JOURNAL OF APPLIED ECONOMETRICS
J. Appl. Econ. **18**: 427–443 (2003)
Published online 22 October 2002 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/jae.695

A NEW COINCIDENT INDEX OF BUSINESS CYCLES BASED ON MONTHLY AND QUARTERLY SERIES

ROBERTO S. MARIANO[a] AND YASUTOMO MURASAWA[b]*

[a] *Department of Economics, University of Pennsylvania, USA and School of Economics and Social Sciences, Singapore Management University, Singapore*
[b] *College of Economics, Osaka Prefecture University, 1-1 Gakuen-cho, Sakai, Osaka 599-8531, Japan*

SUMMARY

Popular monthly coincident indices of business cycles, e.g. the composite index and the Stock–Watson coincident index, have two shortcomings. First, they ignore information contained in quarterly indicators such as real GDP. Second, they lack economic interpretation; hence the heights of peaks and the depths of troughs depend on the choice of an index. This paper extends the Stock–Watson coincident index by applying maximum likelihood factor analysis to a mixed-frequency series of quarterly real GDP and monthly coincident business cycle indicators. The resulting index is related to latent monthly real GDP. Copyright © 2002 John Wiley & Sons, Ltd.

1. INTRODUCTION

There is no doubt that, as a measure of the aggregate state of an economy, real GDP is one of the most important coincident business cycle indicators. Popular US monthly coincident indices of business cycles, however, do not use real GDP; e.g. the composite index (CI) of coincident indicators, currently released by The Conference Board, and the Stock–Watson Experimental Coincident Index (XCI) developed by Stock and Watson (1989). This is presumably because real GDP is quarterly. Without a statistically rigorous method to construct a monthly index from monthly and quarterly series, they ignore quarterly indicators. The Japanese coincident CI uses a quarterly indicator (operating profits), but they simply transform it into a monthly series by linear interpolation.

Another problem of the standard coincident indices is that they lack economic interpretation. Figure 1 compares the CI and the XCI from 1979 to 1983, during which there are two peaks and two troughs. The XCI indicates that the trough in November 1982 is deeper than that in July 1980, while the CI indicates that the two are almost the same depth. In fact, real GDP (seasonally adjusted) is *higher* in the fourth quarter of 1982 (4,915.6 billion chained 1996 dollars) than in the third quarter of 1980 (4,850.3 billion chained 1996 dollars). Such inconsistency can arise because the levels of these indices have no economic interpretation.

This paper proposes a new coincident index of business cycles that uses both monthly and quarterly indicators. Stock and Watson (1991) construct a coincident index (hereafter the S–W coincident index) by applying maximum likelihood (ML) factor analysis to the four monthly coincident indicators that currently make up the coincident CI. This paper extends the S–W coincident index by including quarterly real GDP. The resulting index should improve upon the

* Correspondence to: Yasutomo Murasawa, College of Economics, Osaka Prefecture University, 1-1 Gakuen-cho, Sakai, Osaka 599-8531, Japan. E-mail: murasawa@eco.osakafu-u.ac.jp

