t \) ratio on \( \epsilon_t \)). Following the simulation evidence presented in Ng and Perron (1995), \( p \) was determined by sequential testing rather than by an information criterion, as the latter can induce serious size distortions to the test, which persist as sample size increases. None of the statistics is significant at the 5 per cent level, the critical value of which is -3.44 here.

\[ \text{The Manchester School} \]

© Blackwell Publishers Ltd and The Victoria University of Manchester, 1999.
equation (5) can be written as
$$\Delta y_t = C(1)\varepsilon_t + \Delta C^*(L)\varepsilon_t$$
which, on integrating, becomes
$$y_t = \mu + C(1)\sum_{h=0}^{\infty} \delta_{t-h} + C^*(L)\varepsilon_t$$ (6)

If there is cointegration then $C(1)$ is of reduced rank $k = n - r$ and can be written as the product $\gamma\delta'$, both of which have rank $k$, from which the common trends representation (4) immediately follows, with $\mu' = \mu/\gamma$.

With this cointegration framework, we can introduce a formal definition of convergence in a stochastic environment. Following Bernard and Durlauf (1995, Definition 2.1'), we say that $(y_{1t}, y_{2t}, \ldots, y_{nt})$ converge if the long-term forecasts of each series are equal at a fixed time $t$:
$$\lim_{h \to \infty} E(y_{nt+h} - y_{kt+h} | y_t, y_{t-1}, \ldots) = 0 \quad k \neq n$$

This definition of convergence asks whether the long-run forecasts of the differences from a benchmark series tend to zero as the forecast horizon tends to infinity, and it will be satisfied if the set of ‘benchmark differences’ $y_{nt} - y_{kt}$ are all stationary processes (since the IIP series are indices having different bases, we may allow the benchmark differences to have non-zero means). Given that each $y_{kt}$ is I(1), this is equivalent to requiring $y_{kt}$ and $y_{nt}$ to be cointegrated with cointegrating vector $[1, -1]$, i.e. that there are $r = n - 1$ cointegrating vectors with a triangular representation that has $A = -i$, where $i$ is an $n - 1$ unit vector. Bernard and Durlauf’s (1995) weaker definition (2.2') of convergence is the presence of common trends as defined above.

In the same way that common trends appear in $y_t$ when $C(1)$ is of reduced rank, common cycles appear if $C^*(1)$ is of reduced rank, since $c_t = C^*(L)\varepsilon_t$ is the cyclical component of $y_t$. By definition, $x'y_t$ is I(0) and thus does not contain stochastic trends. From (6) we have, since $x'C(1) = 0$ under cointegration (Banerjee et al., 1993, p. 148),
$$x'y_t = x'\mu + x'C^*(L)\varepsilon_t = x'\mu + x'c_t$$ (7)
so that the cointegrating combinations are themselves linear combinations of the cyclical parts of $y_t$. Vahid and Engle (1993) then consider whether there are other linear combinations of the elements of $y_t$ that do not contain these cyclical parts, i.e. whether there is a set of $s$ linearly independent vectors, gathered together in the $n \times s$ matrix $\phi$, such that
$$\phi'c_t = \phi'C^*(L)\varepsilon_t = 0$$
They show that such a matrix exists if all \( C_i \) have less than full rank and if \( \phi C_i = 0 \) for all \( i \). If it does then, from (6),

\[
\phi' y_i = \phi' \mu' + \phi' C(1) \sum_{i=0}^{r} \epsilon_{i+1} = \phi' \gamma \tau_i
\]

Vahid and Engle then show that, under these circumstances, \( C_i = G \tilde{C}_i \) for all \( i \), where \( G \) is an \( n \times (n-s) \) matrix having full column rank and \( \tilde{C}_i \) may not have full rank. The cyclical component can then be written as

\[
c_i = G \tilde{C}(L) \epsilon_i \equiv G \tilde{c}_i
\]

so that the \( n \)-element cycle \( c_i \) can be written as linear combinations of an \( m = n-s \) cycle \( \tilde{c}_i \), thus leading to the common trend–common cycle representation:

\[
y_i = \gamma \tau_i + G \tilde{c}_i
\]

The number \( s \) of linearly independent ‘cofeature’ vectors making up \( \phi \) can be at most \( n-r \), and these will be linearly independent of the cointegration vectors making up \( \phi \) (Vahid and Engle, 1993, Theorem 1). We might refer to the presence of cofeature vectors of this type as representing a form of ‘short-run’ convergence, to distinguish it from the ‘long-run’ convergence given by the common trends.

This common trend–common cycle representation depends, of course, on the number of cointegrating and cofeature vectors, \( r \) and \( s \), in the system. The number of cointegrating vectors (common trends) can be determined by, for example, the test procedures of Johansen (1988, 1991) and Stock and Watson (1988). The number of cofeature vectors (common cycles) can be determined using the approach of Engle and Kozicki (1993), as extended by Vahid and Engle (1993) to the current context in which there may also be cointegration. To find the rank of the cofeature vector \( \phi \), we consider the VECM equation (3), written more concisely as

\[
\Delta y_i = W_i B + \epsilon_i
\]

where

\[
W_i = (\Delta y_{i-1}, \ldots, \Delta y_{i-p+1}, z_{i-1}, 1)
\]

