JOURNAL OF APPLIED ECONOMETRICS, VOL. 11, 135–151 (1996)

MEASURING UNDERLYING ECONOMIC ACTIVITY

ANTHONY GARRATT
Bank of England, Monetary Analysis HO-4, Threadneedle Street, London EC2R 8AH

STEPHEN G. HALL
Centre for Economic Forecasting, London Business School, Sussex Place, Regent's Park, London NW1 4SA
Email: S.hall@lbs.bon.ac.uk

SUMMARY

Recently, interest in the methodology of constructing coincident economic indicators has been revived by the work of Stock and Watson (1989b). They adopt the framework of the state space form and Kalman filter in which to construct an optimal estimate of an unobserved component. This is interpreted as corresponding to underlying economic activity derived from a set of observed indicator variables. In this paper we apply the Stock and Watson approach to the UK where the observed indicator variables are those that make up the Central Statistical Office (CSO) coincident indicator. The time series properties of the indicator variables are examined and three of the five variables are first difference stationary and are cointegrated, the remaining two are stationary in levels. We then construct two alternative measures of economic activity, each of which deals with the different orders of stationarity of the variables. The first uses the levels of the observed component variables that allows for the cointegrating relationship. The second imposes stationarity on the I(1) variables before the estimation by taking first differences. The levels index is viewed as the preferred specification as it allows for the long-run relationships between the variables and has a superior statistical fit.

1. INTRODUCTION

The measurement of underlying economic activity is an exercise that has attracted interest in many areas for a very long time. The main early work was carried out by Mitchell and Burns (Mitchell and Burns, 1938; Burns and Mitchell, 1946) at the National Bureau of Economic Research (NBER) where lists of leading, coincident and lagging indicators were developed. This work was, much later in 1975, replicated in the UK where the Central Statistical Office (CSO) produces and publishes a set of similar economic indicators. These indicators have had a major influence on the timing and conduct of macroeconomic policy in both the USA and the UK. The most influential of the indicators is the coincident indicator which measures current economic activity. The intuition of such a measure is that while we have many macroeconomic series which measure the level of economic activity, they are all subject to distortion and measurement error and no single measure summarizes the whole economy adequately. Therefore, the coincident indicator combines a range of economic series together to give, in some sense, a better measure of overall economic activity. Leading economic indicators, as constructed by the CSO, are similar in that the component variables are thought to contain information on the future level of economic activity.

Despite the importance attached to coincident and other leading indicators in both the USA and the UK, the methodology of constructing such series has remained unchanged since the

CCC 0883–7252/96/020135–17

Received October 1992
Revised January 1995

early 1970s. The method adopted by the CSO in the UK, for example, is a rather mechanical approach, the econometric foundations for which are unclear. The first step taken is to detrend each series in the group using a five-year moving average process. Each detrended series is then rescaled so that all series in the group have the same average amplitude. The component series, thus adjusted, are averaged in a composite series with a weight of three on GDP and one for the other four variables. It should also be noted that the underlying series are seasonally adjusted before the steps described above are undertaken. Where the series is not already seasonally adjusted, the US Bureau of Census X-11 procedure (Lomas, 1983) is used before the steps described above.

In this paper we suggest that an alternative methodology, developed by Stock and Watson (1989b) for the construction of a coincident indicator, should be adopted by the CSO. The framework is a state space form in which a Kalman filter is used to construct an optimal estimate of an unobserved component which represents a common movement among the set of observable variables that are taken to measure economic activity. The unobservable term is then interpreted as corresponding to an underlying measure of economic activity which represents all the information contained in the comovements of the observable variables. The emphasis is on examining the time series properties of the variables used in the construction of the coincident index and taking the appropriate transformation or linear combination such that we can make a statistically valid inference concerning their comovement. The Stock and Watson framework provides an explicit statistical backdrop in which this can occur. In particular, we are concerned in our example with the case where there is a cointegrating vector between a subset of the observable component variables, which include both I(1) and I(0) variables. These are the same set of variables which the CSO currently uses to construct their coincident index, but it is unclear how they deal with the issue of trends and, in particular, the common stochastic trend which we identify among three of the five component variables that are I(1) processes.

In Section 2 of this paper we describe the Stock and Watson framework. Section 3.1 then establishes the time series properties of the CSO component series that we intend to use to construct an alternative coincident index. Section 3.2 applies the Stock and Watson technique to the problem of forming an indicator of underlying economic activity for the UK. We construct two alternatives for a new measure of economic activity. The first we call the levels form, which takes account of the cointegrating vector found to exist between three of the five component series. The second, the first difference indicator, deals with the differing order of stationarity of the component variables by first differencing the non-stationary I(1) variables and then constructs the index in the Stock and Watson framework. Section 4 concludes.

2. THE STOCK AND WATSON APPROACH

Stock and Watson (1989b) view the problem of identifying a coincident economic indicator as one of extracting a common component from a set of series. Let X_t denote an $n \times 1$ vector of economic series that are assumed to contain some information about the underlying performance of the economy, S_t. Defining an adapted version of Stock and Watson (1989b) the set of measurement equations take the form:

$$X_t = \gamma(L)S_t + u_t \tag{1}$$

where the same S variable effects each of the observed variables but with different weights which may be distributed across time as implied by the matrix $\gamma(L)$ and u_t is an $n \times 1$ vector of idiosyncratic terms that captures all movements in X_t not associated with S. In this model we assume that the variables X_t are driven by the single common factor S_t. The model is then completed by specifying the following two state equations:

$$\zeta(L)S_t = \delta + w_t \tag{2}$$

and[1]

$$\xi(L)u_t = q_t \tag{3}$$

where w_t and q_t are normally distributed with zero mean and given covariance matrix. Therefore S_t and u_t have a general distributed lag form where, when $\zeta(L)$ contains a unit root, the constant term δ can be thought of as a growth term allowing for the presence of a deterministic trend. This model will filter any common information from the X variables into S and all remaining variation in X will be relegated to the idiosyncratic effect. In the Stock and Watson (1989b) application $\gamma(L) = \gamma$ and the measurement equations are in first differences, hence the state variables are implicitly estimated as changes in the underlying series.[2] Stock and Watson's argument for using first differences is that all their X series are integrated (I(1)) series with no cointegrating vector between them. Under these conditions, we would expect to find n independent stochastic trends underlying the data, and reformulating the variables into first differences is appropriate. The complete model is then estimated in this form and the S variable is recovered.

Using first differences, however, leaves certain questions unanswered. How should the model be specified when some variables are in fact cointegrated or when some variables are stationary? If none of the variables are cointegrated, to what extent can we be confident that there is really an underlying series which is driving them all? In particular, if the n series are driven by n separate stochastic trends and thus have no long-run movement in common, the association of any emerging state variable with underlying economic activity may be tenuous.