Received 22 April 2002
Revised 20 June 2002

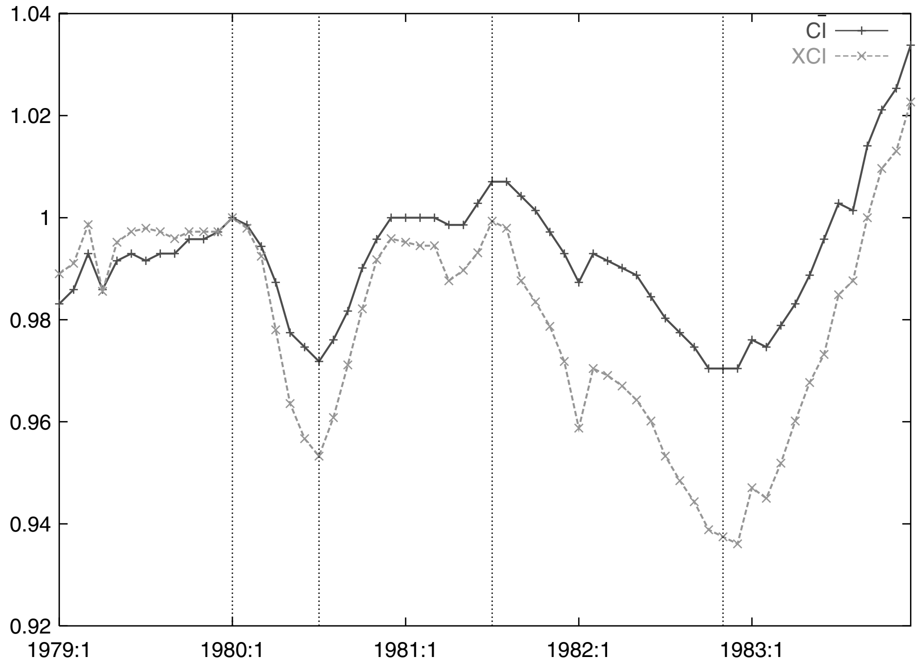

Figure 1. The Conference Board CI and the Stock–Watson XCI from 1979 to 1983 (1980 : 1=1). The vertical lines are the NBER business cycle reference dates. *Sources*: The Conference Board and the home page of James Stock

S–W coincident index, because it uses the most important coincident indicator that the S–W coincident index ignores, namely, real GDP. The resulting index is also related to latent monthly real GDP.

Technically, this paper discusses ML factor analysis of time series with mixed frequencies, i.e. monthly and quarterly. Consider a state-space representation of a factor model, treating quarterly series as monthly series with missing observations. Following Brockwell and Davis (1991, sec. 12.3) and Brockwell, Davis, and Salehi (1991), we fill in missing observations with iid draws from the standard normal distribution independent of the model parameters and rewrite the state-space model accordingly, so that we can apply the Kalman filter to evaluate the likelihood function. Realizations of the draws actually do not matter, because we rewrite the state-space model in that way; hence we do not draw but simply put zeros in practice. Numerical maximization of the likelihood function is straightforward. We use a quasi-Newton method, while Shumway and Stoffer (1982) use an EM algorithm; see also Shumway and Stoffer (2000, sec. 4.4).

The S–W coincident index is essentially the updated estimate of the common factor in the coincident indicators. We prefer the *smoothed* estimate instead, not only because it uses more information, but also because it slightly simplifies formulation of our state-space model. Let y_t be a vector of the indicators (differences in logs) and f_t be the common factor in y_t. Let for $t \geq 1$, $\mathbf{y}_t := (y_1, \dots, y_t)$. Then the S–W coincident index is the updated estimate of the *cumulative* common factor. (One may further take the exponential, but it does not affect the turning points.)

Notice that for $t > 1$,

$$\mathrm{E}\left(\sum_{j=1}^{t} f_j \middle| \mathbf{y}_t\right) \neq \sum_{j=1}^{t} \mathrm{E}(f_j|\mathbf{y}_j)$$

i.e. the updated estimate of the cumulative common factor differs from the cumulative sum of the updated estimates of the common factor. To obtain the left-hand side, Stock and Watson (1991) include the cumulative common factor in the state vector. Among recent extensions of the S–W coincident index that introduce Markov regime-switching into the common factor, Kim and Yoo (1995) and Chauvet (1998) obtain the left-hand side in the same way, but Kim and Nelson (1998) obtain the right-hand side, which is not exactly what we want. This problem does not occur to the smoothed estimate, because we have for $t \leq T$,

$$\mathrm{E}\left(\sum_{j=1}^{t} f_j \middle| \mathbf{y}_T\right) = \sum_{j=1}^{t} \mathrm{E}(f_j|\mathbf{y}_T)$$

Since we can obtain the left-hand side from the right-hand side, we do not need the cumulative common factor in the state vector.

The plan of the paper is as follows. Section 2 sets up a static one-factor model for monthly series, including latent series underlying quarterly series, and derives a state-space model for observable monthly and quarterly series. Section 3 explains estimation of state-space models for mixed-frequency series. Section 4 applies the method to the US quarterly real GDP and monthly coincident indicators to obtain a new coincident index of business cycles, and compares it with other coincident indices. Section 5 concludes.