and \( B \) is defined accordingly. Let \( u_i = \phi_i \Delta y_i \) and \( v_i = \phi_i W_i \), \( i = 1, \ldots, n \), be arbitrary linear combinations of \( \Delta y_i \) and \( W_i \) respectively. The set of \( n \) orthogonal choices of \( \phi_i \) and \( \phi_i \) which yield the maximum correlations between \( u_i \) and \( v_i \) is the canonical correlation set. Each statistically zero canonical correlation thus represents a linear combination of \( \Delta y_i \) that is uncorrelated with all linear combinations of \( W_i \), since by definition it is uncorrelated with the combination that maximizes the correlation between \( u_i \) and \( v_i \). Let \( \phi = (\phi_1, \phi_2, \ldots, \phi_n) \) be the matrix containing the combinations which are associated with the zero canonical correlations: by
definition, \( \phi \) is the \( n \times s \) full rank cofeature matrix. The rank \( s \) of this matrix can be determined by calculating the test statistic
\[
C(p, s) = -(T - p - 2) \sum_{i=1}^{s} \ln(1 - \lambda_i^2)
\]
where \( \lambda_1, \lambda_2, \ldots, \lambda_s \) are the \( s \) smallest squared canonical correlations between \( \Delta y_t \) and \( W_t \). Under the null hypothesis that the rank of \( \phi \) is at least \( s \), this statistic has a \( \chi^2 \) distribution with \( s^2 + sn(p - 1) + sr - sn \) degrees of freedom (Vahid and Engle, 1993). The canonical correlations may be computed using the procedure outlined in Hamilton (1994, Ch. 20.1).

An equivalent approach is to incorporate the \( s \) cofeature vectors, as well as the \( r \) cointegrating vectors, into the VECM (3) directly. Vahid and Engle (1993) point out that the cofeature matrix \( \phi \) is only identified up to an invertible transformation, as any linear combination of the columns of \( \phi \) will also be a cofeature vector. The matrix can therefore be rotated to have an \( s \)-dimensional identity submatrix:
\[
\phi = \begin{bmatrix} I_s & \phi_{(n-s)\times s} \end{bmatrix}
\]
\( \phi' \Delta y_t \) can then be considered as \( s \) pseudo-structural form equations for the first \( s \) elements of \( \Delta y_t \). The system can be completed by adding the unconstrained reduced-form equations for the remaining \( n - s \) equations of \( \Delta y_t \) to obtain the system
\[
\begin{bmatrix} I_s & \phi' \end{bmatrix} \Delta y_t = \begin{bmatrix} 0_{s \times (p+r)} \\ B_1 \ldots B_p \beta' \end{bmatrix} + \mu + v_t
\]
(10)

Here \( v_t \) is iid\( (0, \Sigma_v) \), with \( \Sigma_v - \Sigma_P \geq 0 \). Writing the restricted model in this way makes it clear why there are \( s^2 + sn(p - 1) + sr - sn \) degrees of freedom for the common feature test statistic \( C(p, s) \). The unrestricted reduced-form VECM (3) has \( n[p(p-1) + r] \) parameters, whereas the pseudo-structural model (10) has \( sn - s^2 \) parameters in the first \( s \) equations and \( (n-s)[n(p-1) + r] \) parameters in the \( n-s \) equations which complete the system, so imposing \( s^2 + sn(p - 1) + sr - sn \) restrictions. The pseudo-structural system (10) can be estimated by full information maximum likelihood or some other simultaneous equation estimation technique and a likelihood ratio statistic of the restrictions imposed by the \( s \) cofeature vectors can then be constructed: this will be equivalent to \( C(p, s) \).
5 COMMON TRENDS AND CYCLES IN EUROPEAN INDUSTRIAL PRODUCTION

Before embarking on an analysis of common trends and cycles in our European IIP data, we must first determine the orders of the $n = 6$ dimensional VAR($p$) models in each of the two sample periods. This is typically done by employing an information criterion, and we are able to use the levels representation (1) for this, even though all components of $y_t$ are $I(1)$; see Lütkepohl (1991, Ch. 4.3) or Hamilton (1994, Ch. 18). Three criteria were investigated, the familiar AIC and BIC, and the PIC, recently proposed by Phillips (1996) and Phillips and Ploberger (1994). It was found that setting $p = 1$ minimized the BIC for both periods, but subsequent checks on the VAR(1) residuals showed that significant autocorrelation still remained. The PIC selected high values of $p$, around 12, which seemed to overparameterize the models substantially. The AIC selected $p = 2$ for the Bretton Woods period, while $p = 3$ was found to be appropriate for the ERM period. Since the residuals from these VARs could be regarded as white noise, these settings were used for subsequent analysis.

Conditional upon these settings, we then embarked on examining the cointegration properties of the VARs. Table 2 presents the results from Johansen’s (1988, 1991) maximum likelihood procedure (part (a)) and

<table>
<thead>
<tr>
<th>Null hypothesis: $r = \text{rank}(\Pi)$</th>
<th>10% critical value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$r = 0$</td>
<td>104.17</td>
</tr>
<tr>
<td>$r \leq 1$</td>
<td>57.88</td>
</tr>
<tr>
<td>$r \leq 2$</td>
<td>27.03</td>
</tr>
<tr>
<td>$r \leq 3$</td>
<td>11.48</td>
</tr>
<tr>
<td>$r \leq 4$</td>
<td>4.95</td>
</tr>
<tr>
<td>$r \leq 5$</td>
<td>0.00</td>
</tr>
</tbody>
</table>

**Table 2: Cointegration Results**

(a) Johansen’s maximum likelihood test

(b) Stock and Watson’s common trend test

<table>
<thead>
<tr>
<th>$k$ common trends</th>
<th>5% critical value</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>$-0.06$</td>
</tr>
<tr>
<td>1</td>
<td>$-6.34$</td>
</tr>
<tr>
<td>2</td>
<td>$-11.11$</td>
</tr>
<tr>
<td>3</td>
<td>$-15.85$</td>
</tr>
<tr>
<td>4</td>
<td>$-53.39$</td>
</tr>
<tr>
<td>5</td>
<td>$-122.33$</td>
</tr>
</tbody>
</table>

Notes: * The test statistic is Johansen’s (1988) likelihood ratio trace statistic; his max statistic produces identical inferences. Critical values are reported for a constant term in the null; identical conclusions are obtained when a trend is also included.

