J. Japan Statist. Soc.
Vol. 31 No. 2 2001 129–152

EXAMINING SIMILARITIES AMONG PHASES OF BUSINESS CYCLES —IS THE 1990S US EXPANSION SIMILAR TO THE 1960S'?—

Masaki Katsuura* and Allan P. Layton**

Statistical similarities among the latest long expansion in the U.S. and some other past expansions, in particular that of the 1960s, are examined. Corresponding to the definition of statistical similarity, a test based on the covariance matrices of business cycle component variables for the different expansions is proposed. Among available tests, the test based on partial common principal component analysis is argued to be most appropriate. The test is applied to the components of both GDP and the coincident composite index. As a result, the 1990s expansion is concluded to be statistically similar to that of the 1960s. Using the same method we also examine the statistical similarities of whole cycles (defined on a peak-to-peak basis).

Key words and phrases: Business Cycle; Statistical Similarity; Covariance Matrix Structure; Composite Index; Partial Common Principal Component Analysis.

1. Introduction

An interesting current topic concerning the US economy at the moment is how long the present long expansion, starting from 1991, will continue. Since the last trough in March 1991 the expansion has so far continued for 117 months through until the end of 2000 and is now the longest of the post-war era. In fact, the present expansion in the US economy is so long and strong that it has even produced a view that the US economy may have fundamentally changed and entered a new era of prolonged economic growth supported by high productivity. Many business cycle economists, however, think that the business cycle is still very much alive and believe the US economy will eventually and inevitably go into its next contraction.

The longest previous expansion after World War II occurred in the 1960s, beginning in March 1961 and ending in December 1969 for a total duration of 106 months. Because of its length and strength, the present expansion is sometimes compared with the 1960s expansion. In this comparison the duration of the two expansions is emphasised because they represent the two longest expansions since World War II and, in terms of this duration characteristic, these two expansions are now clearly quite similar.

However, there are several other important characteristics of business cycle

Received August 8, 2000. Revised June 29, 2001. Accepted July 10, 2001.

*Faculty of Economics, Meijo University, 1-501 Shiogamaguchi, Tenpaku-ku, Nagoya, Aichi, 468-8502, Japan. Email: katsuura@meijo-u.ac.jp

**School of Economics and Finance, Queensland University of Technology, GPO Box 2434, Brisbane, Queensland 4001, Australia. Email: a.layton@qut.edu.au

Table 1. Average Growth Rate (%) of the US GDP Components by Expansions.

expansion	start		end	quarters	GDP	C	I	G	X	M
E1	50:Q1	-	53:Q3	15	6.81	3.87	9.82	16.18	3.83	11.67
E2	54:Q3	-	57:Q3	13	3.91	4.23	7.49	0.04	9.19	5.59
E3	58:Q3	-	60:Q2	8	5.78	4.95	14.18	2.46	10.50	7.35
E4	61:Q1	-	69:Q4	35	4.80	4.84	7.16	3.77	6.14	9.15
E5	71:Q1	-	73:Q4	12	5.21	4.73	13.76	−0.69	10.79	5.91
E6	75:Q2	-	80:Q1	20	4.24	3.86	9.85	1.57	7.29	7.95
E7	80:Q4	-	81:Q3	4	4.18	2.01	23.44	0.97	−0.42	8.85
E8	83:Q1	-	90:Q3	31	3.75	3.70	5.31	3.22	8.70	9.13
E9	91:Q2	-	(98:Q4)	(31)	3.13	3.27	8.53	0.49	7.61	10.14
				average	3.98	4.03	5.27	3.21	7.36	7.84

Note: Figures are transformed into annual growth rates. (98:Q4) represents the end of our available data and is not part of the official chronology. C, I, G, X and M denote consumption, investment, government expenditure, exports and imports, respectively.

phases other than duration, viz. amplitude, asymmetry, strength (or depth), the classification of types of phases[1], and so on. For example, the average annual growth rates of GDP components are sometimes used to characterise economic activity in different phases or states of the business cycle. Table 1 shows the average growth rates of the GDP components for all US post-war expansions. We could also discuss similarities among expansions by looking at the similarity or dissimilarity of such growth rates across expansions. Statistical tests, such as tests for equality of means, could then be used to judge whether the expansions were statistically similar or not.

In a more sophisticated way, Pagan (1997) characterises business cycles in many countries in terms of duration and asymmetry using simple statistical models and simulation, and Hess and Iwata (1997) also analyse the business cycle features of duration and amplitude by a very similar method. In these recent studies, univariate time series models are applied to GDP and simulation is carried out to characterise business cycle features. King and Plosser (1994) and Simkins (1994) use a similar approach using real business cycle models. Kim *et al.* (1994) and Fisher *et al.* (1996) characterise business cycles using cross correlations for detrended variables. These authors use real GDP as the representative indicator of business cycle fluctuations. On the other hand, Layton (1998) and others have applied nonlinear time series models, viz. regime switching models developed by Hamilton (1989), to the coincident and leading composite indexes to investigate the characteristics of US business cycles. See also Layton and Smith (2000) and Harding and Pagan (1999a, b).

As noted above, comparisons of business cycle phases often focus on the relative growth rates of GDP and its components or the components of the coincident composite index. However, in comparing different business cycle phases, it is clearly insufficient to frame the analysis solely on the basis of comparing such growth rates. In this paper, we will investigate the statistical similarity between

1 Layton and Smith (2000).

the present expansion and other expansions in the U.S. in a somewhat different way. We examine the extent of overall statistical similarity among expansions using statistical tests based on the variance-covariance matrices (hereafter covariance matrices) of a vector of business cycle variables across different expansions. Expansions are regarded as statistically similar if the covariance matrices have a particular type of similarity.

The suggested approach provides a method to comprehensively assess the degree of phase similarity since the covariance matrix for each phase provides a useful statistical summary of the interrelationships among the business cycle component variables. A non-rigorous very subjective assessment of such covariation could be carried out by "eye-balling" graphs of the various variables under analysis for whatever phases are of interest. However, it is by no means straightforward, using such a graphical approach, to form views about the degree of similarity in variation among the components in the two episodes. The suggested approach allows a degree of statistical rigour to be brought to the issue.

Nonetheless, as noted above, we also recognise that comparing the average growth rates of business cycle component variables across phases is a common and very important aspect of the issue of similarity. Since our approach uses covariance matrices in the comparisons of phases, we focus on the covariation of variables around their means across phases rather than investigating the similarity or dissimilarity in the mean growth rates themselves. In this sense, our approach must be regarded as complementary to the more usual approach.

We should point out at the outset that, as noted by Blanchard and Watson (1982), broadly speaking there are two approaches to analysing features of the business cycle, viz the propagation approach and that of Burns and Mitchell (1946). For the purposes of phase similarity comparisons the former approach would involve specifying and estimating an underlying economic model —most commonly some form of a VAR model— as driving the observations on the vector of component variables under study. The estimated coefficients of such a model across different phases could then be tested for statistical equivalence.

The latter approach would simply emphasise the extent of observed empirical covariation among identified component variables. No explicit behavioural macroeconomic model would be specified and in statistically comparing different phases, this approach would then investigate the nature of such statistical covariation across phases. The current paper suggests a statistical approach to carrying out such comparisons.

It should be noted that the Economic Planning Agency (2000) recently emphasized that the present recovery in Japan, commencing in April 1999, is structurally dissimilar to past expansions. The dissimilarity is analysed by the EPA using many charts of various economic indicators, such as GDP components, profits, labour variables, etc. The approach suggested in this paper will be useful in further examining such a statement and in providing a statistical judgement in terms of the 'overall' relationship among economic indicators used for analysing business cycles.

The paper is organised as follows. In Section 2, the statistical similarity concept used in this paper is defined. Section 3 presents the statistical test used in the paper for examining statistical phase similarities. The test we adopt is based on the 'partial' common principal component analysis proposed by Flury (1987). The proposed method can be applied to both expansions and contractions or to whole cycles, but in this paper our interest is on post World War II US expansions and whole cycles only. We apply the test to the coincident composite index components in Section 4 and to the components of GDP in Section 5. Some extensions of the basic approach are investigated in Section 6 and Section 7 contains some concluding remarks.

2. Definition of statistical similarity among phases of business cycles

Before examining business cycle statistical similarities, we first have to define what we mean by similarity and, indeed, we must also even define what we mean by the business cycle. We will discuss statistical similarity from the point of view of principal component analysis (PCA) and common principal component analysis (CPCA) which generalises PCA to the case of more than one group. Applications of PCA and CPCA to the components of the Japanese diffusion index are seen in Kariya (1986) and Katsuura (1988), respectively, and the optimality of applying PCA to time series data is discussed in Kariya (1993). Also, Stock-Watson's model for business cycle analysis in Stock and Watson (1991, 1993) is an application of a type of dynamic principal component or factor analysis.