2. THE MODEL

2.1. One-factor Model

Let $\{Y_{1,t}\}_{t=-\infty}^{\infty}$ be an N_1-variate random sequence of quarterly indicators observable every third period, and $\{Y_{2,t}\}_{t=-\infty}^{\infty}$ be an N_2-variate random sequence of monthly indicators. Let $N := N_1 + N_2$. Assume that logs of the indicators are integrated of order 1. Let $\{Y_{1,t}^*\}_{t=-\infty}^{\infty}$ be an N_1-variate latent random sequence such that for all t,

$$\ln Y_{1,t} = \frac{1}{3}(\ln Y_{1,t}^* + \ln Y_{1,t-1}^* + \ln Y_{1,t-2}^*) \tag{1}$$

i.e. $Y_{1,t}$ is the geometric mean of $Y_{1,t}^*$, $Y_{1,t-1}^*$, and $Y_{1,t-2}^*$. (The natural log applies to each component of the vector.) Taking the three-period differences, for all t,

$$\begin{aligned}\ln Y_{1,t} - \ln Y_{1,t-3} = {} & \frac{1}{3}(\ln Y_{1,t}^* - \ln Y_{1,t-3}^*) + \frac{1}{3}(\ln Y_{1,t-1}^* - \ln Y_{1,t-4}^*) \\ & + \frac{1}{3}(\ln Y_{1,t-2}^* - \ln Y_{1,t-5}^*)\end{aligned}$$

or

$$
\begin{aligned}
y_{1,t} &= \frac{1}{3}(y^*_{1,t} + y^*_{1,t-1} + y^*_{1,t-2}) + \frac{1}{3}(y^*_{1,t-1} + y^*_{1,t-2} + y^*_{1,t-3}) \\
&\quad + \frac{1}{3}(y^*_{1,t-2} + y^*_{1,t-3} + y^*_{1,t-4}) \\
&= \frac{1}{3}y^*_{1,t} + \frac{2}{3}y^*_{1,t-1} + y^*_{1,t-2} + \frac{2}{3}y^*_{1,t-3} + \frac{1}{3}y^*_{1,t-4} \qquad (2)
\end{aligned}
$$

where $y_{1,t} := \Delta_3 \ln Y_{1,t}$ and $y^*_{1,t} := \Delta \ln Y^*_{1,t}$. We observe $y_{1,t}$ every third period, and never observe $y^*_{1,t}$.

Note that (1) is not the usual accounting identity that links monthly levels to quarterly levels. A quarterly level is usually the *arithmetic* mean (or sum) of the monthly levels in the quarter by definition. Here it is the *geometric* mean, i.e. we define latent monthly levels in such a way. There is a tradeoff: if one wants to work with a linear state-space model, then one must accept (1); if one sticks to the accounting identity, then one must work with a non-linear state-space model, which is rather troublesome. This paper takes the first approach.

Let for all t,

$$
y_t := \begin{pmatrix} y_{1,t} \\ y_{2,t} \end{pmatrix}, \quad y^*_t := \begin{pmatrix} y^*_{1,t} \\ y_{2,t} \end{pmatrix}
$$

where $y_{2,t} := \Delta \ln Y_{2,t}$. Assume a static one-factor model for $\{y^*_t\}_{t=-\infty}^{\infty}$ such that for all t,

$$
\begin{pmatrix} y^*_{1,t} \\ y_{2,t} \end{pmatrix} = \begin{pmatrix} \mu^*_1 \\ \mu_2 \end{pmatrix} + \beta f_t + u_t \qquad (3)
$$

$$
\phi_f(L) f_t = v_{1,t} \qquad (4)
$$

$$
\Phi_u(L) u_t = v_{2,t} \qquad (5)
$$

$$
\begin{pmatrix} v_{1,t} \\ v_{2,t} \end{pmatrix} \sim \text{NID}\left(0, \begin{bmatrix} \sigma_1^2 & 0 \\ 0 & \Sigma_{22} \end{bmatrix}\right) \qquad (6)
$$

where $\beta \in \Re^N$ is a factor loading vector, $\{f_t\}_{t=-\infty}^{\infty}$ is a scalar stationary sequence of the common factor, $\{u_t\}_{t=-\infty}^{\infty}$ is an N-variate stationary sequence of the specific factors, L is the lag operator, $\phi_f(.)$ is a pth-order polynomial on $\Re$, and $\Phi_u(.)$ is a qth-order polynomial on $\Re^{N\times N}$. For identification, assume that (i) the first component of β is 1 and (ii) $\Phi_u(.)$ and Σ_{22} are diagonal; these identification restrictions are standard in factor analysis.