**The test statistic is Stock and Watson’s (1988) $q(6,k)$ statistic, having a null hypothesis of 6 common trends (i.e. zero cointegrating vectors). The alternative thus corresponds to $6-k$ cointegrating vectors. Identical results are obtained using the $q(6,k)$ statistic.**

© Blackwell Publishers Ltd and The Victoria University of Manchester, 1999.
from Stock and Watson’s (1988) common trends test statistic (part (b)). For both samples, the two approaches find cointegration, but they differ in the number of cointegrating vectors uncovered. For the Bretton Woods sample, the Johansen procedure finds $r = 1$ cointegrating vector, so that there are $k = 5$ common stochastic trends, whereas the Stock and Watson approach finds $k = 4$ common trends, i.e. $r = 2$ cointegrating vectors. For the ERM sample, Johansen finds $r = 2$ cointegrating vectors, and hence $k = 4$ common trends, whereas Stock and Watson find $k = 2$ common trends, i.e. $r = 4$ cointegrating vectors. The identification of fewer common trends under the ERM suggests that there has been some partial convergence under this regime in terms of long-run relationships between EU outputs, which complements the findings from the studies of nominal integration cited earlier. However, since $r < n - 1 = 5$, there is no evidence of complete convergence in the Bernard and Durlauf (1995) sense. (Of course, even if $r = 5$, we would still require the restriction $x' = [I_5, -1]$ to hold for convergence.)

In fact, the evidence against complete convergence can be presented in a much more direct way. A simple test of the null of convergence is to investigate the stochastic properties of the ‘benchmark deviations’ $y_{n,t} - y_{k,t}$. Taking Germany to be the benchmark series, the IIP deviations for the two exchange rate regimes are shown in Figs 2 and 3. In the Bretton Woods regime it is clear that the deviations for the Netherlands and the UK are trended, while for Italy, Belgium and France, unit root tests cannot reject the hypothesis that they do contain unit roots. For Belgium and Italy tests which have stationarity as the null (see, for example, Kwiatkowski et al., 1992) are also able to confirm the presence of unit roots. For the ERM sample, tests of both unit roots and stationarity find no compelling evidence of stationarity for any of the deviations series.

The common feature tests are presented in Table 3 and are reported conditional upon the larger number of cointegrating vectors, as there is now a good deal of evidence accumulating that the Johansen testing procedure may underestimate the number of cointegrating vectors (see, in particular, Toda, 1995), and empirical support for this choice is provided below. For the Bretton Woods sample, the tests find that there is a single ($s = 1$) common feature, and hence $m = 5$ common cycles. For the ERM sample, they do not find any common features (i.e. $s = 0$ or $m = 6$ common cycles), although it should be borne in mind that, with $r = 4$ cointegrating vectors, the maximum value that $s$ can take here is 2.

The greater number of common cycles in the ERM could be interpreted as evidence that this regime has witnessed a ‘disintegration’ of short-run links in EU outputs, with an increased prevalence for asymmetric shocks as ERM members have engaged in country-specific
Fig. 2 Deviations from Germany Benchmark in the Bretton Woods Period

<table>
<thead>
<tr>
<th>Null hypothesis</th>
<th>Bretton Woods</th>
<th>ERM</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>C(p, s)</td>
<td>df</td>
</tr>
<tr>
<td>s &gt; 0</td>
<td>1.03</td>
<td>3</td>
</tr>
<tr>
<td>s &gt; 1</td>
<td>33.17</td>
<td>8</td>
</tr>
<tr>
<td>s &gt; 2</td>
<td>87.80</td>
<td>15</td>
</tr>
<tr>
<td>s &gt; 3</td>
<td>128.87</td>
<td>24</td>
</tr>
<tr>
<td>s &gt; 4</td>
<td>180.80</td>
<td>35</td>
</tr>
</tbody>
</table>

© Blackwell Publishers Ltd and The Victoria University of Manchester, 1999.
responses (i.e. in the terminology introduced above, there is no evidence of short-run convergence). One may point, for example, to the numerous exchange rate crises that have affected various members, and to the diverse policy stances that members have taken, both of which might have contributed towards a greater degree of exchange rate flexibility and monetary autonomy.

For the Bretton Woods sample, the two cointegrating vectors obtained using Johansen’s maximum likelihood procedure define the error correction variables as

© Blackwell Publishers Ltd and The Victoria University of Manchester, 1999.
or, in matrix form,

\[ z_t = z_1 \begin{bmatrix} z_{1t} \\ z_{2t} \end{bmatrix} = \zeta'y_t = \begin{bmatrix} 1 & -1.682 & -0.130 & 0.585 & 0.231 & -0.382 \\ 1 & 1.062 & -1.084 & -0.035 & -0.672 & -0.027 \end{bmatrix}y_t \]

(11)

The Johansen maximum likelihood estimators, while statistically unique, may not be economically interpretable. Johansen and Juselius (1994: see also Johansen, 1995, Ch. 7) have considered identifying long-run relations through imposing restrictions on the cointegrating vectors. It is always possible to impose \( r - 1 \) just-identifying restrictions and one normalization on each vector without changing the likelihood function. Additional restrictions on \( z \) will then be over-identifying and capable of being tested. Johansen and Juselius provide conditions under which such restrictions ‘generically’ identify a set of cointegrating vectors, and if these conditions hold the parameters can be estimated by an iterative procedure, known as a switching algorithm. This algorithm is provided in PeFiml9.0 (Doornik and Hendry, 1997), which also automatically checks the identification condition. The cointegrating vectors (11) were found to satisfy three over-identifying restrictions (the accompanying test statistic had a \( p \) value of 0.76), leading to the estimated vectors, with standard errors in parentheses,

\[ z_{1t} = \text{bel}_t - 2.081 \text{fra}_t + 0.672 \text{hol}_t + 0.374 \text{ita}_t - 0.438 \text{uk}_t, \]

\[ (0.226) \quad (0.119) \quad (0.080) \quad (0.127) \]

\[ z_{2t} = \text{bel}_t + \text{fra}_t - \text{ger}_t - 0.737 \text{ita}_t, \]

\[ (0.021) \]

The first vector, \( z_{1t} \), has just the normalization and a just-identifying restriction imposed, while \( z_{2t} \) has three additional over-identifying restrictions. Since this can be written \( \text{ger}_t = \text{bel}_t + \text{fra}_t - 0.74 \text{ita}_t - z_{2t} \), it is seen that there is a proportional long-run relationship between Germany, Belgium and France, with the Netherlands and the UK not appearing.