2.1. Definition of business cycles

An appropriate definition of the business cycle has been discussed by a number of authors. The most commonly used definition is to regard business cycles as significant cyclical fluctuations in "aggregate economic activity" as defined by Burns and Mitchell (1946)[2]. By this definition, one approach is to define business cycle phases by analysing a group of selected related economic variables such as appropriate measures of income, output, sales and employment. Another approach is simply to use cyclical movements in GDP as a proxy for the business cycle.

Using GDP, one widely used convenient rule of thumb to determine peaks of the business cycle is the period prior to that for which GDP falls for at least two consecutive quarters. An analogous rule is applied to identify trough dates. Some countries, for example Australia, use this simple rule to establish a business cycle chronology. Other countries, in particular the U.S., use a much more elaborate statistical approach involving a range of economic indicators to determine the business cycle chronology.

When cyclical fluctuations in GDP are analysed, the demand components of it, viz. consumption, private investment, government expenditure and net exports are often given much attention. If we adopt GDP as a proxy for the business cycle, it is natural then to regard business cycles as the combined fluctuations in these components and therefore in looking at similarities across phases it would be sensible to investigate fluctuations in these GDP components. Analogously,

2 See Diebold and Rudebush (1996), King and Plosser (1994).

following the approach of Burns and Mitchell a group of appropriately identified macroeconomic variables deemed to be representative of the business cycle would be the object of the analysis.

Therefore, irrespective of whether one prefers to use the composite index or GDP, business cycles may be regarded as a weighted average of fluctuations in the relevant components. If we assume linearity, we could therefore express the business cycle as

$$f_t = \beta_1 x_{1t} + \beta_2 x_{2t} + \cdots + \beta_p x_{pt} \tag{2.1}$$

or

$$f_t = \sum_{i=0}^{r} \beta_{i1} x_{1,t-i} + \sum_{i=0}^{r} \beta_{i2} x_{2,t-i} + \cdots + \sum_{i=0}^{r} \beta_{ip} x_{p,t-i} \tag{2.2}$$

where f denotes the business cycle, x is a variable in the coincident index or GDP, β is the weight attached to the variable, p is the number of variables and r is number of lags.

2.2. Definition of similarities

As noted in the introduction one approach to investigating statistical similarities or dissimilarities between business cycle phases is to measure the degree of covariation among the various component variables taken to represent the business cycle, i.e. if correlations among variables between two phases are statistically similar, those two phases may be regarded as statistically similar. For example, consider the situation in which consumption, as a GDP component, shows similar average growth between two phases. This suggests a certain similarity between the two phases but this is only part of the picture. Variation around this average could be quite different across two phases as could the extent of covariation between consumption and other components under study. We cannot conclude whether these two phases are similar or not in respect of consumption unless we also consider the variation in consumption and its relation to other variables in the two episodes as well. A useful overall measure of the interrelationships among variables is their covariance matrix[3]. Therefore, a potentially useful measure of statistical similarity is the extent to which the covariance matrices of different phases are statistically the same. As noted earlier, however, for the purposes of this paper, our analysis of similarities abstracts from incorporating an explicit statistical comparison of component variable mean growth rates.

Whilst the above discussion is quite intuitive, we now formally define statistical similarity based on equation (2.1) or (2.2)[4]. Consider the question of statistical similarity between two phases, viz. phase 1 and 2. Analogous to equation (2.1) we would have:

$$\begin{aligned} f_t &= \beta_1^{(1)} x_{1t} + \beta_2^{(1)} x_{2t} + \cdots + \beta_p^{(1)} x_{pt} \qquad (t = \tau_1, \tau_1 + 1, \ldots, \tau_1 + T_1), \\ f_t &= \beta_1^{(2)} x_{1t} + \beta_2^{(2)} x_{2t} + \cdots + \beta_p^{(2)} x_{pt} \qquad (t = \tau_2, \tau_2 + 1, \ldots, \tau_2 + T_2) \end{aligned}$$

3 Because a correlation matrix can be regarded as the covariance matrix of standardised data, we use the term covariance matrix and correlation matrix interchangeably.

4 In the following discussion we use the definition (2.1) for simplicity. The definition (2.2) which incorporates the time series structure of the included variables will be discussed in Section 6.

(where the phases are denoted by the parenthesised superscripts). In these equations, the statistical similarity between the two phases would be judged by whether the corresponding coefficients in the two equations are statistically equal or not. This can be easily generalized to three or more phases.

In this discussion, however, there are two problems: the first is how to estimate the β coefficients, and the other is how to relate those expressions to the covariance structure. To answer these questions simultaneously, (partial) common principal component analysis, explained in the next subsection, is useful.

2.3. Hierarchy of similarities among covariance matrices

Flury (1988) defined a hierarchy of similarities among the structure of covariance matrices in terms of five levels. This similarity hierarchy recognises that while strict equality of covariance matrices may not exist it may nonetheless be the case that there may exist at least one linear combination of the underlying variables which remains statistically invariant across different groups (or episodes). The five levels are:

(i) equality; $\Sigma_1 = \Sigma_2 = \cdots = \Sigma_k$

(ii) proportionality; $\Sigma_i = \rho_i \Sigma_1$

(iii) common principal component model; $\Sigma_i = B\Lambda_i B'$,
where $B = (\vec{\beta}_1, \vec{\beta}_2, \ldots, \vec{\beta}_p)$, $\Lambda_i = \mathrm{diag}(\lambda_{i1}, \lambda_{i2}, \ldots, \lambda_{ip})$ with $\lambda_{ij} = \vec{\beta}_j' \Sigma_i \vec{\beta}_j$ and $\vec{\beta}_j$s denote the common eigenvectors.

(iv) partial common principal component model; $\Sigma_i = B^{(i)}\Lambda_i B^{(i)\prime}$,
where $B^{(i)} = (\vec{\beta}_1, \ldots, \vec{\beta}_q, \vec{\beta}_{q+1}^{(i)}, \ldots, \vec{\beta}_p^{(i)})$, $q \le p$ and the first q eigenvectors only are common.

(v) arbitrary,
where $\Sigma_i (i = 1, 2, \ldots, k)$ denotes the i-th covariance matrix in k groups (phases), $\vec{\beta}$ is eigenvector of order p, B and $B^{(i)}$ are orthogonal square matrices of order p, Λ_i is a diagonal square matrix of order p, and p is the number of variables. The strongest form of statistical similarity is (i), and the weakest is (iv).

The common principal component (CPC) model in level (iii) is proposed by Flury (1984). Principal components are common across groups if there exists a common orthogonal matrix whose columns are eigenvectors which are common for all groups. Note that in the usual principal component analysis (PCA), which assumes stage (v) implicitly, $B_i' \Sigma_i B_i = \Lambda_i$, where Λ_i is diagonal with eigenvalues of Σ_i. In normal PCA with $k = 2$, we can express the principal components using original data as follows:

$$f_{jt} = \beta_{1j}^{(1)} x_{1t} + \beta_{2j}^{(1)} x_{2t} + \cdots + \beta_{pj}^{(1)} x_{pt} \quad (t = \tau_1, \tau_1 + 1, \ldots, \tau_1 + T_1; j = 1, \ldots, p), \tag{2.3}$$

$$f_{jt} = \beta_{1j}^{(2)} x_{1t} + \beta_{2j}^{(2)} x_{2t} + \cdots + \beta_{pj}^{(2)} x_{pt} \quad (t = \tau_2, \tau_2 + 1, \ldots, \tau_2 + T_2; j = 1, \ldots, p). \tag{2.4}$$

Therefore, in the CPC model of level (iii), statistical similarity of covariance matrices is defined as

$$\beta_{ij}^{(1)} = \beta_{ij}^{(2)} = \beta_{ij} \qquad (i, j = 1, 2, \ldots, p) \tag{2.5}$$

by which (2.3) and (2.4) would be expressed without superscripts.

The CPC model assumes p common eigenvectors. But, in practice, we would not necessarily need to require all p eigenvectors to be the same in analysing statistical similarities in business cycles phases. In other words, the definition of statistical similarity in (2.5) may be somewhat too strict. We would argue that provided there exists at least one distinct linear combination of the component variables which is statistically invariant across two business cycle phases that some level of similarity exists across the phases. This situation corresponds to the case in which just one common eigenvector exists and this suggests that, for business cycle analysis purposes, partial common principal component analysis may be the more relevant approach.