Since we never observe $y^*_{1,t}$, we cannot estimate (3) directly. Hence we consider the corresponding dynamic one-factor model for $\{y_t\}_{t=-\infty}^{\infty}$ such that for all t,

$$
\begin{aligned}
\begin{pmatrix} y_{1,t} \\ y_{2,t} \end{pmatrix} &= \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} + \begin{pmatrix} \beta_1 \left(\frac{1}{3}f_t + \frac{2}{3}f_{t-1} + f_{t-2} + \frac{2}{3}f_{t-3} + \frac{1}{3}f_{t-4}\right) \\ \beta_2 f_t \end{pmatrix} \\
&\quad + \begin{pmatrix} \frac{1}{3}u_{1,t} + \frac{2}{3}u_{1,t-1} + u_{1,t-2} + \frac{2}{3}u_{1,t-3} + \frac{1}{3}u_{1,t-4} \\ u_{2,t} \end{pmatrix} \qquad (7)
\end{aligned}
$$

where $\mu_1 := 3\mu^*_1$, $(\beta'_1, \beta'_2)' := \beta$, and $(u'_{1,t}, u'_{2,t})' := u_t$.

2.2. A State-space Representation

Assuming that $p, q \leq 4$, a state-space representation of (7) is

$$s_t = Fs_{t-1} + Gv_t \tag{8}$$

$$y_t = \mu + Hs_t \tag{9}$$

where

$$s_t := \begin{pmatrix} f_t \\ \vdots \\ f_{t-4} \\ u_t \\ \vdots \\ u_{t-4} \end{pmatrix}$$

$$v_t := \begin{pmatrix} v_{1,t} \\ v_{2,t} \end{pmatrix}$$

$$F := \begin{bmatrix} \begin{matrix} \phi_{f,1} & \dots & \phi_{f,p} & o'_{5-p} \\ 1 & & 0 & 0 \\ & \ddots & & \vdots \\ 0 & & 1 & 0 \end{matrix} & O_{5\times 5N} \\ O_{5N\times 5} & \begin{matrix} \Phi_{u,1} & \dots & \Phi_{u,q} & O_{N\times(5-q)N} \\ I_N & & 0 & O_{N\times N} \\ & \ddots & & \vdots \\ 0 & & I_N & O_{N\times N} \end{matrix} \end{bmatrix}$$

$$G := \begin{bmatrix} 1 & o'_N \\ 0 & o'_N \\ \vdots & \vdots \\ o_N & I_N \\ o_N & O_{N\times N} \\ \vdots & \vdots \end{bmatrix},$$

$$H := \begin{bmatrix} \frac{\beta_1}{3} & \frac{2\beta_1}{3} & \beta_1 & \frac{2\beta_1}{3} & \frac{\beta_1}{3} & \frac{1}{3}I_{N_1} & O_{N_1\times N_2} & \frac{2}{3}I_{N_1} & O_{N_1\times N_2} & \dots \\ \beta_2 & & O_{N_2\times 4} & & & O_{N_2\times N_1} & I_{N_2} & & O_{N_2\times N} & \dots \end{bmatrix}$$

where o_n is the $n \times 1$ zero vector and $O_{m\times n}$ is the $m \times n$ zero matrix.

This representation follows Stock and Watson (1991, p. 68). Note, however, that when $q < 4$, $u_{2,t-(q+1)}, \dots, u_{2,t-4}$ are irrelevant state variables, and removing them speeds up the computation.

3. ESTIMATION

3.1. Likelihood Function

Let θ be the parameter vector. Let $\{y_{1,t}^+\}_{t=-\infty}^{\infty}$ be such that for all t,

$$y_{1,t}^+ := \begin{cases} y_{1,t} & \text{if } y_{1,t} \text{ is observable} \\ z_t & \text{otherwise} \end{cases}$$

where z_t is a random draw from a distribution that does not depend on θ. Let for $t \geq 1$, $\mathbf{y}_t := (y_1, \ldots, y_t)$ and $\mathbf{y}_t^+ := (y_1^+, \ldots, y_t^+)$. Since z_t's are independent of $\mathbf{y}_T$ by construction, we can write a joint pdf of $\mathbf{y}_T^+$ as

$$f(\mathbf{y}_T^+; \theta) = f(\mathbf{y}_T; \theta) \prod_{t \in A} f(z_t)$$

where $y_{1,t}$ is missing for $t \in A \subset \{1, \ldots, T\}$; thus the likelihood function of θ given $\mathbf{y}_T$ and that given $\mathbf{y}_T^+$ are equivalent up to scale. We work with the latter because $\mathbf{y}_T^+$ does not contain missing observations.

The distribution of z_t can be anything as long as it does not depend on θ; we assume that $z_t \sim \mathrm{N}(0, I_{N_1})$ for convenience. Since the ML estimator of θ given $\mathbf{y}_T^+$ does not depend on z_t, we simply set $z_t = 0$ for its realization.