The general issue of unit roots and common stochastic trends involve (1) unit roots in the S_t process, (2) unit roots in the u_t process and (3) factors of (1-L) in $\gamma(L)$. For example, if some components of X_t are I(1) processes that cointegrate, then S_t can be the stochastic trend common to these components.[3] In this instance S_t will be I(1) and therefore some rows of $\gamma(L)$ can contain (1-L), so that ΔS_t loads on the remaining I(0) components. For the case where the remaining variables are I(1) but do not form part of the cointegrating vector then a separate unit root may exist in u_t.

Identifying long-run movements among variables requires imposing the correct number of stochastic trends (see Robertson and Wickens, 1992, for a more complete description). When specifying a model it is important to impose the correct number of cointegrating vectors or stochastic trends; if the form of the model is unrestricted then the estimates will be consistent but not efficient, while if the levels terms are omitted or inappropriately restricted, the model will be misspecified and the estimates will no longer be consistent. However, if more than $n - r$ (r being the number of cointegrating vectors) stochastic trends are suppressed, the model will be misspecified.

[1] This format is not truly a state space one as the lag structure is more general. The move to a true state space form is, however, trivial (see Stock and Watson, 1991).
[2] Strictly this is not entirely accurate as in Stock and Watson (1989b) the monthly index produced does allow for the presence of lags in the employment series.
[3] For the case where there is only one stochastic trend.

3. CONSTRUCTION OF A NEW COINCIDENT INDEX FOR THE UK

The objective of this empirical application is to construct an alternative composite coincident index to that of the CSO using the technique of Stock and Watson (1989b) described in the previous section.

3.1. Time Series Properties of the CSO Component Variables

For a direct comparison the vector of indicator variables is the same as those used by the CSO and comprise the following: Gross Domestic Product average income at factor cost 1985 = 100 (GDP), Output of the production industries 1985 = 100 (IP), Confederation of British Industry (CBI) quarterly survey of below-capacity utilization in percentage terms (CBICU), the volume of Retail Sales 1985 = 100 (RES), CBI quarterly survey of expected change in stocks of raw materials in percentage balance terms (CBIRM).

Casual inspection of the five component variables used by the CSO (see Figures 1–3) would suggest two distinct categories: first, a trended, non-stationary group *GDP*, *IP* and *RES* (in all that follows natural logs are used for these variables). Second, the CBI variables *CBICU* and *CBIRM*, which have no obvious trend and are possibly a stationary process with large variance. In the first group each variable clearly has a strong upward trend component and comovements between them are clear. However, the idiosyncratic component of each series looks significant such that additional information is gained. For example, *IP* looks more volatile than both *GDP* and *RES* and different growth rates in economic activity would be implied from each series, especially in the period from the early 1970s through to 1980. The second group shows high

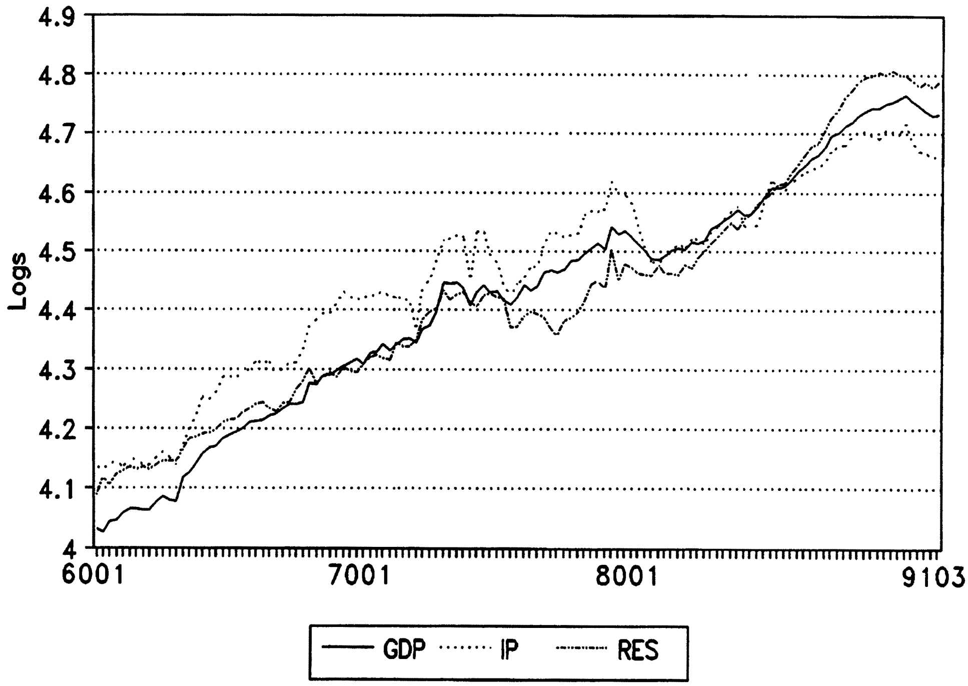

Figure 1. CSO trended component variables

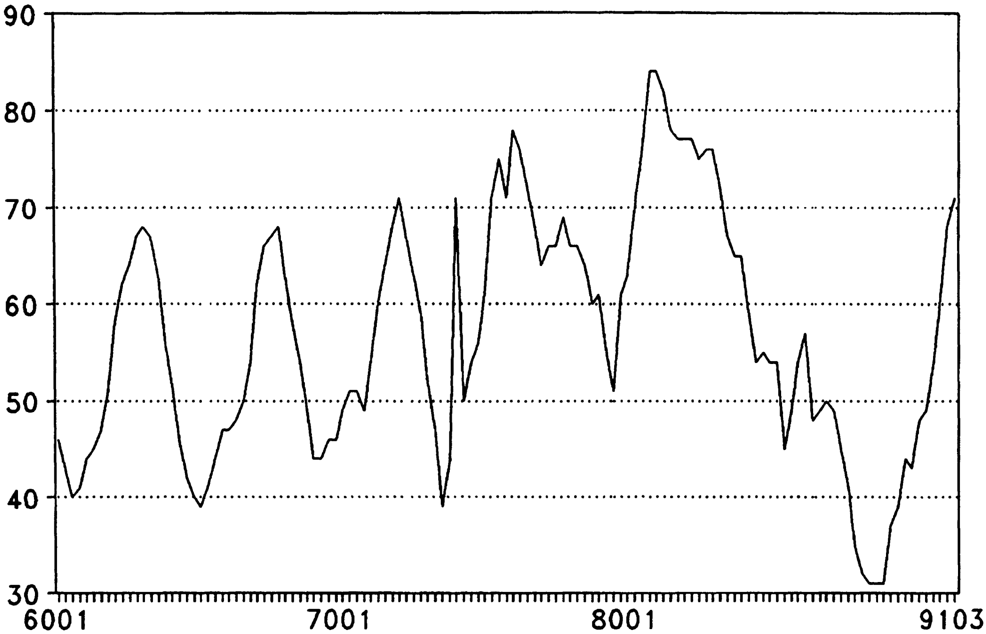

Figure 2. CSO stationary component variable. *CBICU*—percentage below-capacity utilization