Given this, using partial CPC model described by level (iv) similarity, we can express a definition of statistical similarity as:

$$\beta_{ij}^{(1)} = \beta_{ij}^{(2)} = \beta_{ij} \qquad (i = 1, 2, \ldots, p; j = 1, 2, \ldots, q). \tag{2.6}$$

Note that the difference between (2.5) and (2.6) is the range of j. And, using the expression involving the original data, we can express

$$\begin{aligned} f_{jt} &= \beta_{1j}x_{1t} + \beta_{2j}x_{2t} + \cdots + \beta_{pj}x_{pt} \\ &\qquad (j = 1, 2, \ldots, q \text{ for all } t), \\ f_{jt} &= \beta_{1j}^{(1)}x_{1t} + \beta_{2j}^{(1)}x_{2t} + \cdots + \beta_{pj}^{(1)}x_{pt} \\ &\qquad (j = q+1, \ldots, p, t = \tau_1, \tau_1 + 1, \ldots, \tau_1 + T_1) \\ f_{jt} &= \beta_{1j}^{(2)}x_{1t} + \beta_{2j}^{(2)}x_{2t} + \cdots + \beta_{pj}^{(2)}x_{pt} \\ &\qquad (j = q+1, \ldots, p, t = \tau_2, \tau_2 + 1, \ldots, \tau_2 + T_2). \end{aligned} \tag{2.7}$$

In equation (2.7), the greater is q, the greater the degree of statistical similarity the two phases would have. This means that for phases of business cycles to be regarded as, in some sense, statistically similar, we require at least level (iv) similarity with $q = 1$, and we believe this criteria might be useful for examining similarities of covariance matrices among business cycle phases. Of course, if CPCs exist it would be desirable if they accounted for a significant proportion of the total variance in the underlying component variables.

When statistical similarity is obtained for just $q = 1$, this would be analogous to the single index model adopted by Stock and Watson (1991, 1993) which in turn was theoretically based upon Sargent and Sims (1977). The larger is q(maximum = the number of variables in the computed covariance matrix), the stronger is the degree of similarity. In conclusion, it may be worthwhile to explicitly point out that what the approach basically amounts to is determining whether there exists at least one latent unobserved variable (a principal component), itself a linear combination of the variables under study, which remains statistically invariant across phases.

3. Testing similarities among covariance matrices

3.1. Testing equality of covariance matrices

To test the equality of k covariance matrices, corresponding to the strongest level of similarity as defined in level (i), the null and alternative hypotheses are:

$$H_0 : \Sigma_1 = \Sigma_2 = \cdots = \Sigma_k$$
$$H_A : \Sigma_i \neq \Sigma_j \qquad (i \neq j).$$

The likelihood ratio test statistics of above hypothesis is,

$$\lambda_1 = \frac{\prod_{i=1}^{k} |S_i|^{T_i/2}}{|\sum_{i=1}^{k} S_i|^{T/2}} \cdot \frac{T^{kT/2}}{\prod_{i=1}^{k} T_i^{kT_i/2}} ,$$

where S_i is i-th sample covariance matrix, T_i is the number of observations in the i-th group (phase) and $T = \sum_{i=1}^{k} T_i$. Since the asymptotic null distribution of $-2\log\lambda_1$ is χ^2 distribution with $(k-1)p(p+1)/2$ degrees of freedom, the usual procedure for testing the null hypothesis can be carried out[5]. Some test statistics which adjust λ_1 have also been proposed.

However, in examining statistical similarities of business cycle phases defined by (2.6), we do not require strict equality of covariance matrices. Clearly, if covariance matrices are equal, identical eigenvectors are always obtained, but even if covariance matrices are not strictly equal, there nevertheless could be cases in which one or more (statistically) identical eigenvectors are obtained, viz. the case of the existence of (partial) CPC in the covariance matrices. As noted in the previous section, statistical similarity of phases defined in this paper requires only that at least one eigenvector exists which is statistically the same across alternative phases. See Katsuura (1997) for examples in which H_0 is rejected by the test for equality of covariance matrices, but statistically similar (statistically identical) eigenvectors are obtained for alternative sample periods in stock price data. Nevertheless, for completeness and for expositional purposes, in the next section we will also report the results of this test for the business cycle data which we analyse.

3.2. Testing common principal components hypothesis[6]

When testing the existence of CPC among k covariance matrices, the following is the null hypothesis to be tested:

$$H_{CPC} : B'\Sigma_i B = \Lambda_i$$
$$H_A : \text{not } H_{cpc}.$$

The test statistic for this hypothesis is,

$$\chi^2_{CPC} = -2\log\frac{L(\hat{\Sigma}_1, \hat{\Sigma}_2, \ldots, \hat{\Sigma}_k)}{L(S_1, S_2, \ldots, S_k)} \tag{3.1}$$

5 See Anderson (1984) pp. 404-426.

6 On testing level (ii) of similarity, viz. proportionality of covariance matrices, see Flury (1988) chap. 5. We omit this issue in this paper because we are not interested in proportionality of covariance matrices in terms of business cycle analysis using growth rates of the component variables.

$$= \sum_{i=1}^{k} n_i \log \frac{|\operatorname{diag} F_i|}{|F_i|} = \sum_{i=1}^{k} n_i \log \frac{\prod_{j=1}^{p} f_{jj}^{(i)}}{\prod_{j=1}^{p} l_{ij}}$$

where $\hat{\Sigma}$ is the maximum likelihood estimate (MLE) of Σ in the CPC hypothesis, $F_i = \hat{B}' S_i \hat{B}$, $\hat{B}$ is the maximum likelihood estimator of CPC, $f_{jj}^{(i)}$ is the (j,j) element of F_i and l_{ij} is the j-th eigenvalue of F_i. The asymptotic null distribution of χ^2_{CPC} is χ^2 distribution with $(k-1)p(p-1)/2$. For further details of the estimation and the statistical test in CPCA, see Flury (1984, 1988).

If the null hypothesis can not be rejected, the CPC hypothesis is accepted, implying that, according to the above definition, similarities among phases are found to exist. Katsuura (1997) used this statistic for examining statistical similarities defined by (2.5) for the Japanese diffusion index components, but no business cycle statistical similarity was found[7].

3.3. Testing the partial common principal components hypothesis

In Section 2, we explained that statistical similarities among phases of business cycles are captured by (2.6) for at least $q = 1$. Therefore, it is appropriate to test the existence of such similarity based on the partial CPC model (level (iv) similarity). Estimation and test statistics are similar to that of the CPC model. For further details of partial CPC concept, estimation and statistical test, see Flury (1987, 1988).

The partial CPC hypothesis and its alternative are

$$H_{CPC}(q) : B^{(i)\prime} \Sigma_i B^{(i)} = \Lambda_i$$
$$H_A : \text{not } H_{cpc}(q),$$

and the test statistic for this hypothesis is

$$\chi^2_{CPC}(q) = -2 \log \frac{L(\tilde{\Sigma}_1, \tilde{\Sigma}_2, \ldots, \tilde{\Sigma}_k)}{L(S_1, S_2, \ldots, S_k)} = \sum_{i=1}^{k} n_i \log \frac{|\tilde{\Sigma}_t|}{|S_i|}, \tag{3.2}$$

where $\sim$ denotes the MLE of the partial CPC model assuming a particular value for q. The asymptotic distribution of $\chi^2_{CPC}(q)$ under the partial CPC hypothesis is χ^2 distribution with $(k-1)q(2p-q-1)/2$ degrees of freedom. The test procedure is the same as that of CPC except for the selection of q , the number of partial CPC. It is possible to determine the value of q in a step-wise fashion, for example, in accordance with the magnitude of diagonal elements of estimated Λ_i . Note that the cases of $q = p-1$ and $q = p$ are equivalent and correspond to the CPC case in the previous subsection.

MLE of $B^{(i)}$, which includes both common and specific eigenvectors, provides information about the influence of individual variables on partial CPCs. As such, the identified variable may be expected to contribute significantly to business cycle fluctuations as expressed by (2.7). These estimated eigenvectors

7 Katsuura (1997) uses levels data —not transformed into growth rates— to test similarities. Therefore, we can not directly compare this earlier work with the results in the following chapter.

will therefore be useful in interpreting the characteristics of the phases as well as those of the obtained principal components.

In summary, we use the test statistic (3.2) to examine statistical similarities among business cycle phases as defined in (2.6). The reason we use the test based on partial CPCA, as opposed to ordinary CPCA, relates to the definition of statistical similarity used here and its relation to the concept of the business cycle outlined earlier. We defined statistical similarity among phases as the existence of at least one common eigenvector across phases. This corresponds to the case of $q = 1$ in (2.6). Also, as explained in the next section, we select common eigenvectors in a step-wise fashion starting from the common component with the largest explanatory power. Also, if we implement the test in a step-wise fashion for as $q = 1, 2, \ldots, p$, the case of $q = p$ is equivalent to a test based on ordinary CPCA which can test for the stronger form of similarity. For this reason, the test based on partial CPC is adopted as the test for examining statistical similarity of phases of business cycles.

4. Empirical results using the coincident index components

In this section, we examine statistical similarities between the present and past US expansions using the ECRI (Economic Cycle Research Institute) coincident index components.

The dates of the phases, i.e., contractions and expansions, are determined in the U.S. by the NBER (National Bureau of Economic Research). The official chronology and durations of phases are shown in Table 2 on both a monthly and quarterly basis. The empirical analyses in the present and following sections are based on these chronologies. Available data are categorised according to the expansion dates, and, for each expansion, the relevant sample covariance matrix is calculated.