We derive a state-space model for $\{y_t^+\}_{t=-\infty}^{\infty}$ next, so that we can apply the Kalman filter to evaluate the likelihood function of θ given $\mathbf{y}_T^+$. Write (9) as

$$\begin{pmatrix} y_{1,t} \\ y_{2,t} \end{pmatrix} = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} + \begin{bmatrix} H_1 \\ H_2 \end{bmatrix} s_t$$

Then we have for all t,

$$\begin{pmatrix} y_{1,t}^+ \\ y_{2,t} \end{pmatrix} = \begin{pmatrix} \mu_{1,t} \\ \mu_2 \end{pmatrix} + \begin{bmatrix} H_{1,t} \\ H_2 \end{bmatrix} s_t + \begin{pmatrix} w_{1,t} \\ 0 \end{pmatrix}$$

where

$$\mu_{1,t} := \begin{cases} \mu_1 & \text{if } y_{1,t} \text{ is observable} \\ 0 & \text{otherwise} \end{cases},$$

$$H_{1,t} := \begin{cases} H_1 & \text{if } y_{1,t} \text{ is observable} \\ 0 & \text{otherwise} \end{cases},$$

$$w_{1,t} := \begin{cases} 0 & \text{if } y_{1,t} \text{ is observable} \\ z_t & \text{otherwise} \end{cases}.$$

Thus we have a state-space model for $\{y_t^+\}_{t=-\infty}^{\infty}$ such that for all t,

$$s_t = F s_{t-1} + G v_t \tag{10}$$

$$y_t^+ = \mu_t + H_t s_t + w_t \tag{11}$$

where

$$\mu_t := \begin{pmatrix} \mu_{1,t} \\ \mu_2 \end{pmatrix}, \quad H_t := \begin{bmatrix} H_{1,t} \\ H_2 \end{bmatrix}, \quad w_t := \begin{pmatrix} w_{1,t} \\ 0 \end{pmatrix}$$

Let for $t \geq 1$,

$$\mu_{t|t-1}(\theta) := \mathrm{E}(y_t^+ | \mathbf{y}_{t-1}^+; \theta)$$
$$\Sigma_{t|t-1}(\theta) := \mathrm{var}(y_t^+ | \mathbf{y}_{t-1}^+; \theta)$$

where $\mathbf{y}_0^+ := \emptyset$. Then for $t \geq 1$,

$$f(y_t^+ | \mathbf{y}_{t-1}^+; \theta) = (2\pi)^{-N/2} \det(\Sigma_{t|t-1}(\theta))^{-1/2} \exp\left(-\frac{1}{2}(y_t^+ - \mu_{t|t-1}(\theta))' \Sigma_{t|t-1}(\theta)^{-1} (y_t^+ - \mu_{t|t-1}(\theta))\right)$$

The log-likelihood function of θ given $\mathbf{y}_T^+$ is

$$\ln L(\theta; \mathbf{y}_T^+) = -\frac{NT}{2} \ln 2\pi - \frac{1}{2} \sum_{t=1}^{T} \ln \det \Sigma_{t|t-1}(\theta)) - \frac{1}{2} \sum_{t=1}^{T} (y_t^+ - \mu_{t|t-1}(\theta))' \Sigma_{t|t-1}(\theta)^{-1} (y_t^+ - \mu_{t|t-1}(\theta))$$

To evaluate this, we must evaluate $\{\mu_{t|t-1}(\theta), \Sigma_{t|t-1}(\theta)\}_{t=1}^{T}$. Let for $t, s \geq 0$,

$$\hat{s}_{t|s} := \mathrm{E}(s_t | \mathbf{y}_s^+; \theta)$$
$$P_{t|s} := \mathrm{var}(s_t | \mathbf{y}_s^+; \theta)$$

From (11), for $t \geq 1$,

$$\mu_{t|t-1}(\theta) = \mu_t + H_t \hat{s}_{t|t-1}$$
$$\Sigma_{t|t-1}(\theta) = H_t P_{t|t-1} H_t' + \Sigma_{ww,t}$$

where

$$\Sigma_{ww,t} := \begin{cases} O_{N \times N} & \text{if } y_{1,t} \text{ is observable} \\ \begin{bmatrix} I_{N_1} & O_{N_1 \times N_1} \\ O_{N_1 \times N_2} & O_{N_2 \times N_2} \end{bmatrix} & \text{otherwise} \end{cases}.$$

Given θ, we can evaluate $\{\hat{s}_{t|t-1}, P_{t|t-1}\}_{t=1}^{T}$ using the Kalman filter.