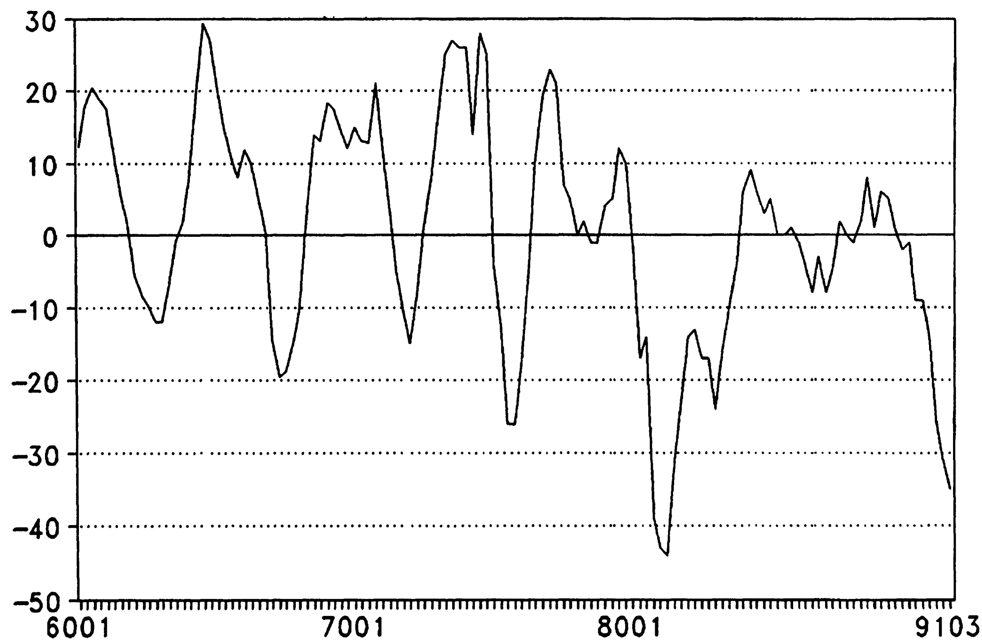

Figure 3. CSO stationary component variable. *CBIRM*—expected percentage change in raw materials

volatility, in particular the expectations variable *CBIRM*, and while comovements are apparent sometimes the general pattern is less clearly defined.

So that we may more formally establish what type of trend (if any) best characterizes each variable, we apply the ADF testing procedure for both deterministic trend and a unit root using the following form:[4]

$$\Delta Y_t = c + \alpha Y_{t-1} + \beta t + \sum_{i=1}^{q} \gamma_i \Delta Y_{t-i} + \nu_t$$

where Y is any of our variables, t represents a time trend and q is the number of lagged dependent variables required to induce no serial correlation in the residuals υ_t. Table I reports the Dickey–Fuller (DF) statistics, or the Augmented Dickey–Fuller (ADF) where appropriate for the coefficient α and the standard t-ratio for the coefficient β as tests for unit roots and deterministic trends respectively. For all the five variables the deterministic time trend is insignificant.

In the same calculation the value of the DF or ADF statistic strongly suggests the presence of a stochastic trend in the three variables *GDP*, *RES* and *IP*. For the case of the CBI variables the ADF statistic suggests no stochastic trend for the *CBIRM* variable but is borderline on the presence of a stochastic trend for the *CBICU* variable. We therefore report an additional ADF(1) statistic for the CBI variables, where only a constant is included in the regression. From this we conclude that no stochastic trend is present in either, contradicting the above result, as both are below the critical value of $-2{\cdot}95$. Therefore, we decide to model the variables *GDP*, *IP* and *RES* as I(1) and *CBICU* and *CBIRM* as I(0).[5]

Before proceeding we need to test the single common factor assumption. This entails establishing the number of common stochastic trends contained in our information set of five variables. In this instance we are interested in the three I(1) variables as the I(0) variables will not effect the number of common stochastic trends. This will simplify the Johansen maximum likelihood estimation. We conduct two tests for a single common stochastic trend. First, we test for cointegration among our three I(1) variables, where the number of stochastic trends is $n-r$, where $n=3$ and r is the number of cointegrating vectors. Hence in our case we look for two

Table I. Testing for deterministic and stochastic trends

	GDP	*IP*	*RES*	*CBICU*	*CBIRM*
DF	−2·11	−2·27	−1·84	—	—
t	1·88	1·84	1·76	—	—
ADF(1)	—	—	—	−2·65	−4·36
ADF(1)				−2·95	−4·06
t	—	—	—	0·22	−1·8

Critical values (100 observations) for the DF and the ADF(1) is $-3{\cdot}45$, both a time trend and a constant are assumed to be present. The second ADF(1) statistic is computed for the case where only a constant is included in the regression and therefore the critical value for 100 observations is $-2{\cdot}89$.

[4] In this regression the t-statistic for H_0: $\beta=0$ is not $\sim N(0,1)$ when $\alpha=1$. The critical values can be obtained using simulations constructed from a DGP with $\alpha=1$ and $\beta=0$.

[5] If in fact, contrary to the implication of the tests carried out here, a deterministic trend is present in any of the variables then we allow for this by including a constant in the state equation.

cointegrating vectors among our three variables. The second test is suggested by Geweke (1977) which tests the single-factor restriction at a range of frequencies.

Table II(a) reports the results of the cointegration analysis. In the first instance we test for cointegrating vectors between *GDP* and *RES* respectively. A cointegrating relationship in both cases is a necessary condition for two cointegrating vectors between all three variables. The results in Table II(a) suggest that both the variables *RES* and *IP* are cointegrated with *GDP* and therefore potential for a single common stochastic trend between the three variables exists. However, when combining all three variables in the analysis, we find only one clear cointegrating vector. The test for two cointegrating vectors is very close to the critical value (18·64 versus 19·96), suggesting the possibility of a second cointegrating term, but, nonetheless the results imply the presence of a second common trend. Hence we report the second of our tests for a single common stochastic trend in Table II(b).

Table II(b) reports the Geweke (1977) test for a single common factor restriction at a range of frequencies. Our assertion of two cointegrating vectors and therefore a single common trend (all at zero frequency), although not directly tested, is consistent with the results. For all frequencies examined, $0{\cdot}27\pi$ through to and including $0{\cdot}87\pi$, the single common factor restriction is accepted. This, in combination with the marginal result reported in Table II(a), we take this to imply a common factor at zero frequency. As most of the evidence favours the single common factor we proceed using this assumption.

Table II(a). Johansen normalized cointegrating vectors: dependent variable: *GDP*

RES	0·82	—	0·62
IP	—	1·67	0·41
$r=0$	35·53	37·99	50·32
1	5·49	8·71	18·64
2	—	—	4·02

Critical values for the likelihood ratio tests for the case of two variables are 19·96, 9·24 and for three variables 34·91, 19·96 and 9·24. Note that a constant is included in the error-correction term. The cointegration tests are based on maximal eigenvalues of stochastic matrix.