After the elimination of insignificant coefficients, the pseudo-structural form (10) was obtained as

\[ \begin{bmatrix} 1 & \phi^* \\ 0, I \end{bmatrix} \Delta y_t = \begin{bmatrix} \theta^* \\ B^*, \beta^* \end{bmatrix} \begin{bmatrix} \Delta y_{t-1} \\ z_{t-1} \end{bmatrix} + \mu + \nu_t \]

with

\( \theta, B, \beta, \phi^* \) parameters.

© Blackwell Publishers Ltd and The Victoria University of Manchester, 1999.
\[
\hat{\phi}^* = \begin{pmatrix}
0.163 & -0.607 & -0.302 & 0.032 & -0.093 \\
0  & 0.105 & -0.283 & 0 & 0.155 \\
0  & 0.432 & -0.431 & 0 & 0 \\
0  & 0.238 & 0 & -0.400 & 0.221 \\
0  & 0 & -0.158 & -0.383 & 0.643 \\
0  & 0.173 & 0 & 0 & -0.193
\end{pmatrix},
\]

\[
\hat{\mu}^* = (0.485 0 0 0)
\]

\[
\hat{B}^*_1 = \begin{pmatrix}
0.1879 & -0.0792 \\
0 & 0.0039 \\
-0.0034 & 0 \\
-0.0029 & 0 \\
0 & 0.0009
\end{pmatrix}
\]

The 29 (common cycle and zero) restrictions imposed on the unrestricted VECM have a \(p\) value of 0.49 (the associated likelihood ratio test statistic is 28.5, distributed as \(\chi^2(29)\)). Note that the ‘factor loading’ matrix \(\hat{\beta}^*\) contains significant elements in each column, so that both cointegrating vectors enter into the VECM, providing support for our choice of \(r = 2\).

For the ERM system, the four cointegrating vectors yield the set of identified error corrections

\[
z_{1t} = \text{bel}_t + 2.206 \text{ita}_t - 3.016 \text{uk}_t,
\]

\[
(0.332) (0.362)
\]

\[
z_{2t} = \text{fra}_t - 0.754 \text{uk}_t
\]

\[
(0.072)
\]

\[
z_{3t} = \text{ger}_t + 2.840 \text{ita}_t - 3.774 \text{uk}_t
\]

\[
(0.433) (0.498)
\]

\[
z_{4t} = \text{hol}_t - 1.955 \text{ita}_t - 2.756 \text{uk}_t
\]

\[
(0.305) (0.351)
\]

This has Phillips's (1991) triangular normalization with a single identifying zero restriction in \(z_2\) (with \(p\) value 0.77). As there are no cofeature vectors, the system has a standard VECM representation, as in (3). After deletion of insignificant coefficients (the set of zero restrictions has a \(p\) value of 0.95, the test statistic being 44.24, distributed as \(\chi^2(61)\)), we obtain the model shown below. The factor loading matrix \(\hat{\beta}^*\) contains significant
elements in all columns, so that all four cointegrating vectors do indeed appear in the VECM.

\[
\hat{B}_1 = \begin{bmatrix}
-0.408 & 0.366 & 0 & 0 & -0.114 & -0.276 \\
0 & -0.408 & 0 & 0 & 0 & 0 \\
-0.129 & 0 & -0.372 & 0 & 0 & 0.172 \\
0 & 0 & 0 & -0.333 & 0 & 0 \\
0 & 0 & 0 & 0 & -0.485 & -0.461 \\
0 & 0 & 0 & -0.109 & 0 & -0.170
\end{bmatrix}
\]

\[
\hat{B}_2 = \begin{bmatrix}
-0.426 & 0.212 & 0 & 0 & -0.138 & 0 \\
0 & -0.282 & 0 & 0.076 & 0 & 0 \\
-0.071 & 0 & -0.119 & 0 & 0 & 0 \\
0 & -0.399 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & -0.256 & -0.465 \\
-0.079 & 0 & 0 & 0 & 0 & 0
\end{bmatrix}
\]

\[
\hat{\beta} = \begin{bmatrix}
-0.293 & 0 & 0.188 & 0 \\
0 & -0.125 & 0 & 0 \\
0.133 & 0 & -0.165 & 0.072 \\
0 & 0 & 0.276 & -0.439 \\
-0.093 & 0.072 & 0 & 0 \\
0 & -0.045 & -0.048 & 0.072
\end{bmatrix}
\]

\[
\hat{\mu}' = (0.194 \ 0.143 \ -0.129 \ 0.327 \ 0 \ 0)
\]

Our finding of cointegration in both periods implies that the levels VAR (1) is an appropriate representation, albeit with restrictions imposed on the \( A_i \) matrices by the various common trends and, in the Bretton Woods period, common cycles. In analysing the dynamic structure of the models, it is convenient to rewrite (1) as

\[
y_t = B(L)\Delta y_{t-1} + Ay_{t-1} + \mu + \epsilon_t \quad (12)
\]

where

\[
A = A(1) - I = \sum_{i=1}^{p} A_i
\]

For the Bretton Woods model, we have

© Blackwell Publishers Ltd and The Victoria University of Manchester, 1999.
\[ A_1 = I_6 + \left[ -\phi \beta' \right] (\beta \beta' \phi') \quad A_2 = -\left[ -\phi \beta' \right] B_1 \quad A_p = 0, \ p > 2 \]