4.1. Data

As discussed in Section 2, a widely accepted definition of empirical business cycles employs the components of a coincident composite index to determine a

Table 2. Monthly and Quarterly Chronology of Business Cycles in the U.S.

(a) Monthly

Trough	Peak	expansion	contraction	total
	Nov-48		11	
Oct-49	Jul-53	45	10	55
May-54	Aug-57	39	8	47
Apr-58	Apr-60	24	10	34
Feb-61	Dec-69	106*	11	11
Nov-70	Nov-73	36*	16	16
Mar-75	Jan-80	58*	6	6
Jul-80	Jul-81	11	16	27
Nov-82	Jul-90	92*	8	8
Mar-91	(Jan-99)	94*		

(b) Quarterly

Trough	Peak	expansion	contraction	total
	48:Q4		4	
49:Q4	53:Q3	15	3	18
54:Q2	57:Q3	13	3	16
58:Q2	60:Q2	8	3	11
61:Q1	69:Q4	35*	4	4
70:Q4	73:Q4	12	5	17
75:Q1	80:Q1	20*	2	2
80:Q3	81:Q3	4	5	9
82:Q4	90:Q3	31*	2	2
91:Q1	(98:Q4)	31*		

Note: Date and quarters in parenthesis are not the official chronology, but the end of the data. Figures are number of months or quarters of phases and whole cycles. * denotes the phases used in this analysis.

phase chronology. Therefore, we use these variables to examine statistical similarities among phases in this section. One advantage is that most of these variables are available monthly which allows us to use a greater number of observations than in the case of quarterly data as in the following section.

The variables used in this section are five of the components of the ECRI coincident composite index. Although the ECRI coincident composite index consists of six components, we use only the components with monthly observations and the sole quarterly based component (real GDP) is excluded. It should also be noted that, apart from the unemployment rate, the other variables are those used in the construction of the US coincident index as was once computed and published for many years by the US Department of Commerce and which is now published by the Conference Board. This provides strong motivation for using these variables as the basis for our monthly investigation. It is also the case that the ECRI coincident index is known to have turning points which are very similar to the official US business cycle chronology (see Layton (1998)) lending support to incorporating unemployment into the analysis. The components used here are then as follows;

Personal income (C1)
Industrial production (C2)
Manufacturing and trade sales (C3)
Unemployment rate (C4)
Non-farm employees (C5).

All of these series are seasonally adjusted as is the usual case in constructing a composite index. If unadjusted data are used, the seasonality would be detected as a component of similarities, which is not desirable for the purpose of our analysis. This paper is concerned with examining statistical similarities in business cycles; i.e. our interest is focused upon the cyclical fluctuations in the business cycle variables.

The available data spanned January 1959 to January 1999 and the subsample periods are determined by the monthly chronology shown in Table 2. However, since the distribution of the test statistic in (3.2) requires a large number of observations, we only use the expansions whose durations are greater than thirty months[8]. As a result, we analyse five expansions, viz., March 1961-December 1969 (E4; 60s), December 1970-November 1973 (E5; early 70s), April 1975-January 1980 (E6; late 70s), December 1982-July 1990 (E8; 80s) and the present expansion, April 1991-(E9; 90s) [9]. The components are transformed into month-to-month growth rates except for C4. Since C4, unemployment rate, is already expressed as a ratio, it is transformed into month-to-month differences. By transforming the data into growth rates and differences we should ensure the resulting data have the property of stationarity. In estimating the covariance

8 In order to use as many as possible of the same expansions used in the quarterly analysis in the following section, we adopt in that section the criteria of incorporating expansions with more than twenty quarters. Nevertheless, the outcome is that we are able only to analyse four expansions in the quarterly analysis as compared with five expansions in the monthly analysis.

9 The figures after E (Expansion) correspond to the orders of the post war expansions shown in Table 2.

matrix and PCs, stationarity is an important property. See Kariya (1993) on the role of the stationarity in PCA.

It is possible to apply PCA or CPCA to correlation matrices as well as covariance matrices. The use of correlation matrices is often effective when the variables are measured in a variety of units. However, as we transform the data into growth rates, the problem of different units does not exist here. The reason for using covariance matrices, rather than correlation matrices, is that information on the variances of variables is very important for characterising business cycles. For example, investment is often said to have an important role in business cycle fluctuations partly because of its great variability. If we use correlation matrices, the variances of component variables are all set to one, which means important information relating to their variance is disregarded with each variable effectively regarded as being equally variable within and between expansions. Therefore, if we had used correlation matrices, it may have been the case that even stronger similarity would have been observed because of the identical variances [10]. For this reason, we use covariance matrices as the basis of the test for statistical similarity of business cycle phases.

4.2. Results using coincident index components

Using the calculated covariance matrices for each expansion, the results of a standard PCA for each expansion are shown in Table 3. Note that the covariance matrices were calculated for the zero lag only. As all variables under study are regarded as "coincident" —in the sense of their turning points being within three months of each other— choosing to calculate contemporaneous covariations is quite natural. Nonetheless, one can argue that non-zero lags should be included also. This is investigated in Section 6 below.

The principal components (PCs) are ordered according to the magnitude of eigenvalues in the tables. Interpretation of the components should be carried out in a careful way by looking at the relative magnitudes of the elements of the eigenvectors. In particular, we need to pay special attention to the largest eigenvector element. For example, for the expansions in the 1960s and 1990s, the first PC in 1960s corresponds to the second PC in 1990s because the largest element in both of these eigenvectors corresponds to the same underlying variable, viz. C3. In general, C3, manufacturing and trade sales, has the greatest value in the first PC for each expansion except in the case of 1990s where it is observed to be the largest element in the second PC. C2, industrial production, has the second largest weight in the first PC for all expansions except 1990s where it has the second largest weight in the second PC. If we interpret these PCs as indicative of the business cycle, fluctuations in C3 and C2 are relatively more important. This interpretation, in which business cycle fluctuations are captured mainly by variation in industrial production and manufacturing and trade sales, seems to be intuitive and reasonable.

10 In fact, we also calculated the test statistics by using correlation matrices, and this tendency was in evidence. For example, 1980s and 1990s were similar at the value of $q = 5$ using correlation matrices, and their similarity is observed only for $q = 1$ by covariance matrices as shown in Table 4.

Table 3. Results of PCA for Each Expansion for Coincident Composite Index Components.

(a) E4 (Mar.61-Dec.69)

	1st PC	2nd PC	3rd PC	4th PC	5th PC
C1	0.0437	0.1370	0.9758	−0.1353	−0.0941
C2	0.4040	0.8994	−0.1540	−0.0647	−0.0071
C3	0.9115	−0.4097	0.0107	−0.0264	−0.0237
C4	0.0129	−0.0080	0.0559	−0.2697	0.9612
C5	0.0627	0.0671	0.1446	0.9508	0.2581
eigenvalue	1.1106	0.3784	0.1220	0.0272	0.0421
proportion	0.6609	0.2252	0.0726	0.0162	0.0251
cum.prop.	0.6609	0.8861	0.9587	0.9749	1.0000

(b) E5 (Dec.70-Nov.73)

	1st PC	2nd PC	3rd PC	4th PC	5th PC
C1	0.2473	0.0462	−0.0917	0.2200	0.9380
C2	0.3848	−0.1294	−0.0825	0.8578	−0.3043
C3	0.8868	−0.0291	0.0453	−0.4419	−0.1243
C4	0.0034	−0.1685	0.9777	0.0957	0.0805
C5	0.0664	0.9757	0.1636	0.1067	−0.0745
eigenvalue	1.3331	0.0179	0.0539	−0.3309	0.2827
proportion	0.6604	0.0089	0.0267	0.1639	0.1401
cum.prop.	0.6604	0.6693	0.6960	0.8599	1.0000

(c) E6 (Apr.75-Jan.80)

	1st PC	2nd PC	3rd PC	4th PC	5th PC
C1	0.1520	−0.0145	−0.0786	0.0635	0.9831
C2	0.3854	−0.9159	−0.0663	−0.0511	−0.0750
C3	0.9064	0.3970	−0.0399	−0.0306	−0.1355
C4	−0.0106	−0.0018	−0.5616	0.8217	−0.0964
C5	0.0826	−0.0573	0.8200	0.5633	0.0155
eigenvalue	1.4736	0.3536	0.0272	0.0168	0.1194
proportion	0.7403	0.1776	0.0137	0.0085	0.0600
cum.prop.	0.7403	0.9179	0.9316	0.9400	1.0000

(d) E8 (Dec.82-Jul.90)