Table II(b). Geweke test for the single-factor restriction at a range of frequencies

Frequency	Ordinates (of 131)	Test statistic ($\chi^2(11)$)
$0{\cdot}12\pi$	3–31	7·7
$0{\cdot}37\pi$	35–63	10·7
$0{\cdot}62\pi$	67–95	10·7
$0{\cdot}88\pi$	99–127	7·7
Overall		36·9($\chi^2(44)$)

The test was constructed using all five variables, where first differences of the I(1) variable and levels of the I(0) were used, $\pi=(T/2+1)$ where $T=128$, which allowed for dropping two observations because of pre-whitening the data. The critical value of chi-squared 11 is 19·68.

3.2. Constructing a Measure of Economic Activity

In this section we report two alternative measures of economic activity each of which results from applications of the framework outlined in Section 2, but deals with the time series properties of the component variables in different ways. The first keeps the component variables in there levels form but allows a single common trend component S_t to exist between the cointegrated I(1) variables. The same common trend, plus its lagged value, is estimated for the stationary variables. The common trend represents the measure of economic activity and in this instance is modelled as a random walk with an estimated drift term. The second approach deals with the differing orders of integration by first differencing the I(1) variables, ignoring the long-run relationship and then combining a set of stationary component variables. Here the common trend component, as before, is the unobserved component but is modelled as a stationary AR(1).

The appropriate comparison or point of reference for these measures of economic activity is not obvious given that the notion of underlying economic activity is, by definition, unobservable. Therefore there is no direct comparison with an actual series, as implicit in an exercise of this type is the assumption that GDP taken in isolation is not a sufficient measure of economic activity. Making a comparison with the CSO coincident index is also difficult, as a main point of the paper is to highlight that the method currently used by the CSO incorrectly deals with variables with different orders of stationarity. The comparison in the USA is made a little easier as additional points of reference exist in the *ex-post* NBER-dated cyclical turning points. However, in the UK the closest comparative measure is the CSO coincident index. Therefore while acknowledging the limitations, our only choice is to use GDP and the CSO coincident indicators as a reference, the justification being that, although imperfect, these measures do provide a guide to the likely path of economic activity. Hence if the new measure were completely different it is likely to be implausible or would require strong justification. We adopt a different comparison in each of the two cases. For the levels index our comparison is with GDP in both the levels and percentage change, plus reference to the other two I(1) variables *RES* and *IP*. In the case of the difference index, the more appropriate comparison is with the CSO coincident index and its percentage change.

In the estimation we use a likelihood function that concentrates out a single diagonal element of the states covariance matrix (see Harvey, 1989, p. 433). The algorithm employed is Davidson–Fletcher–Powell, where the first and second derivatives were evaluated numerically. Throughout we report the non-smoothed common component from the Kalman filter. Therefore it is an estimate of $E[S_t \mid X_t]$ or in the case of the second index $E[DS_t \mid X_t]$.[6]

First, we examine the levels coincident index (S) plotted in Figure 4, where the maximum likelihood estimates are reported in Table III(a). The common factor, S_t, enters each of the five equations, contemporaneously in the case of the I(1) variables equation but also with a lag in the case of the I(0) CBI variables. The significance of the common factor effect is strong in all cases, particularly with retail sales, where the coefficient is 0·941 with a t-statistic of 22·57. The model shows no sign of misspecification, each equations error term being modelled as an AR(2) process, and the residuals show no serial correlation.[7]

[6] Confidence intervals for S_t can be computed using the covariance matrix of the Kalman filter. However, we do not compute these as in this instance it is not clear what the null hypothesis would be.

[7] The choice of an AR(2) process was based on preliminary OLS estimation where for the I(1) variables we regressed the variables of interest on *GDP* (for *GDP* we used its the lagged value) and for the I(0) variables the first difference of *GDP*. We then examined the AR properties of the resulting residuals. Parsimony also played a part, as the estimation problem is already large and hence increasing the numbers of parameters to be estimated would make any inference more difficult.

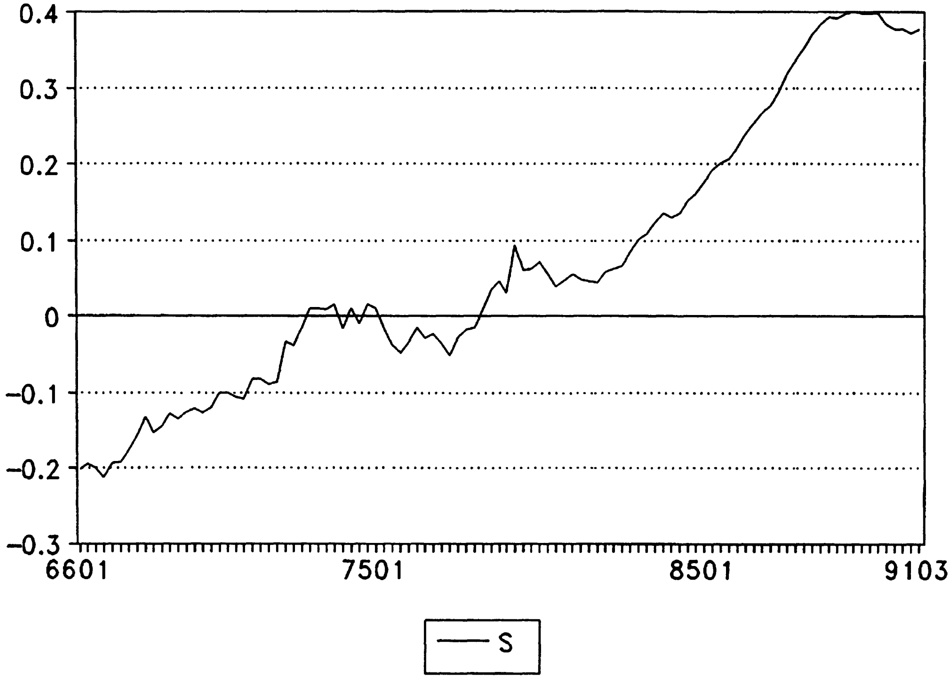

Figure 4. Levels coincident index

The level index is quantitatively very similar to our three I(1) variables, with a stronger trend growth from the late 1980s onwards (see Figure 5) and with coinciding turning points. In Figure 5 the series S is neither as volatile as industrial production nor as smooth as GDP and most closely corresponds to the series *RES*, whose estimated coefficient, as noted above, is large. Figure 6 plots the annual percentage change of *GDP* and our levels indicator (S). There is clear comovement between the two series (correlation coefficient between the two series is 0·76) and the implied percentage changes are of the same order of magnitude. For the period 1967 : Q1 through to around 1976 the percentage changes in economic activity and the timing are very similar. However, the index S implies a more pronounced fall in economic activity in 1976 with a more volatile and ultimately higher percentage increase around 1979. The 1980–81 recession and subsequent recovery are quantitatively similar throughout the 1980s, with the stronger trend growth of the index S being more apparent in the late 1980s.