That is,

\[ y_t = B_1 A y_{t-1} + A y_{t-1} + \mu + e_t \]

where

\[
B_1 = -A_2 = \begin{bmatrix}
0 & 0.33 & -0.21 & -0.11 & 0 & 0 \\
0 & 0.11 & -0.28 & 0 & 0 & 0.16 \\
0 & 0.43 & -0.42 & 0 & 0 & 0 \\
0 & 0.24 & 0 & -0.40 & 0 & 0.22 \\
0 & 0 & -0.16 & -0.38 & 0 & 0.64 \\
0 & 0.17 & 0 & 0 & 0 & -0.19
\end{bmatrix}
\]

and

\[
A = A_1 + A_2 = \begin{bmatrix}
0.98 & 0.08 & -0.02 & -0.02 & -0.02 & 0.01 \\
0.11 & 0.53 & 0.08 & 0.13 & 0.13 & -0.08 \\
0 & 0 & 1.00 & 0 & 0 & 0 \\
0 & 0.01 & 0 & 1.00 & 0 & 0 \\
0 & 0.01 & 0 & 0 & 1.00 & 0 \\
0 & 0 & 0 & 0 & 0 & 1.00
\end{bmatrix}
\]

For there to be no Granger causality running from \( y_i \) to \( y_j \) requires that \( B_{i,j}^\perp \) and \( A_{i,j} \) are both zero. From examination of the two matrices, we observe that Belgium and France are Granger-caused by all countries. Germany is only Granger-caused by France, while the Netherlands is only Granger-caused by France and the UK, so that the Netherlands and Germany are independent. Italy is not Granger-caused by Belgium (and possibly France) and the UK is only Granger-caused by France. We should emphasize that we are not providing formal tests of Granger non-causality here. For conventional Wald tests of non-causality in standard VECMs to be asymptotically valid as \( \chi^2 \) criteria, we require detailed information about the rank conditions of certain submatrices of the cointegrating matrix \( \alpha \) and the ‘loading’ matrix \( \beta \) (see Toda and Phillips, 1993). While this information can be obtained, at least in principle, we prefer to approach the issue of causality testing in a more direct fashion using an approach to be outlined below.

We also note that, apart from the first two rows of \( A \), all off-diagonal elements are either zero or tiny in magnitude. This suggests that innovations to Germany, the Netherlands, Italy and the UK are predominantly transitory, while those to Belgium and France are more persistent.

For the ERM model, we have

\[ A_1 = I_6 + \beta \phi' + B_1 \quad A_2 = B_2 - B_1 \quad A_3 = -B_2 \quad A_p = 0, \ p > 3 \]

© Blackwell Publishers Ltd and The Victoria University of Manchester, 1999.
Thus
\[ y_t = B_1 \Delta y_{t-1} + B_2 \Delta y_{t-2} + A y_{t-1} + \mu + \epsilon_t \]

where
\[
A = A_1 + A_2 + A_3 = \begin{bmatrix}
0.71 & 0 & 0.19 & 0 & -0.11 & 0.17 \\
0 & 0.87 & 0 & 0 & 0 & 0.09 \\
0.13 & 0 & 0.83 & 0.07 & -0.03 & 0.02 \\
0 & 0 & 0.28 & 0.56 & -0.09 & 0.07 \\
-0.09 & 0.07 & 0 & 0 & 0.79 & 0.23 \\
0 & -0.05 & -0.05 & 0.07 & 0 & 1.02
\end{bmatrix}
\]

There is much greater evidence of feedback and persistence in the ERM system. Far fewer elements of \( A \) are zero or very small, even though there are only two unit roots in the system, the other four all being real, with magnitudes 0.91, 0.80, 0.64 and 0.44.

6 Levels VAR Modelling

The approach taken above to modelling the two systems has been to pretest the data to establish whether reduced rank models are appropriate, and then to engage in a simplification search to obtain a parsimonious, yet congruent, empirical representation with which to investigate the dynamic interactions between the various series. While care has been taken to ensure that the restrictions imposed in moving from the initial levels VAR to the reduced rank VECMs actually reported are data consistent (recall the reported \( \chi^2 \) tests), some doubts may be harboured about the legitimacy of all the restrictions that have been imposed to arrive at these models. Moreover, as we have discussed above, formal testing of Granger non-causality hypotheses is complicated in such models, particularly after a model reduction sequence has been performed. We therefore investigate two alternative approaches that seek to avoid any form of pre-testing apart from setting the order \( p \) of the levels VAR, yet still provide consistent estimation of the system even in the presence of an unknown number of unit roots and cointegrating vectors and allow Granger non-causality hypotheses to be tested directly.

One approach, termed FM–VAR regression by Phillips (1995), builds on earlier work by Phillips and Hansen (1990) on fully modified (FM) ordinary least squares (OLS) regression, which provides optimal estimates of (single) cointegrating regressions. To define the FM–VAR estimator, it is convenient to rewrite (12) as
\[ y_t = F x_t + \epsilon_t \]  
\[ (13) \]

where \( x_t = (\Delta y_{t-1}, \ldots, \Delta y_{t-p+1}, 1, y_{t-1}) = (w_t, y_{t-1}) \) and \( F = (B_1, \ldots, B_{p-1}, \mu, A) \).