	1st PC	2nd PC	3rd PC	4th PC	5th PC
C1	0.0249	0.0572	0.9651	−0.2468	0.0610
C2	0.3383	0.9327	−0.0865	−0.0759	0.0490
C3	0.9389	−0.3435	−0.0097	−0.0218	0.0038
C4	−0.0247	−0.0525	−0.0681	−0.0350	0.9954
C5	0.0533	0.0783	0.2373	0.9652	0.0557
eigenvalue	1.2400	0.3079	0.1212	0.0192	0.0267
proportion	0.7231	0.1795	0.0707	0.0112	0.0155
cum.prop.	0.7231	0.9026	0.9732	0.9845	1.0000

(e) E9 (Apr.91-Jan.99)

	1st PC	2nd PC	3rd PC	4th PC	5th PC
C1	0.9561	−0.2925	−0.0183	−0.0066	0.0067
C2	0.0841	0.2158	0.9679	−0.0916	−0.0343
C3	0.2806	0.9314	−0.2320	0.0036	−0.0071
C4	0.0022	0.0195	0.0557	0.2770	0.9591
C5	0.0129	0.0095	0.0773	0.9565	−0.2809
Eigenvalue	0.7163	0.5609	0.1609	0.0110	0.0207
proportion	0.4873	0.3816	0.1095	0.0075	0.0141
cum.prop.	0.4873	0.8689	0.9784	0.9859	1.0000

Note: C1; personal income, C2; industrial production, C3; manufacturing and trade sales; C4; unemployment rate, C5; non-farm employees.

The period used in the 1990s expansion, of course, does not include the whole expansion. This may be the cause of the different pattern of the elements of eigenvectors in the PCs. Alternatively, we could conclude that the core features of the present expansion are slightly different to those of the other expansions.

We calculate the test statistic (3.2) for examining statistical similarity of expansions for different values of q in Table 4. In selecting eigenvectors for a value of q, we extract and define the j-th partial CPC according to the j-th largest sum of corresponding diagonal elements in CPCA, viz. $\sum_{i=1}^{k} f_{jj}^{(i)}$ in (3.1), because the estimators and algorithm of partial CPCA are based on CPCA. Moreover, the values of q are increased in step-wise fashion from one to p. Recall that if $q = p$ or $p - 1$, it corresponds to the CPC model.

Calculated χ^2 values for test statistic (3.2) between the 1990s and other expansions for each value of q are shown in Table 4. A significant calculated value means that the partial CPC hypothesis, $H_{CPC}(q)$, is rejected, which implies no statistical similarity exists between the relevant expansions. As we have defined statistical similarity among phases as the existence of at least one partial CPC, the expansion in the late 1970s is not statistically similar to the present expansion. However, the other expansions, the 1960s, early 1970s and 1980s expansions can be said to be statistically similar to the 1990s expansion. Looking at the value of q, the maximum possible value of $q = 5$ is obtained in testing for statistical

Table 4. χ^2 Statistics for Testing Similarities for Coincident Composite Index Components.

Expansions	$q=1$	$q=2$	$q=3$	$q=4$ $(q=5)$	accepted number of CPC	$-2\log(\lambda_1)$
60s-90s	5.9544	13.5628	14.7493	18.3027	5	553.5911*
Early 70s-90s	5.0328	7.9354	9.1759	12.3227	5	420.5473*
Late 70s-90s	11.6799*	18.0128*	19.5802*	20.2862*	0	436.5553*
80s-90s	6.6287	26.2592*	29.9854*	29.9879*	1	520.1603*
critical value	9.4877	14.0671	16.9190	18.3070		24.9958

Note: * denotes the finding that the partial CPC hypothesis or, in the case of column seven, the hypothesis of strict covariance matrix equality, is rejected at 5% significance level. Column six is the maximum accepted number of significant partial CPC (q). Column seven presents the test statistics for testing for the existence of level (i) similarity, viz. the strict equality of the covariance matrices of the phases.

similarity between both the 1990s and the 1960s and 1990s and early 1970s expansions. This corresponds to level (iii) similarity in Section 2, implying the statistical similarity between the 1990s expansion and that of the 1960s and early 1970s expansions is very strong. Statistical similarity between the present and the immediately previous expansion also exists, but it is not as strong. Also, to emphasise the advantages of our proposed definition of similarity, the test statistics, $-2\log\lambda_1$, for the strict statistical equality of covariance matrices, viz. level (i) similarity, are reported in the last column of Table 4. Using these statistics, no similarities (equalities) are accepted which would imply that level (i) similarity is absent as far as the business cycle expansions under study are concerned. However, as we argued earlier, level (i) similarity —strict equality of covariance matrices of different phases— is too strict a definition of similarity than is required for business cycle analysis purposes.

These similarities or dissimilarities between expansions may be at least partially explained by the general economic situations prevailing during the expansions in question. For example, the expansion in the late 1970s began after the previous deep contraction caused by the first oil crisis occurring at the end of 1973. The US economy subsequently recovered very rapidly from the deep trough in March 1975. We can see a similar situation in the expansion in the 1980s after the second oil crisis and very tight monetary policy in the early 1980s brought about the very pronounced contraction in 1981/82. Again, the US economy recovered very rapidly from the trough in November 1982. For example, using the unemployment rate (C4) as one indicator of the speed and strength of employment growth in each expansion, it was decreasing from the very earliest months of each of these two expansions and continued to fall throughout almost the entire duration of the expansions.

However, as the contraction preceding the present expansion was not as deep or as protracted as the above contractions, during its early months, the rate of growth in the economy was considerably more modest. Considering the unemployment rate again, it continued to increase for some 20 months after the official NBER trough in March 1991 until November 1992, and only then did it begin to decline. In the case of the 1960s expansion, although the economy

Table 5. Estimated (Partial) CPC for Coincident Composite Index Components.

(a) E4 (60s) and E9(90s) ($q = 5$)

	1st CPC	2nd CPC	3rd CPC	4th CPC	5th CPC
C1	0.0347	0.9981	0.0485	0.0050	−0.0122
C2	0.3135	−0.0578	0.9438	−0.0040	−0.0869
C3	0.9482	−0.0177	−0.3170	−0.0078	−0.0101
C4	0.0156	−0.0039	0.0165	0.9809	0.1933
C5	0.0351	0.0079	0.0780	−0.1944	0.9771
E4					
diagonal	1.1028	0.1221	0.3832	0.0394	0.0328
proportion	0.6563	0.0727	0.2281	0.0235	0.0195
cum.prop.	0.6563	0.7290	0.9570	0.9805	1.0000
E9					
diagonal	0.5743	0.6976	0.1659	0.0209	0.0112
proportion	0.3907	0.4746	0.1129	0.0142	0.0076
cum.prop.	0.3907	0.8653	0.9782	0.9924	1.0000

(b) E5 (early 70s) and E9 (90s) ($q = 5$)

	1st CPC	2nd CPC	3rd CPC	4th CPC	5th CPC
C1	0.2402	0.9704	−0.0227	0.0016	−0.0063
C2	0.2762	−0.0466	0.9529	−0.0507	−0.1048
C3	0.9300	−0.2368	−0.2812	0.0015	−0.0011
C4	0.0146	−0.0035	0.0563	0.9959	0.0698
C5	0.0307	0.0012	0.0963	−0.0754	0.9920
E5					
diagonal	1.3177	0.2834	0.3407	0.0554	0.0213
proportion	0.7842	0.1687	0.2028	0.0329	0.0127
cum.prop.	0.7842	0.9528	1.1556	1.1886	1.2012
E9					
diagonal	0.6009	0.6751	0.1620	0.0204	0.0116
proportion	0.4088	0.4593	0.1102	0.0138	0.0079
cum.prop.	0.4088	0.8681	0.9783	0.9921	1.0000

(c) E8 (80s) and E9 (90s) ($q = 1$)

	common PC	Specific PC							
		E8				E9			
	1st CPC	2nd PC	3rd PC	4th PC	5th PC	2nd PC	3rd PC	4th PC	5th PC
C1	0.0225	0.0586	0.9650	−0.2452	0.0688	0.9996	−0.0144	−0.0061	0.0052
C2	0.2912	0.9487	−0.0864	−0.0726	0.0495	0.0066	0.9517	−0.0958	−0.0169
C3	0.9557	−0.2940	−0.0059	−0.0083	−0.0094	−0.0259	−0.2922	−0.0038	0.0240
C4	−0.0079	−0.0562	−0.0688	−0.0053	0.9960	−0.0023	0.0488	0.2919	0.9552
C5	0.0357	0.0833	0.2379	0.9667	0.0265	0.0077	0.0796	0.9516	−0.2946
E4									
diagonal	1.2370	0.3104	0.1212	0.0195	0.0269				
proportion	0.7213	0.1810	0.0707	0.0114	0.0157				
cum.prop.	0.7213	0.9023	0.9730	0.9843	1.0000				
E9									
diagonal	0.5741					0.7009	0.1625	0.0110	0.0214
proportion	0.3906					0.4768	0.1106	0.0075	0.0145
cum.prop.	0.3906					0.8674	0.9779	0.9855	1.0000

Note: The results are only shown for the maximum number of q in Table 4. "Diagonal" denotes diagonal elements of estimated Λ_i in (partial) CPC model.

expanded in the early part of the expansion in a similar way to the late 1970s and 1980s expansions with unemployment decreasing from the trough, the economy did not always expand sufficiently rapidly throughout the entire expansion; the unemployment rate increased from the beginning of 1964 to mid 1965, then began to decline again[11]. These unique features of the expansions in question may explain the strong statistical similarities between both the 1990s and 1960s and the 1990s and early 1970s expansions, the weaker similarity between the 1990s and 1980s expansions, and the absence of any similarity between the 1990s and late 1970s expansions.