According to the levels index S, the start of the most recent recession occurs at approximately the same time as GDP began to fall in around 1989. Generally, using GDP as a comparison, our alternative measure of economic activity implies important differences at specific points, both in terms of the timing and the size of changes in economic activity. However, while these differences generate interest in the alternative measure the comparisons are sufficiently close and the orders of magnitude of the differences maintain the plausibility.

The second of our indices, the first difference index (DS), maximum likelihood estimates are reported in Table III(b) and is plotted in Figure 7. The common factor, which now has an estimated coefficient on its lagged value of 0·95, enters all five equations only contemporaneously. The common factor term DS is strongly significant in the equations for the I(0) CBI variables but insignificant with small coefficients (zero in the case of *RES*) for the

Table III(a). Maximum likelihood estimates of the model (1) to (3) using CSO coincident indicator component variables: levels

(A) *Measurement equations*:

$GDP = S_t + u_{1t}$ (normalizing restriction imposed at one)

$IP = 0{\cdot}643S_t + u_{2t}$
(16·71)

$RES = 0{\cdot}941S_t + u_{3t}$
(22·57)

$CBICU = -111{\cdot}0S_t + 111{\cdot}0S_{t-1} + u_{4t}$

$CBIRM = -55{\cdot}0S_t + 55{\cdot}0S_{t-1} + u_{5t}$

Likelihood ratio test of restriction $\chi^2(2) = 4{\cdot}8$

(B) *State equations*:

$S_t = 0{\cdot}0057 + 1{\cdot}0S_{t-1} + w_t$; $\sigma_w^2 = 0{\cdot}0012$
(1·7)

$u_{1t} = -0{\cdot}815u_{1t-1} - 0{\cdot}682u_{1t-2} + q_{1t}$; $\sigma_{q1}^2 = 0{\cdot}0012$
(5·98) (5·23)

$u_{2t} = -2{\cdot}516u_{2t-1} - 1{\cdot}229u_{2t-2} + q_{2t}$; $\sigma_{q2}^2 = 0{\cdot}00038$
(11·7) (5·49)

$u_{3t} = -0{\cdot}186u_{3t-1} - 0{\cdot}918u_{3t-2} + q_{3t}$; $\sigma_{q3}^2 = 0{\cdot}00038$
(0·80) (2·62)

$u_{4t} = -1{\cdot}49u_{4t-1} - 1{\cdot}11u_{4t-2} + q_{4t}$; $\sigma_{q4}^2 = 0{\cdot}171$
(13·53) (7·41)

$u_{5t} = -1{\cdot}21u_{5t-1} - 0{\cdot}723u_{5t-2} + q_{5t}$; $\sigma_{q5}^2 = 0{\cdot}171$
(25·99) (15·88)

(C) *Innovations from the Kalman filter*:

Autocorrelation function (Box–Pearce statistic in parentheses):

	u_{1t}	u_{2t}	u_{3t}
Lag: 1	−0·093(0·01)	−0·0296(0·09)	0·004(0·00)
4	−0·053(2·04)	−0·047(4·08)	0·039(3·33)
8	−0·029(6·75)	0·059(9·56)	0·245(14·15)

	u_{4t}	u_{5t}
1	0·047(0·23)	0·217(5·10)
4	−0·131(2·44)	−0·152(10·24)
8	0·166(5·83)	0·159(16·57)

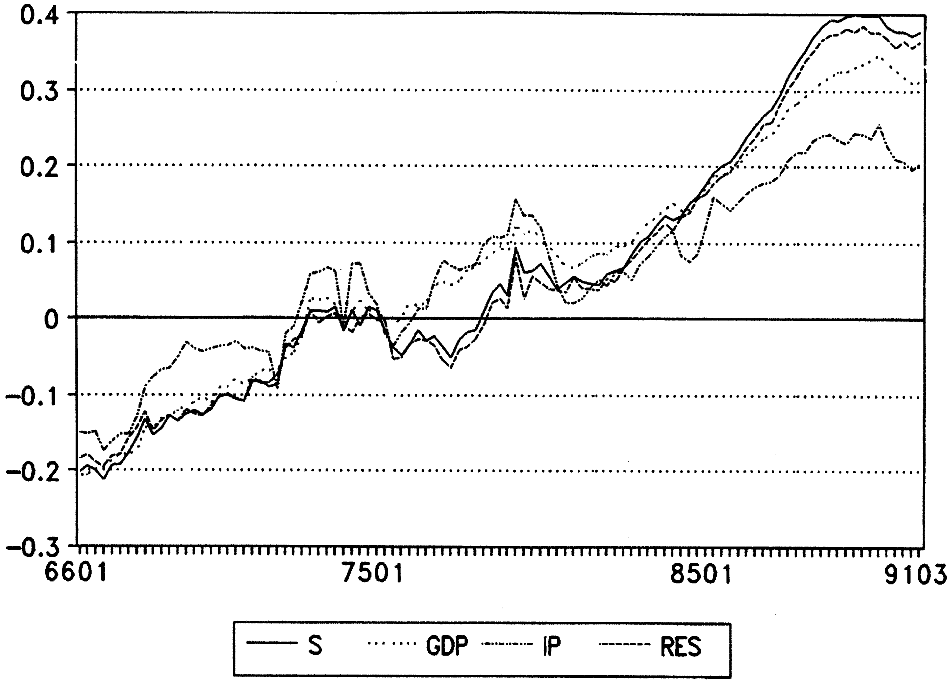

Figure 5. Levels coincident index

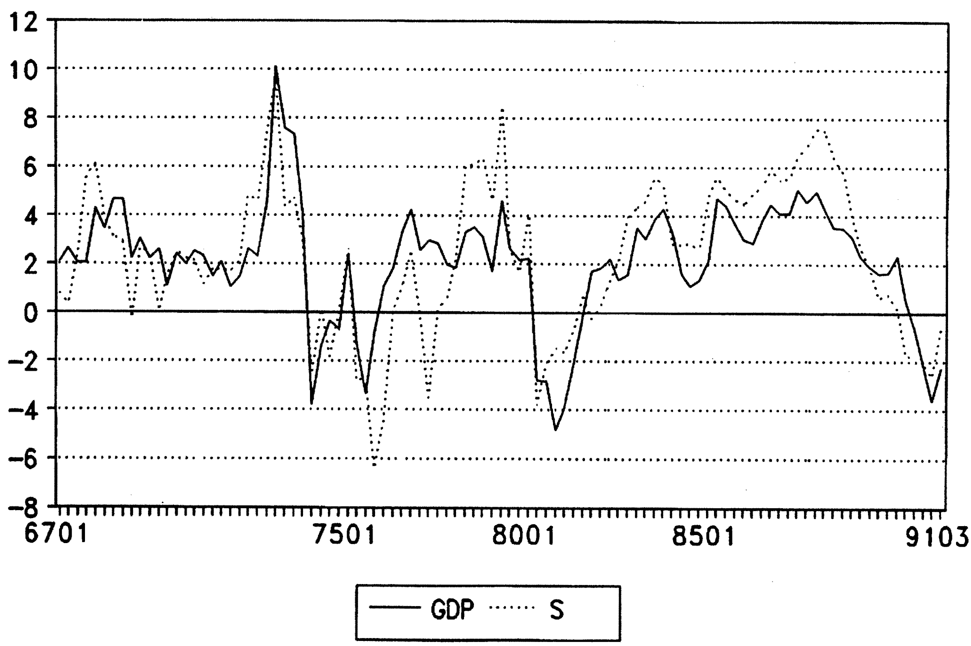

Figure 6. GDP and level coincident index. Annual percentage change

Table III(b). Maximum likelihood estimates of the model (1) to (3) using CSO coincident indicator component variables: first differences