The FM–VAR estimator of \( F \) is then defined as
\[ \hat{F} = [Y'W : Y'Y_1 - \hat{\Omega}_\eta^{-1}(\hat{\Delta}Y_{11} - T\hat{\Delta}Y_{11})][X'X]^{-1} \]

Here, \( Y, Y_1, W, \Delta Y_{11} \) and \( X \) are data matrices constructed from the variables in (13), e.g. \( Y = (y_1, \ldots, y_T) \). \( \hat{\Omega}_\eta \) and \( \hat{\Omega}_\eta \) are kernel estimates of the long-run covariance matrices of \( \Delta Y_{11} = y - \hat{F}_x, \Delta Y_{11} \) and \( \Delta Y_{11} \), respectively. (The long-run covariance matrices of two covariance stationary series, \( u_t \) and \( v_t \), is defined as \( \Omega_m = \sum_{\omega=-\infty}^{\infty} E(u_{t+k}v_t') \). Kernel estimates use weighted moving sums of sample covariances to non-parametrically estimate a long-run covariance matrix.)

\[ \hat{F} = Y'X(X'X)^{-1} = [Y'W : Y'Y_1][X'X]^{-1} \]

is the OLS estimator of \( F \). \( \hat{\Delta} \) is a kernel estimate of the one-sided long-run covariance of \( \Delta Y_{11} = \Delta u = \sum_{\omega=0}^{\infty} E(u_{t+k}v_t') \). The OLS estimator is thus ‘modified’ to take into account the endogeneity brought about by the inclusion of the \( \Delta(1) \) regressor \( Y_{11} \); see Phillips (1995) for technical details and extended discussion. A routine for computing the FM–VAR estimator within the GAUSS language is provided by Ouliaris and Phillips (1995).

Hypothesis tests on \( F \) that use the FM–VAR regression estimator \( \hat{F} \) may be constructed from the asymptotic approximation

\[ T^{1/2}(\hat{F} - F) \sim N[0, \Sigma_p \otimes T(X'X)^{-1}] \]

Suppose that we wish to test the set of \( q \) restrictions

\[ H_0: R \text{ vec}(F) = r \]

where \( R \) is of full row rank \( q \). The Wald statistic of this hypothesis is

\[ W_p = T(R \text{ vec}(\hat{F}^*) - r)[R(\hat{\Sigma}_p \otimes T(X'X)^{-1})R']^{-1}(R \text{ vec}(\hat{F}^*) - r) \]

Phillips (1995) shows that the \( \tilde{\chi}^2(q) \) distribution is an upper bounding variate for this statistic, so that the usual \( \chi^2 \) critical values can be used to construct tests that have conservative size.

In this framework, causality restrictions can be set up in the following way. The hypothesis that \( y_j \) has no Granger-causal effect on \( y_i \), which imposes the restrictions \( B_{11} = \ldots = B_{p-1,1} = A_{11} = 0 \), can be written as

\[ H_0: (R_i \otimes R_j) \text{ vec}(F) = 0 \]

where \( R_i = e_i \) and \( R_j = I_p \otimes e_i', k = i, j \), being a selection vector. The associated Wald statistic then has a \( \chi^2(p) \) limit distribution.

The second approach has been developed by Toda and Yamamoto (1995) and Saikkonen and Lütkepohl (1996), who show that OLS estimates of the parameters in (13) are consistent if, rather than the \( x_t \) being used as regressors, we use the augmented set \( x_t' = (x_t' : \Delta Y_{1p}) \), i.e. a VAR\((p+1)\) is fitted rather than the correctly specified VAR\((p)\), with the additional matrix of coefficients, \( B_p \), being regarded as nuisance parameters whose presence nevertheless induces consistency of \( \hat{F} \), the
OLS estimator of $F$ using $x_t$, and allows conventional Wald tests of $H_0$ to have standard asymptotic $\chi^2$ distributions.

Both approaches were used to estimate $F$ in (13) and to construct Granger causality tests. The test statistics obtained from the FM-VAR approach are not reported, however, as they bear little resemblance to the causal patterns obtained from the VECM models. Those from the augmented VAR approach are reported in Table 4. There is reasonable consistency between the patterns of causality found in the VECM models and these test statistics, although the latter tend to uncover fewer statistically significant causal relationships, possibly because the additional lags required for consistency also make the tests inefficient. In both eras, Belgium and France tend to be caused by the others. There is support for German leadership over Belgium, France and the Netherlands in the ERM era, a necessary condition for this being German Granger causality but no feedback.

7 Estimating the Permanent Components of Industrial Production

The presence of cointegration in both systems implies, as we have seen, that the vector of European IIP may be thought of as being ‘driven’ by a smaller number of common trends or permanent components. As these are unobservable and must therefore be estimated, an important issue is whether such components are identified. The Stock and Watson (1988) common trends formulation (4) achieves identification by imposing the

<table>
<thead>
<tr>
<th>i → j</th>
<th>Belgium</th>
<th>France</th>
<th>Germany</th>
<th>Netherlands</th>
<th>Italy</th>
<th>UK</th>
</tr>
</thead>
<tbody>
<tr>
<td>(a)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Belgium</td>
<td>—</td>
<td>6.44</td>
<td>3.17</td>
<td>2.28</td>
<td>3.19</td>
<td>2.37</td>
</tr>
<tr>
<td>France</td>
<td>2.71</td>
<td>—</td>
<td>15.36</td>
<td>8.59</td>
<td>9.68</td>
<td>0.71</td>
</tr>
<tr>
<td>Germany</td>
<td>5.37</td>
<td>10.64</td>
<td>—</td>
<td>1.81</td>
<td>0.53</td>
<td>2.30</td>
</tr>
<tr>
<td>Netherlands</td>
<td>1.23</td>
<td>11.30</td>
<td>0.74</td>
<td>—</td>
<td>0.54</td>
<td>4.00</td>
</tr>
<tr>
<td>Italy</td>
<td>3.34</td>
<td>4.23</td>
<td>4.51</td>
<td>4.03</td>
<td>—</td>
<td>9.43</td>
</tr>
<tr>
<td>UK</td>
<td>0.74</td>
<td>8.98</td>
<td>1.09</td>
<td>5.64</td>
<td>0.85</td>
<td>—</td>
</tr>
</tbody>
</table>

| (b)  |         |        |         |             |      |    |
| Belgium | —      | 9.18   | 9.92    | 2.68        | 8.72 | 9.52|
| France  | 1.01   | —      | 9.86    | 6.57        | 2.37 | 2.33|
| Germany | 2.35   | 1.07   | —       | 2.70        | 0.53 | 4.30|
| Netherlands | 1.25 | 3.53   | 7.46    | —           | 0.35 | 1.51|
| Italy   | 2.83   | 4.19   | 0.54    | 3.97        | —    | 14.86|
| UK      | 5.96   | 9.87   | 3.63    | 7.29        | 4.15 | —   |