We should note here in passing that transitivity does not always hold in this type of analysis; for example, the χ^2 statistic relating to the late 1970s and 1980s expansions is 7.517 for $q = 1$ (not shown in Table 4), and similarities between the late 1970s and 1980s expansions and between the 1980s and 1990s expansions exist, but no statistical similarity exists between the late 1970s and 1990s expansion. This is not unusual because of the possibility of the existence

11 The discussion of the unemployment rate or the strength of recovery is consistent with the results obtained in Layton and Smith (2000). They categorised expansions into two types, viz. a "slow or normal growth" expansion and a "fast growth" expansion. For example, the expansion in 1980s began as a "fast growth" expansion and then switched into a "normal or slow growth" expansion, whilst the 1990s recovery was categorised as a "slow or normal growth" expansion for the first few years. Such characterisations of these two expansions are consistent with the weaker similarity found in this analysis.

of different sorts of pairwise statistical similarities.

Now, looking at the CPC results for the 1960s and 1990s expansions in Table 5, the first CPC seems to be a mixed vector of the first PC in the 1960s expansion and the second PC in the 1990s expansion because the coefficient of C3 is much greater than those of the other variables. Furthermore, recall that this first CPC by definition accounts for the largest proportion of aggregate variation across both phases (even though in the case of 1990s the percentage variation explained is only .39). Given this, and the fact that it is statistically significantly common, the implication is that it is important as a representation of the US business cycle. In this representation C3, manufacturing and trade sales, is found to be an important underlying component variable, not only for the first CPC of the 1960s-1990s comparison, but also for the first CPC for each of the other two pairwise comparisons (see Table 5).

5. Empirical results for the GDP components

5.1. Data

In this section, we use the GDP components to examine statistical similarities among expansions. As GDP and its components are published quarterly by the Bureau of Economic Analysis (BEA), we refer to the quarterly chronology presented in Panel (b) of Table 2. We use four components of GDP in the analysis, viz. private final consumption expenditure (C), gross domestic capital formation (I), government final consumption expenditure (G) and exports (X). Although it is possible to use net exports (exports minus imports) as the fourth variable, it often takes negative values since the 1980s and the variability in its growth rate is extremely high. We could also use imports as the fifth variable. However, as it does not represent products produced in the economy it is difficult to interpret how it is directly relevant as far as the business cycle is concerned. Therefore, we use the above four GDP components only. All of the components are seasonally adjusted as in the case of composite index components.

As mentioned in the previous section, we investigate only the expansions whose periods are at least twenty quarters. According to this criterion, four expansions, viz. 1961:Q1-1969:Q4 (E4), 1975:Q2-1980:Q1 (E6), 1982:Q4-1990:Q3 (E8) and the present expansion, 1991:Q2-(E9) are used in this analysis. The duration of the expansion in the early 1970s, which was found to have statistical similarity to the present expansion in the previous section, is too short to examine in this quarterly analysis. Though the expansion in the late 1970s also comprises a relatively small number of observations (20 quarters), it is the only expansion which was not statistically similar to the present expansion in the case of the coincident composite index components and we therefore include it to further clarify the situation. As before, the data for calculating covariance matrices are transformed into quarter-to-quarter growth rates.

5.2. Results for GDP components

Analogous results for the US GDP components are shown in Tables 6–8. From Table 6, the first PC in the 1960s expansion seems to reflect export variation. For the other three expansions the first PC is interpreted as representing investment fluctuations, while the second PC is related mainly to export fluctuations. Intuitively this is because the variability in investment and exports is usually much greater than that of consumption or government expenditure. Together, these first two PCs account for about 95% of total variation in all four components for each of the expansions under study.

Calculated χ^2 test statistics in Table 7 imply that the present expansion is statistically similar to the expansions in the 1960s and 1980s but is again statistically dissimilar to the expansion in the late 1970s. Whilst these results are broadly consistent with those in the previous section including the results of the test for strict statistical equality between covariance matrices, there are some interesting differences. One issue is the fact that, in Table 7, the null hypothesis of CPC is rejected for $q = 1$ but paradoxically is accepted for $q = 2$ and 3. This is logically inconsistent and is most likely explained by the relatively small sample size available for the late 1970s expansion (viz. 20 quarters). A second issue is that the statistical similarity between the 1980s and 1990s expansions is now found to be much stronger than in the case of the coincident composite index components analysis.

Each of the 1960s-1990s and the 1980s-1990s expansions comparisons shows

Table 6. Results of PCA for Each Expansion for GDP Components.

(a) E4 (60s)

	1st PC	2nd PC	3rd PC	4th PC
C	−0.0098	0.0189	0.0661	0.9976
I	−0.3270	0.9447	0.0116	−0.0219
G	0.0134	−0.0091	0.9977	−0.0658
X	0.9449	0.3273	−0.0095	0.0037
eigenvalue	38.2685	11.0777	1.4927	0.4014
proportion	0.7468	0.2162	0.0291	0.0078
cum.prop.	0.7468	0.9630	0.9922	1.0000

(b) E6 (late 70s)

	1st PC	2nd PC	3rd PC	4th PC
C	0.0508	−0.0440	−0.3145	0.9469
I	0.9351	−0.3453	−0.0268	−0.0751
G	0.1022	0.1365	0.9349	0.3113
X	0.3355	0.9275	−0.1625	−0.0289
eigenvalue	15.5950	9.7270	0.3899	0.3022
proportion	0.5995	0.3739	0.0150	0.0116
cum.prop.	0.5995	0.9734	0.9884	1.0000

(c) E8 (80s)

	1st PC	2nd PC	3rd PC	4th PC
C	0.0017	−0.0558	0.1589	0.9857
I	0.9992	−0.0042	0.0394	−0.0083
G	−0.0403	−0.2031	0.9640	−0.1668
X	−0.0040	0.9776	0.2095	0.0216
eigenvalue	17.0645	2.8108	0.8175	0.2699
proportion	0.8140	0.1341	0.0390	0.0129
cum.prop.	0.8140	0.9481	0.9871	1.0000

(d) E9 (90s)

	1st PC	2nd PC	3rd PC	4th PC
C	−0.0145	−0.0620	0.9574	0.2818
I	0.9103	0.3929	0.0009	0.1304
G	−0.1284	−0.0173	−0.2828	0.9504
X	−0.3933	0.9173	0.0590	−0.0189
eigenvalue	5.4873	4.3225	0.1207	0.5405
proportion	0.5240	0.4128	0.0115	0.0516
cum.prop.	0.5240	0.9369	0.9484	1.0000

Table 7. χ^2 Statistics for Testing Similarities for GDP Components.

Expansions	$q = 1$	$q = 2$	$q = 3$ ($q = 4$)	accepted number of CPC	$-2\log(\lambda_1)$
60s-90s	2.2678	4.8782	6.3809	4	141.4050*
Late 70s-90s	9.3463*	9.7430	9.8909	0	115.1431*
80s-90s	1.8811	4.4749	5.0003	4	109.0858*
critical value	7.8147	11.0705	12.5916		18.3070

Note: See notes in Table 4.

Table 8. Estimated CPC for GDP Components.

(a) E4 (60s) and E9 (90s) ($q = 4$)

	1st CPC	2nd CPC	3rd CPC	4th CPC
C	−0.0164	−0.0192	0.1899	0.9815
I	−0.3247	0.9432	0.0706	−0.0006
G	0.0275	−0.0640	0.9793	−0.1903
X	0.9453	0.3255	−0.0009	0.0224
E4				
diagonal	38.2594	11.0337	1.5176	0.4295
proportion	0.7467	0.2153	0.0296	0.0084
cum.prop.	0.7467	0.9620	0.9916	1.0000
E9				
diagonal	4.8364	4.9474	0.5569	0.1303
proportion	0.4619	0.4725	0.0532	0.0124
cum.prop.	0.4619	0.9344	0.9876	1.0000

(b) E8 (80s) and E9 (90s) ($q = 4$)

	1st CPC	2nd CPC	3rd CPC	4th CPC
C	−0.0118	−0.0513	0.2373	0.9700
I	0.9977	0.0062	0.0679	−0.0042
G	−0.0669	−0.0063	0.9689	−0.2382
X	−0.0072	0.9986	0.0179	0.0483
E8				
diagonal	17.0497	2.7329	0.9026	0.2774
proportion	0.8133	0.1304	0.0431	0.0132
cum.prop.	0.8133	0.9437	0.9868	1.0000
E9				
diagonal	5.2925	4.4895	0.5665	0.1225
proportion	0.5054	0.4288	0.0541	0.0117
cum.prop.	0.5054	0.9342	0.9883	1.0000

Note: See notes in Table 5.