(A) *Measurement equations*:

$\Delta GDP = DS_t + u_{1t}$ (normalizing restriction imposed at one)

$\Delta IP = \underset{(0\cdot39)}{0\cdot053}DS_t + u_{2t}$

$\Delta RES = \underset{(0\cdot001)}{0\cdot0}DS_t + u_{3t}$

$CBICU = \underset{(15\cdot31)}{-227\cdot9}DS_t + u_{4t}$

$CBIRM = \underset{(9\cdot42)}{114\cdot44}DS_t + u_{5t}$

(B) *State equations*:

$DS_t = \underset{(-0\cdot002)}{-0\cdot00015} + \underset{(8\cdot47)}{0\cdot949}DS_{t-1} + w_t;\ \sigma_w^2 = 0\cdot248$

$u_{1t} = \underset{(0\cdot05)}{-0\cdot095}u_{1t-1} - \underset{(0\cdot12)}{0\cdot173}u_{1t-2} + q_{1t};\ \sigma_{q1}^2 = 0\cdot248$

$u_{2t} = \underset{(0\cdot39)}{-0\cdot053}u_{2t-1} - \underset{(0\cdot21)}{0\cdot03}u_{2t-2} + q_{2t};\ \sigma_{q2}^2 = 0\cdot303$

$u_{3t} = \underset{(0\cdot39)}{-0\cdot038}u_{3t-1} - \underset{(1\cdot11)}{0\cdot15}u_{3t-2} + q_{3t};\ \sigma_{q3}^2 = 0\cdot016$

$u_{4t} = \underset{(0\cdot056)}{-0\cdot085}u_{4t-1} - \underset{(1\cdot02)}{1\cdot64}u_{4t-2} + q_{4t};\ \sigma_{q4}^2 = 0\cdot527$

$u_{5t} = \underset{(10\cdot04)}{5\cdot5}u_{5t-1} - \underset{(2\cdot55)}{2\cdot11}u_{5t-2} - \underset{(5\cdot11)}{1\cdot94}u_{5t-3} + q_{5t};\ \sigma_{q5}^2 = 0\cdot0104$

(C) *Innovations from the Kalman filter*:

Autocorrelation function (Box–Pearce statistic in parentheses):

	u_{1t}	u_{2t}	u_{3t}
Lag: 1	0·098(0·04)	−0·119(1·54)	−0·069(0·51)
4	0·014(1·93)	0·018(3·99)	−0·006(3·43)
8	−0·027(2·94)	−0·103(11·63)	0·018(4·37)

	u_{4t}	u_{5t}
1	−0·052(0·29)	0·063(0·43)
4	−0·145(4·21)	−0·098(18·59)
8	−0·022(8·48)	0·368(43·20)

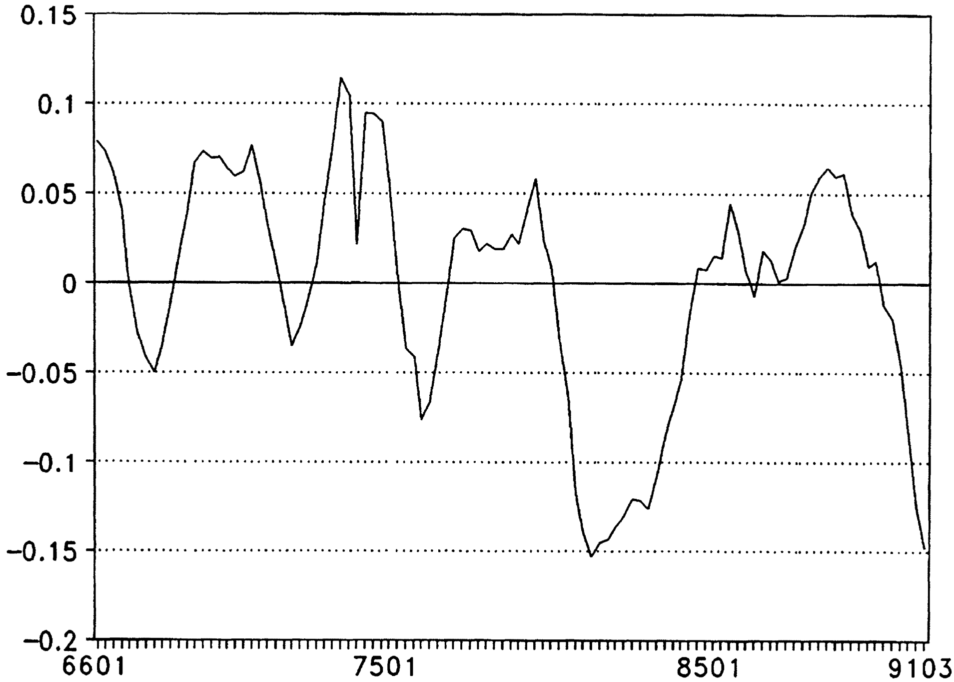

Figure 7. First difference coincident index

group of variable which are I(0) in first differences. Therefore the resulting index DS reflects movements in the CBI variables and *GDP* whose coefficient is imposed to be one. The error terms are again AR(2) processes, except the *CBIRM* variables which has an AR(3) form. Panel (C) in Table III(b) reports the autocorrelation function that shows no serial correlation in the Kalman filter innovations. In comparison with the levels form the general specification fits less well as there the common trend component was strongly significant when explaining all the five variables. In addition, the variances of the error terms in Table III(a) are considerably smaller.