Notes: a Statistics have a limiting $\chi^2(2)$ distribution; critical values are 10 per cent, 4.60; 5 per cent, 5.99. 
b Statistics have a limiting $\chi^2(3)$ distribution; critical values are 10 per cent, 6.25; 5 per cent, 7.81.
conditions that the permanent component $\tau_t$ is a random walk and that both this and the *transitory* component $c_t$ are driven by the same innovation $e_t$, so that they are perfectly correlated. Quah (1992) and Gonzalo and Granger (1995) investigate an alternative, less restrictive, definition of a permanent–transitory decomposition and we follow the latter in estimating the permanent components driving European industrial production. Consider therefore the VECM (3), where $y_t$ is now assumed to be written in mean deviation form for simplicity:

$$
\Delta y_t = B(L)\Delta y_{t-1} + \beta x'y_{t-1} + e_t
$$

(14)

As in (4), $y_t$ can be explained in terms of a smaller number, $k = n - r$, of I(1) variables, known as common factors and denoted $f_t$, plus some I(0) components $\tilde{y}_t$:

$$
y_t = G_1 f_t + \tilde{y}_t
$$

(15)

The first identification condition is that the common factors $f_t$ are linear combinations of $y_t$,

$$
f_t = Jy_t
$$

so that

$$
\tilde{y}_t = (I - G_1 J)y_t = G_2 x'y_t
$$

for some appropriately defined matrix $G_2$. The second identification condition is that $\tilde{y}_t$ does not Granger-cause $G_1 f_t$ in the long run: see Gonzalo and Granger (1995, Definition 1) for further details and Granger and Lin (1995) for a formal definition of long-run Granger causality. This implies that, in the VAR representation of $(\Delta f_t, \tilde{y}_t)$,

$$
\begin{bmatrix}
H_{11}(L) & H_{12}(L) \\
H_{21}(L) & H_{22}(L)
\end{bmatrix}
\begin{bmatrix}
\Delta f_t \\
\tilde{y}_t
\end{bmatrix}
= 
\begin{bmatrix}
u_{1t} \\
u_{2t}
\end{bmatrix}
$$

the total multiplier of $\Delta f_t$ with respect to $\tilde{y}_t$ is zero, i.e. $H_{12}(1) = 0$. Now pre-multiplying (16) by $J$, and noting that $J\Delta y_t = \Delta f_t$ and $x'y_t = G_2^{-1}\tilde{y}_t$, yields

$$
\begin{bmatrix}
I - B(L) & J \beta G_2^{-1}L \\
H_{21}(L) & H_{22}(L)
\end{bmatrix}
\begin{bmatrix}
\Delta f_t \\
\tilde{y}_t
\end{bmatrix}
= 
\begin{bmatrix}
J e_t \\
u_{2t}
\end{bmatrix}
$$

so that the condition $H_{12}(1) = 0$ is satisfied if $J = \beta_\perp$, where $\beta_\perp$ is the orthogonal complement of $\beta$ such that $\beta_\perp \beta = 0$ (see Banerjee et al., 1993, p. 147, for a formal definition of an orthogonal complement). Note that $J\Delta y_t$ has the ‘common feature’ (Engle and Kozicki, 1993) of not containing the error correction $z_{t-1}$. Substituting these two identifying conditions into (15) yields

$$
I = G_1 \beta_\perp + G_2 x'
$$

© Blackwell Publishers Ltd and The Victoria University of Manchester, 1999.
or

\[
\begin{bmatrix}
G_1 \\
G_2
\end{bmatrix} = [\beta'_\perp, \gamma]^\top
\]

That is,

\[
G_1 = \alpha_\perp (\beta'_\perp \gamma_{\perp})^{-1} \quad G_2 = \beta (\alpha \beta)^{-1}
\]

The Gonzalo–Granger decomposition (17) and the Stock–Watson representation (4) are closely linked. The Stock–Watson common trends component is

\[
\gamma \tau_t = \gamma \delta \Delta^{-1} e_t = C(1) \Delta^{-1} e_t
\]

Now, it can be shown (Banerjee et al., 1993, pp. 148–150) that

\[
C(1) = \alpha_\perp (\beta'_\perp \Psi \gamma_{\perp})^{-1} \beta'_\perp
\]

where

\[
\Psi = I - B(1) + \Pi
\]

After some algebra, this can be written as

\[
C(1) = G_1[I - \beta'_\perp B(1)G_1]^{-1} \beta'_\perp
\]

Returning to the VECM (14), pre-multiplying by \(\beta'_\perp\) and substituting \(G_1 f_t + G_2 z_t\) for \(y_t\) yields, on noting that \(\beta'_\perp G_1 = I\) and \(\beta'_\perp G_2 = 0\) and on rearranging,

\[
[I - \beta'_\perp B(L) L G_1] \Delta f_t = \beta'_\perp B(L) G_2 z_{t-1} + \beta'_\perp e_t
\]

The random walk component of \(f_t\) (in the multivariate Beveridge–Nelson (1981) sense) is

\[
[I - \beta'_\perp B(1)G_1]^{-1} \beta'_\perp \Delta^{-1} e_t
\]

which, when compared with the definition of \(C(1)\) above, shows that the Stock–Watson common trend is the random walk component of the Gonzalo–Granger common factor \(f_t\): since they differ only by \(I(0)\) components, they will be cointegrated.

The Gonzalo–Granger decomposition can be computed as long as we have available estimates of the orthogonal complements \(\gamma_{\perp}\) and \(\beta'_{\perp}\). An estimate of the former is obtained as a byproduct of maximum likelihood estimation (see Johansen, 1995) and, as Gonzalo and Granger (1995) point out, there is a duality between \(\beta'_{\perp}\) and \(\gamma\), so that an estimate of \(\beta'_{\perp}\) can be obtained by solving an eigenvalue problem that is related to that which provides the estimates of \(\alpha\) and \(\gamma_{\perp}\). Gonzalo and Granger provide details of the estimation procedure, which is followed here.