Table 9. χ^2 Statistics for Testing Similarities for Quarterly Coincident CI Components.

Expansions	$q = 1$	$q = 2$	$q = 3$	$q = 4$ ($q = 5$)	accepted number of CPC
60s-90s	6.3380	11.6261	12.4080	18.0450	5
Late 70s-90s	0.9628	5.5680	8.9999	24.1092*	3
80s-90s	3.9416	15.0208*	16.7418	17.0350	1
critical value	9.4877	14.0671	16.9190	18.3070	

Note: See notes in Table 4.

strong similarity and corresponds to level (iii) similarity. The estimated CPCs are presented in Table 8. For the 1960s and 1990s expansions, the first CPC consists of the first PC in the 1960s and the second PC in the 1990s which seems to represent export variation, and the second CPC is again interpreted as representing investment variation. For the 1980s and 1990s expansions, the first CPC can be interpreted as loading heavily on the investment variable for both expansions, while the second CPC mainly reflects export variation. Again, this is quite consistent with the commonly held view that cyclical fluctuations in the economy are the result mainly of investment and foreign trade fluctuations.

As noted above, an interesting issue pertains to the relatively stronger similarity found in this analysis between the 1980s and 1990s expansions compared with that in the previous section. The stronger similarity is most likely caused by the different definition of the business cycle implicit in the analyses in the two sections. In the previous section, although GDP is not included because it is a quarterly series, it is certainly regarded as one indicator of the business cycle, but only one of several. The finding that the 1960s expansion is very similar to the 1990s expansion in this section is obtained only in terms of the components of this one single indicator of the business cycle. Therefore, the result is not inconsistent with the analysis of the composite index components in which the statistical similarity between the 1960s and 1990s expansions was found to exist but was somewhat weaker. Importantly, these components include employment measures which are considered by many business cycle analysts to be extremely important aspects of the business cycle. We have already discussed the fact that the patterns of employment variation were quite different between the 1980s and

1990s expansions. Obviously this differential employment aspect is not captured in the GDP components analysis. Furthermore, another reason for the stronger similarity using the GDP components is due to growth in investment (whose weight is the largest in the first CPC) increasing relatively rapidly during the early period of both expansions and then, in both cases, its growth moderating considerably in the later periods of the expansions.

Moreover, another reason for the stronger similarity in GDP components would be explained by aggregation issues. The quarterly series, in general, could be considered to contain less irregular fluctuations than monthly series because of smoothing effects. To check this, the same tests are conducted on quarterly series of coincident composite index components. Each component (C1–C5) is aggregated by using a three month moving average, then they are transformed into growth rates, and similarities are tested using the same procedures. The results are shown in Table 9 (the result for early 70s is omitted because of the small sample size). Comparing it to Table 4, while similarity between the expansions of the 60s and 90s is the same, similarity between the late 70s and 90s expansions becomes stronger; accepted similarity is changed from $q = 1$ to $q = 3$. Moreover, the result for the expansions of the 80s and 90s become unstable for higher levels of q. From these observations, it could be said that quarterly series tend to provide somewhat stronger similarity than the monthly series.

6. Some extensions

6.1. Peak-to-peak cycle similarity

The suggested approach can also be used to test for the statistical similarities of whole cycles. Results are provided in Table 10. As is evident for the coincident index components, the results are broadly similar to the analysis of the expansions. The most substantive difference is that the strength of the similarity between the 1960s cycle and the 1990s cycle is much weaker ($q = 5$ to

Table 10. x^2 Statistics for Testing Similarities for Peak-to-Peak Cycles.

(a) Coincident Index Components

Cycles	$q = 1$	$q = 2$	$q = 3$	$q = 4$ ($q = 5$)	accepted number of CPC
60s-90s	5.3971	18.7874*	20.6705*	20.6992*	1
Early 70s-90s	3.3478	5.9115	6.3357	8.6920	5
Late 70s-90s	18.3448*	20.1739*	20.1988*	21.6388*	0
80s-90s	6.2643	26.4458	29.3956	31.0914*	1
critical value	9.4877	14.0671	16.9190	18.3070	

(b) GDP Components

Cycles	$q = 1$	$q = 2$	$q = 3$ ($q = 4$)	accepted number of CPC
60s-90s	3.4401	5.4664	7.9857	4
Late 70s-90s	5.1932	7.5103	7.9048	4
80s-90s	5.8812	7.3101	7.611	4
critical value	9.4877	14.0671	16.9190	

Note: See notes in Table 4.

$q = 1$) when data from each preceding recession are included in the analysis. In relation to the GDP components, the significant difference is that incorporating the preceding recession serves to strengthen quite substantially ($q = 0$ to $q = 4$) the degree of similarity in evidence between the late 1970s cycle and that of the 1990s. The above two material changes seem quite dramatic —especially in relation to the extent of the rather modest corresponding increases in data used in the peak-to-peak analysis. The explanation, however, lies in Tables 4 and 7 where it may be seen that the respective calculated test statistics really are quite close to their critical values.

6.2. Incorporating lags

The analysis of Section 4 and 5 was based on analysing contemporaneous covariation among the variables under study. Although a motivation and justification for this was provided in Section 4 it may nonetheless be of interest to see what impact the incorporation of lags may have on the analysis.

Equation (2.7) now becomes

$$f_{jt} = \sum_{i=0}^{r} \beta_{i,1j} x_{1,t-i} + \sum_{i=0}^{r} \beta_{i,2j} x_{2,t-i} + \cdots + \sum_{i=0}^{r} \beta_{i,pj} x_{p,t-i} \qquad (j = 1, 2, \ldots, q),$$

where r denotes the number of lags to be included. In the absence of any priors as to which lags should be incorporated, one method of doing this is to expand progressively the dimensionability of the covariance matrices being analysed to include increasing lag lengths and compute all the various lagged cross covariances. We would then investigate the equivalence of the resulting covariance matrices (level (i) similarity) or the existence of at least one linear combination of this much expanded set of variables which remains invariant across business cycle phases (level (iv) similarity). This is clearly a very strong requirement, particularly if many of the additional cross-covariances may in effect have population values of zero.

Table 11. Number of Similar CPCs for the 1960s and 90s Expansion Including Lagged Variables.

(a) Coincident CI Components

Lags (months)	$q=1$	$q=2$	$q=3$	$q=4$	$q=5$	$q=6$	$q=7$	accepted number of CPC	max. q
0	5.9544	13.5628	14.7493	18.3027	18.3027			5	5
3	24.0504	54.3031*	79.3991*	90.4929*	104.9417*	124.6516*	137.1755*	1	20
6	46.6368	90.8843*	144.9079*	168.5701*	201.1947*	217.6053*	251.3922*	1	35
8	70.2695*	130.0404*	207.4928*	242.4910*	285.1436*	303.6856*	352.1093*	0	45

(b) GDP Components

lags (quarters)	$q=1$	$q=2$	$q=3$	$q=4$	$q=5$	$q=6$	$q=7$	accepted number of CPC	max. q
0	2.2678	4.8782	6.3809	6.3808				4	4
1	8.0540	17.9598	24.5173	28.4057	30.2286	32.2809	32.3135	8	8
2	26.9493*	38.4903*	51.4725*	54.6518*	66.0592*	79.3825*	83.1900*	0	12
4	54.8821*	127.3820*	146.1098*	173.6406*	203.9141*	215.7908*	230.5611*	0	20

Note: Accepted number of q is determined by the same way as in Table 4 and 7. x^2 for $q > 7$ are omitted. Max. q means the possible number if the strongest similarity exist, which is calculated by (number of lags $+1) \times p$.

Table 12. χ^2 Statistics for Testing Similarities among Three Phases.

(a) Coincident Index Components.

Cycles	$q = 1$	$q = 2$	$q = 3$	$q = 4$ $(q = 5)$	accepted number of CPC
60s-Early 70s-90s	12.8866	22.0308	24.4149	29.9220	5
critical value	15.5073	23.6848	28.8693	31.4104	

(b) GDP Components

Cycles	$q = 1$	$q = 2$	$q = 3$ $(q = 4)$	accepted number of CPC
60s-80s-90s	9.21486	11.5287	13.1175	4
critical value	12.5916	18.3070	21.0261	

Note: See notes in Table 4.