The index *DS* and the relevant comparisons are reported in Figures 7–9. In this instance we compare the index to the CSO coincident index as these two indices are best viewed as attempts at measuring the turning points in economic activity rather than the level. Constructing a levels form from what is a change term as in Stock and Watson would be a possibility, but we view the index *S* as being the appropriate notion of the levels index and therefore only in Figure 10 do we compare the annual growth rates of the two indices *S* and *DS*. The *DS* index plotted in Figure 7 reveals a broadly similar pattern and timing of the movement in economic activity to that implied by the CSO index plotted in Figure 8. However, the *DS* index highlights two periods in the early and late 1970s where economic growth is significantly lower than the rate implied by the CSO index. A further difference in the implied rates of growth in economic activity from the two series arise for the early 1980s to the 1985 period, where the *DS* index suggests that economic growth was much greater. A more revealing comparison would be the annual percentage changes of the two series plotted in Figure 9. The actual size of the implied change in economic activity is of a similar order of magnitude and the precise timing of the changes does differ, but only by odd quarters in a majority of cases. The correlation coefficient between the two series is 0·72. The main differences occur in the early 1970s (around 1973)

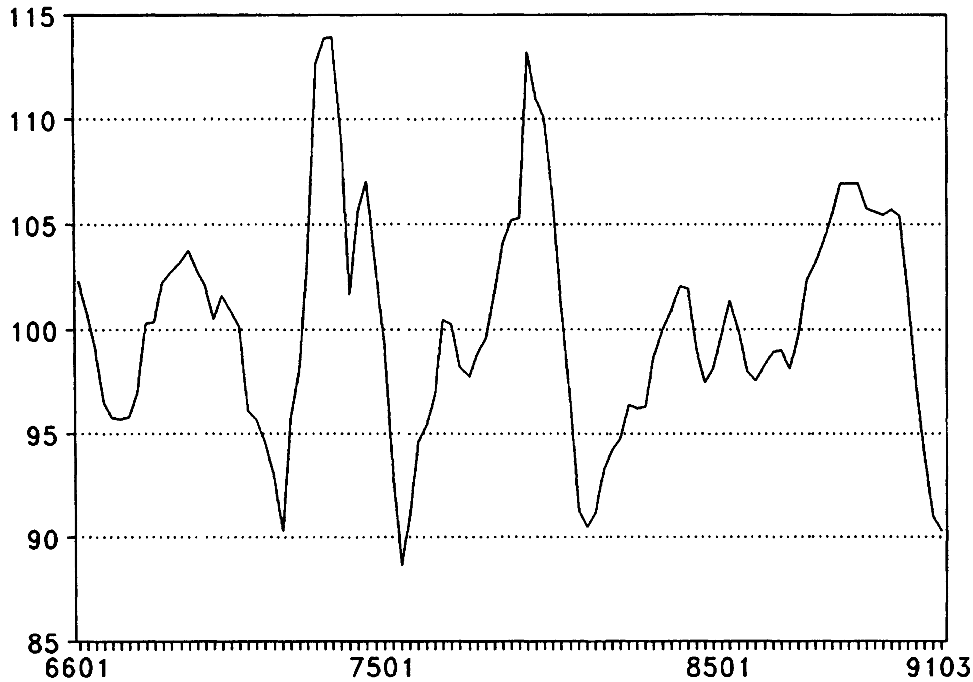

Figure 8. CSO coincident index

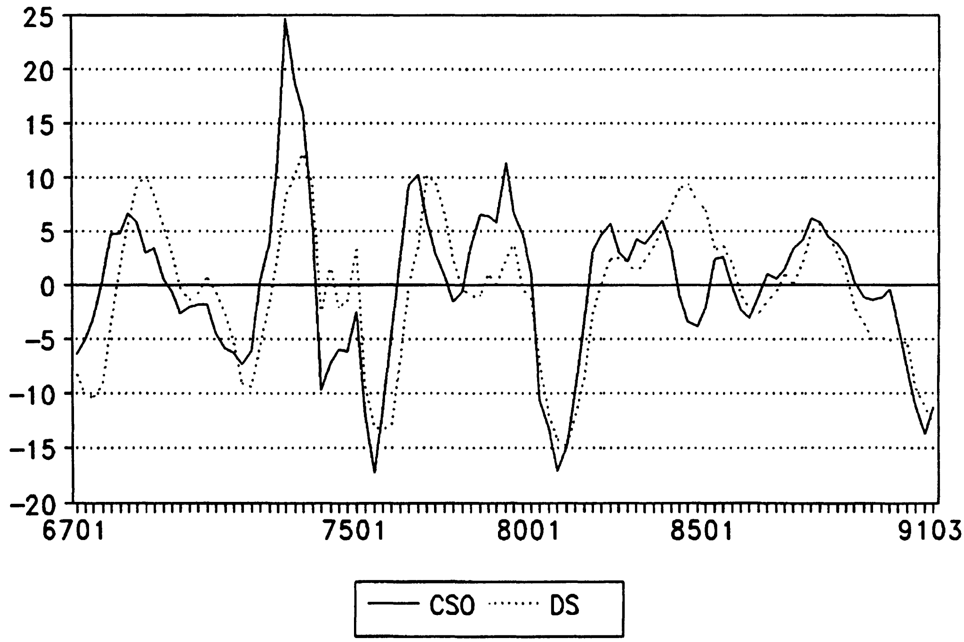

Figure 9. CSO and first difference coincident index. Annual percentage change

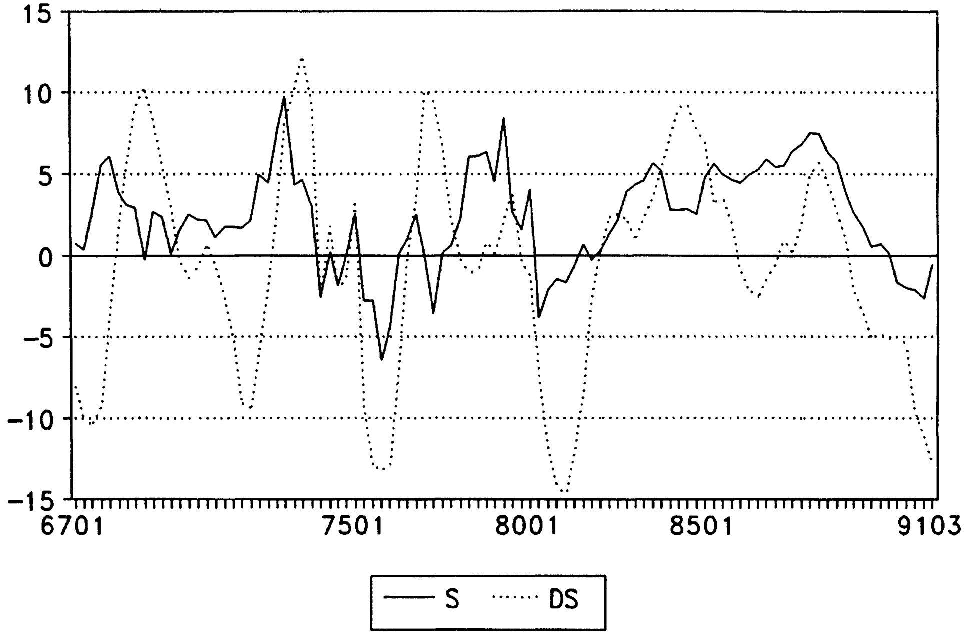

Figure 10. Level and first difference coincident index. Annual percentage change

where the CSO index implies a much greater growth in economic activity (25% versus 12%) at a slightly earlier point, followed by a much larger fall. This is the one point where the levels index (S) and GDP imply virtually the same percentage change of 10%, lending a degree of support to the DS index. The second major difference occurs around 1985, where the DS index implies a growth of close to 10% while the CSO index implies a negative growth of close to 5%.[8] Referring back to Figure 6, we note that during the 1985 period both GDP and our level index S imply falls in the growth of economic activity, therefore in this instance lending support to the inference implied by the CSO index. The inference of both indices would be approximately the same after the 1985 period as the two indices move closely together, suggesting the beginning of the most recent UK recession at approximately the same time, the beginning of 1990.