For the Bretton Woods system, which has \(r = 2\) cointegrating vectors, there are four common factors and \(\beta'_\perp\) and \(G_1\) are \(6 \times 4\) matrices. For the
Fig. 4 Annual Growth Rates, Bretton Woods

© Blackwell Publishers Ltd and The Victoria University of Manchester, 1999.
ERM system, where there are \( r = 4 \) cointegrating vectors, \( f_t \) is of dimension 2 and \( \beta_L \) and \( G_1 \) are \( 6 \times 2 \) matrices. Rather than report these estimated matrices, which convey little in isolation, we present information about the permanent components, defined as \( y_t^p = G_1 \beta_L y_t \), graphically.

We look first at the Bretton Woods decompositions. For Germany, the Netherlands, Italy and the UK, the weight given to the home country in the permanent component dominates the rest, so that their observed and permanent growth rates are highly correlated (correlations range from 0.97 to 0.997), and deviations between the two are very small and short-lived. This is not the case for Belgium and France, however, where the weights are much more equally spread. As a consequence, the divergences between observed and permanent growth rates, shown in Fig. 4, are larger and more persistent, which mirrors the findings presented earlier from the VECM analysis. A rather different pattern emerges from the ERM decompositions, but again one that is consistent with the previous VECM analysis. Here the correlations between the observed and permanent growth rates are substantially weaker—indeed, there is almost no correlation (−0.04) between them for Italy—although the UK does appear to be an exception to this, the correlation here remaining high at 0.94. Rather than plot growth rates, we show in Fig. 5 the levels of the

![Fig. 5 Permanent Components, ERM](image-url)
permanent components, which should be compared with the corresponding plot of the observed levels in Fig. 1. The presence of the four cointegrating vectors, and hence two common factors, is clearly seen as the components follow a distinct non-stationary but cyclical pattern. Interestingly, in the final year of the sample, 1994, observed growth was higher than permanent growth in all six countries, whereas in the previous year observed growth was less than permanent growth in all. We conjecture that this was a consequence of the relatively tight monetary policy in ERM economies before the 1993 widening of the permitted bands of exchange rate fluctuation to ±15 per cent, after which monetary policy relaxed considerably.

8 Summary and Conclusions

We may draw the following conclusions from our analysis. Our results provide only mixed evidence that the ERM period, when compared with Bretton Woods, has been associated with an increased convergence of EU industrial production. Long-run convergence has increased with the presence of fewer common trends, whereas the absence of common cycles suggests that the ERM period has been associated with a short-run divergence of output among EU members, and this is confirmed by the computation of the permanent components. Increased convergence in the long run reflects the sentiment of earlier studies on inflation and interest rates, where Caporale and Pittis (1993), Hafer and Kutan (1994) and Thom (1995) analyse common shared trends in the ERM and find evidence of partial convergence. It is our view that the long-run convergence of output has been facilitated by relaxed capital controls and increased goods and factor mobility in the EU. On the other hand, short-run convergence has been restricted by the prevalence of asymmetric exchange rate shocks (thus lending support to De Grauwe, 1994) and by exchange rate bands which have permitted some limited scope for monetary autonomy. Thus, in contrast to earlier studies on inflation and interest rates, we cannot argue that the ERM has provided even partial short-run convergence.

Our results also differ from those of Artis and Zhang (1997), who detrend IIP data for 15 OECD economies and examine contemporaneous or maximum cross-country correlations. They find that the ERM countries have shifted their business cycle affiliation from the USA to Germany during the ERM period. Our methodology, based on common cycles and permanent components, offers a more complete empirical view of macroeconomic interdependence, although we do concede that, just as Hafer and Kutan (1994) and Koedijk and Kool (1992), we do not include the USA in our analysis. This could, of course, be done but would certainly increase the complexity of the modelling.
Our investigation suggests that, while the instances of Granger causality increase during the ERM, there is only limited evidence of German economic leadership. As with inflation and interest rates, output relationships are interactive in nature. While the UK is characterized by the greatest degree of independence, in terms of the extent of non-causality by other EU members, we find little evidence of a distinct core grouping based on the degree of commitment to ERM membership.

The study highlights the need for three key considerations when making judgements about international economic convergence. First, one should explicitly acknowledge the nature of the exchange rate regime that is in operation. The studies by Serletis and Krichel (1992), Engle and Kozicki (1993) and Calcagnini (1995) examine common trends and cycles for G7 and OECD output for periods of nearly 40 years. Their findings are biased by the exclusion of any formal demarcation of the data on the basis of exchange rate regimes. Indeed, Calcagnini finds evidence in favour of only a single common cycle in the case of G7 per capita output, thereby implying that the G7 economies have experienced short-run convergence. Second, one should distinguish between long-run and short-run convergence. Our evidence suggests that a more complete picture of economic integration is formed when both types of convergence are considered. A weakness of many studies on inflation and interest rates is that it has usually been the case that only long-run convergence has been considered. Third, if production in the EU was to become more specialized as members proceed towards EMU, it is possible that an increased vulnerability might reduce the probability of a single common cycle despite the absence of monetary autonomy.

We can identify a number of avenues for future research. First, a study of common trends and cycles at a more disaggregated level would be useful in order to investigate convergence for particular industries across the EU. This may assist in the identification of those areas of industry in the EU which have experienced the greatest degree of convergence. Second, a study which re-examines inflation and interest rate convergence on the basis of common cycles would test the robustness of the earlier conclusions on convergence.

REFERENCES
Common Trends and Cycles in European Industrial Production 585


© Blackwell Publishers Ltd and The Victoria University of Manchester, 1999.