In any event the results of such an analysis for just the comparison of the 1960s expansion and that of the 1990s (only this single comparison was done owing to computational considerations) are presented in Table 11. The results are probably largely as could be expected given the supposed 'coincident' nature of the variables used in this analysis. In the case of the coincident index components the degree of similarity is strongest with no lags, but nonetheless weak similarity remains in evidence when up to six months of lags of the variables are also included. In respect of the GDP components strong similarity exists for zero and 1 lag (three months) and thereafter similarity is not found.

6.3. Similarity among several phases

We have so far examined pairwise statistical similarities, but it is straightforward to extend the analysis to consider several (k) phases simultaneously within the framework of partial common principal component analysis described in Sections 2 and 3.

Table 12 shows results of the tests for examining "three-phase" similarities for selected phases for the composite index and GDP components. The phases selected are the 60s, early 70s and 90s expansions for the coincident index components, and the 60s, 80s and 90s expansions for the GDP components because of the strong similarities to the 90s expansion which were found in the previous sections. Although transitivity in similarities does not always hold in this type of analysis, Table 12 indicates that the highest level of similarity, viz. similarity for $q = 5$ for composite index components and $q = 4$ for GDP components, still exist for the three phases when considered as a group. The estimated first CPC for the composite index components (the results are not shown) is similar to that in Table 5, i.e. the weight for C3 is largest. For the GDP components the first CPC is explained by exports because of the largest weight on exports, while the second CPC is interpreted as the investment fluctuation. These results can be explained by similar economic interpretations described in the previous sections.

7. Conclusion

In this paper we have proposed a new method for examining the existence and strength of business cycle phase similarity. The method uses the framework of partial common principal component analysis and tests for the existence of common eigenvectors in the covariance matrices of business cycle component variables across different business cycle phases. We applied the method to the US coincident composite index components as well as to the components of GDP in order to examine the statistical similarities between the present long expansion of the 1990s and earlier expansions, particularly that of the 1960s.

It is sometimes argued that the expansions in the 1960s and the 1990s are similar because of their long durations. Duration, however, is only one aspect of an expansion and there are other dimensions of the issue of similarity which desirably should be examined. This paper focuses on statistical similarities of the covariance matrices of selected variables from different expansions. As a result, we found statistical similarity between the expansions in the 1960s and the 1990s in terms of the components of both GDP and the coincident composite index. Furthermore, the statistical similarity observed in the two expansions may be said to be quite strong owing to the fact that the level of observed similarity corresponds to a relatively high level in the hierarchy of covariance matrix similarity.

We also found that the present expansion is statistically similar to expansions in the early 1970s (only for the components of coincident composite index) and the 1980s, but not that of the late 1970s. We speculate that, in the latter case, the lack of similarity may be due to the quite unique supply side impacts of the first major oil crisis in the mid 1970s which brought about the contraction which immediately preceded the expansion of the later 1970s. Quite similar results are obtained for the GDP components and the coincident composite index components for the other three expansions. More stable results, however, were obtained for both sets of components in the comparison of the expansions of the 1990s and 1960s and this provides quite strong evidence of statistical similarity between the two expansions. However, the present expansion has not yet finished. When we include data for the whole expansion, the results may well be changed and, indeed, certain aspects of the similarity may be strengthened, e.g. in relation to the proportion of variation explained by the CPCs, the statistical similarity of the eigenvectors, etc.

There are some problems with the approach suggested in this paper. First, as we have to separate the data according to the business cycle chronology, some phases include only a small number of observations; especially if we use quarterly data. This can cause some results to be very imprecise. Moreover, the method is based on ML estimation which assumes large samples. If we were to try to examine the similarities between contractions or between contractions and expansions, this problem would become even more serious. A possible approach to this problem is to apply the least squares method developed by Clarkson (1988). However, the statistical properties of the least squares estimator are not as yet known.

Another problem is that the results will be affected very much by the component variables selected. This, in turn, is related to the definition of the business cycle which one employs. Nonetheless, we feel we have provided reasonable justification for our choice for this paper. Moreover, recall that the use of covariance matrices does not utilise potentially important information concerning phase similarity which could be inferred from an investigation of differences in component variable mean growth rates across phases. The method does not incorporate such considerations but, instead, investigates the degree of similarity in the patterns of covariation among the component variables around their respective means across different phases.

Despite these concerns, we believe the method as demonstrated in this paper could be a useful additional tool in business cycle analyses.

Acknowledgements

The authors are grateful to the editor and the referee for their helpful comments, and also to Mr. Anirvan Banerjee, the director of research of ECRI, for providing data of composite index components.

References

Anderson, T. W. (1984), *An Introduction to Multivariate Statistical Analysis* (2nd ed.), John Wiley & Sons, New York.

Blanchard, O. J. and M. W. Watson (1982), Are Business Cycles All Alike?, in Gordon, R. J. (eds), *The American Business Cycles*, 123–179, Univ. of Chicago Press, Chicago.

Burns, A. F. and W. C. Mitchell (1946), *Measuring Business Cycles*, NBER, New York.

Clarkson, D. B. (1988) A Least-squares Version of Algorithm AS 211, the F-G Diagonalization Algorithm, *Applied Statistics*, **37**, 317–321.

Diebold, F. X. and G. D. Rudebusch (1996), Measuring Business Cycles: A Modern Perspective, *The Review of Economics and Statistics*, **78**, 67–77.

Economic Planning Agency (2000), *White Paper on Japanese Economy 2000*, Printing Bureau, Ministry of Finance (in Japanese).

Fisher, L. A., G. Otto and G. M. Voss (1996), Australian Business Cycle Facts, *Australian Economic Papers*, **35**, 300–320.

Flury, B. N. (1984), Common Principal Components in k Groups, *Journal of the American Statistical Association*, **79**, 892–898.

Flury, B. N. (1987), Two Generalization of the Common Principal Component Analysis, *Biometrika*, **74**, 56–69.

Flury, B. N. (1988), *Common Principal Components and Related Multivariate Methods*, John Wiley & Sons, New York.

Hamilton, J. D. (1989), A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle, *Econometrica*, **57**, 357–384.

Harding, D. and A. Pagan (1999a), Knowing the cycle, *Melbourne Institute Working Paper*, No. 12/99.

Harding, D. and A. Pagan (1999b), Dissecting the cycle, *Melbourne Institute Working Paper*, No. 13/99.

Hess, G. D. and S. Iwata (1997), Measuring and Comparing Business-Cycle Features, *Journal of Business and Economic Statistics*, **15**, 432–444.

Kariya, T. (1986), *Econometric Analysis: Thoughts and Practice*, Tokyo, Toyokeizai-shinpousha (in Japanese).

Kariya, T. (1993), *Quantitative Methods for Portfolio Analysis*, Dordrecht, Kluwer Academic Publishers.

Katsuura, M. (1988), Analysis of Components in Diffusion Index by Common Principal Component Analysis, *Annals of Economic Studies* (*Graduate School of Economics*, *Waseda University*), **28**, 47–63 (in Japanese).

Katsuura, M. (1997), Tests on Structural Changes in Applying Principal Component Analysis to Time Series Data, *Meijo Shogaku* (*The Journal of Commerce and Economics, Meijo University*), **47**(3), 15–38 (in Japanese).

Kim, K., R. A. Buckle and V. B Hall (1994), Key Features of New Zealand Business Cycles, *The Economic Record*, **70**, 56–72.

King, R. G. and C. I. Plosser (1994), Real Business Cycles and the Test of the Adelmans, *Journal of Monetary Economics*, **33**, 405–438.

Layton, A. P. (1998), A Further Test on the Influence of Leading Indicators on the Probability of US Business Cycle Phase Shifts, *International Journal of Forecasting*, **14**, 63–70.

Layton, A. P. and D. Smith (2000), A Further Note on the Three Phases of the US Business Cycles, *Applied Economics*, **32**(9), 1133–1143.

Pagan, A. (1997), Towards an Understanding of Some Business Cycle Characteristics, *The Australian Economic Review*, **30**(1), 1–15.

Sargent, T. J. and C. A. Sims (1977), Business Cycle Modelling without Pretending to Have Too Much A Priori Economic Theory, in *New Methods in Business Cycle Research*, *Proceeding from a Conference*, Federal Reserve Bank of Minneapolis, 45–109.

Simkins, S. P. (1994), Do Real Business Cycle Models Really Exhibit Business Cycle Behavior?, *Journal of Monetary Economics*, **33**, 381–404.

Stock, J. H. and M. W. Watson (1991), A Probability Model of the Coincident Economic Indicators, in K. Lahiri and G. H. Moore (eds.), *Leading Economic Indicators*: *New Approaches and Forecasting Records*, Cambridge University Press, 63–89, Cambridge.

Stock, J. H. and M. W. Watson (1993), A Procedure for Predicting Recessions with Leading Indicators: Econometric Issues and Recent Experience, in J. H. Stock and M. W. Watson (eds.), *Business Cycles*, *Indicators and Forecasting*, University of Chicago Press, 255–284, Chicago.