Figure 10 plots the growth rates of the two alternative indices S and DS. The DS index shows large and volatile movements in economic growth throughout the period while the S index shows a movement and volatility of a smaller order of magnitude. This is a direct result of the importance of the CBI variables in this formulation. However, the important difference between the indices is the implied timing and direction of changes in economic activity. There are three instances where the two indices imply opposite movements in economic activity, 1969–70, 1977 and 1985. In addition, the timing of changes in economic activity do not coincide in the periods 1973–4 nor in 1980. Therefore it is important which of the two indices we prefer, as this will imply significant differences in the movement of economic activity.

The first difference index, DS, we interpret as similar to the practice of the CSO in the sense that the trends of the component series are removed before the construction of the coincident

[8] An explanation for this would lie in the movement of the variable *CBICU*, which falls by a large amount during this period.

index. The index *DS* is also comparable to the index of Stock and Watson who first difference a set of I(1) variables that are not cointegrated. In Stock and Watson this was the correct procedure as no cointegrating vector was found to exist between their variables but for the CSO we believe that this is not the case. In their case detrending (in ours first differencing) the variables before the construction of the coincident index removes a significant part of the information; that three of the five variables are I(1) and have a common stochastic trend. Therefore the long-run relationship is ignored and a misspecification is induced. This is highlighted by comparing the results in Tables III(a) and III(b), which shows that the levels form does have a superior fit to the first difference form. The implication is that the information is in the levels of variables rather than in the noise around a detrended variable. Therefore our preferred specification is in levels as it considers the long-run cointegrating relationship, which is reflected in goodness of fit.

4. CONCLUSION

In this paper the methodology of Stock and Watson (1989b) is suggested as a preferred method of constructing a coincident index to the procedure currently employed by the CSO. It provides a framework in which the time series properties of the component variables used in constructing the index and the relationships between them may be dealt with in an explicit way. In our example, where we construct two new coincident indices for the UK, we allow for the different orders of integration in the component variables and for the fact that three of the five component variables form a cointegrating vector. The first is in a levels form where we allow for the cointegrating vector between the three I(1) variables, the second index is constructed after stationarity has been imposed on I(1) variables by first differencing. The two indexes are then assessed by comparing them with the annual percentage change in GDP and the CSO respectively and are shown to have different implications for economic activity. We prefer the levels formulation, as it allows for the long-run relationship between the three I(1) variables and has a superior fit in the common factor formulation.

ACKNOWLEDGEMENTS

This paper was written while the first author worked at the London Business School. The views expressed in this paper are solely those of the authors and do not represent those of the London Business School or the Bank of England. We would like to thank the anonymous referees of this journal for a number of useful comments and suggestions and in particular we would like to thank Mark Watson for his help and encouragement.

REFERENCES

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

Cochrane, J. H. (1991), 'A critique of the application of unit root tests', *Journal of Economic Dynamics and Control*, **15**, No. 2, 275–84.

Crowder, M. J. (1976), 'Maximum likelihood estimation for dependent observations', *Journal of the Royal Statistical Society*, Series B, **38**, No. 1, 45–53.

Economic Trends No. 257, March (1975), 'Cyclical indicators for the United Kingdom economy', 95–8.

Engle, R. F. and C. W. J. Granger (1987), 'Co-integration and error correction representation, estimation and testing', *Econometrica*, **35**, 251–76.

Engle, R. F. and S. Yoo (1991), 'Cointegrated economic time series: an overview with new results', in R. F. Engle and C. W. J. Granger (eds), *Long Run Economic Relationships: Readings in Cointegration*, Oxford University Press, New York.

Geweke, J. (1977), 'The dynamic factor analysis of economic time series', In D. J. Aigner and A. S. Goldberger (eds), *Latent Variables in Socio-Economic Models*, North-Holland, Amsterdam, Chapter 19.

Hall, S. G., K. Cuthbertson and M. P. Taylor (1992), *Applied Econometric Techniques*, University of Michigan Press, Ann Arbor.

Harvey, A. C. (1987), 'Applications of the Kalman filter in econometrics', in T. F. Bewley (ed.), *Advances in Econometrics: Fifth World Congress*, Vol. 1, Econometric Society Monograph No. 13, Cambridge University Press, Cambridge.

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

Johansen, S. (1988), 'Statistical analysis of cointegration vectors', *Journal of Economic Dynamics and Control*, **12**, 231–54.

Lomas, D. (1983), *Economic Trends*, November.

Mitchell, W. C. and A. F. Burns (1938), *Statistical Indicators of Cyclical Revivals*, NBER, New York.

Robertson, D. and M. R. Wickens (1992), 'Stochastic trends representation, cointegration and economic activity', mimeo.

Sargent, T. J. and C. A. Sims (1977), 'Business cycle modeling without pretending to have too much *a-priori* economic theory', in C. Sims *et al.* (eds), *New Methods in Business Cycle Research*, Federal Reserve Bank of Minneapolis, Minneapolis.

Stock, J. H. and M. W. Watson (1988), 'A new approach to the leading economic indicators', manuscript, Kennedy School of Government, Harvard University.

Stock, J. H. and M. W. Watson (1989a), 'A simple MLE of cointegrating vectors in higher order integrated systems', NBER Technical Working Paper No. 83.

Stock, J. H. and M. W. Watson (1989b), 'New indexes of leading and coincident economic indicators', *NBER Macroeconomics Annual*, 351–94.

Stock, J. H. and M. W. Watson (1991), 'A probability model of coincident economic indicators', in K. Lahiri and G. H. Moore (eds), *Leading Economic Indicators: New Approaches and Forecasting Records*, Cambridge University Press, New York, pp. 63–85.

Schweppe, F. (1965), 'Evaluation of likelihood functions for Gaussian signals', *IEEE Transactions on Information Theory*, **11**, 61–70.

References

Co-Integration and Error Correction: Representation, Estimation, and Testing
Robert F. Engle; C. W. J. Granger
Econometrica, Vol. 55, No. 2. (Mar., 1987), pp. 251-276.
Stable URL:
http://links.jstor.org/sici?sici=0012-9682%28198703%2955%3A2%3C251%3ACAECRE%3E2.0.CO%3B2-T