3.2. Kalman Filter

Initial state

To start the Kalman filter, one must specify $\hat{s}_{1|0}$ and $P_{1|0}$, the unconditional mean and variance of s_1. Given stationarity, we can show that

$$\hat{s}_{1|0} = 0 \tag{12}$$
$$\mathrm{vec}(P_{1|0}) = (I_{(5+5N)^2} - F \otimes F)^{-1} \mathrm{vec}(G \Sigma_{vv} G') \tag{13}$$

The second equation involves inversion of a large matrix.

Alternatively, one can simply set

$$\hat{s}_{0|0} = 0 \tag{14}$$

$$P_{0|0} = 0 \tag{15}$$

which implies that

$$\hat{s}_{1|0} = 0 \tag{16}$$

$$P_{1|0} = G\Sigma_{vv}G' \tag{17}$$

The resulting estimator of θ is asymptotically equivalent to the ML estimator.

Updating
The Kalman gain matrix is for $t \geq 1$,

$$B_t := P_{t|t-1}H_t'(H_t P_{t|t-1}H_t' + \Sigma_{ww,t})^{-1} \tag{18}$$

The updating equations for the state vector and its variance–covariance matrix are for $t \geq 1$,

$$\hat{s}_{t|t} = \hat{s}_{t|t-1} + B_t(y_t^+ - \mu_t - H_t\hat{s}_{t|t-1}) \tag{19}$$

$$P_{t|t} = P_{t|t-1} - B_t H_t P_{t|t-1} \tag{20}$$

Prediction
The prediction equations for the state vector and its variance–covariance matrix are for $t \geq 1$,

$$\hat{s}_{t|t-1} = F\hat{s}_{t-1|t-1} \tag{21}$$

$$P_{t|t-1} = FP_{t-1|t-1}F' + G\Sigma_{vv}G' \tag{22}$$

Combining the updating and prediction equations, we obtain $\{\hat{s}_{t|t-1}, P_{t|t-1}\}_{t=1}^{T}$.

3.3. Fixed-interval Smoothing

We are interested in estimation of the common factor, i.e. the first component of the state vector. The smoothing equation for the state vector is for $t = 1, \ldots, T$,

$$\hat{s}_{t|T} = \hat{s}_{t|t} + P_{t|t}F'P_{t+1|t}^{-1}(\hat{s}_{t+1|T} - \hat{s}_{t+1|t}) \tag{23}$$

See Hamilton (1994, sec. 13.6) for the derivation.

In practice, one may not want to invert $P_{t+1|t}$ numerically when its dimension is large. The following algorithm by de Jong (1988, 1989) is useful in such cases; see also Koopman (1998). Let for $t = 1, \ldots, T + 1$,

$$r_t := P_{t|t-1}^{-1}(\hat{s}_{t|T} - \hat{s}_{t|t-1})$$

Then for $t = 1, \ldots, T + 1$,

$$\hat{s}_{t|T} = \hat{s}_{t|t-1} + P_{t|t-1}r_t$$

Inserting (19) into (23), for $t = 1, \ldots, T$,

$$\hat{s}_{t|T} = \hat{s}_{t|t-1} + B_t(y_t^+ - \mu_t - H_t\hat{s}_{t|t-1}) + P_{t|t}F'P_{t+1|t}^{-1}(\hat{s}_{t+1|T} - \hat{s}_{t+1|t})$$

Comparing the previous two equations, for $t = 1, \ldots, T$,

$$\begin{aligned} P_{t|t-1}r_t &= B_t(y_t^+ - \mu_t - H_t\hat{s}_{t|t-1}) + P_{t|t}F'P_{t+1|t}^{-1}(\hat{s}_{t+1|T} - \hat{s}_{t+1|t}) \\ &= B_t(y_t^+ - \mu_t - H_t\hat{s}_{t|t-1}) + P_{t|t}F'r_{t+1} \end{aligned}$$

or using (18) and (20),

$$\begin{aligned} r_t &= P_{t|t-1}^{-1}B_t(y_t^+ - \mu_t - H_t\hat{s}_{t|t-1}) + P_{t|t-1}^{-1}P_{t|t}F'r_{t+1} \\ &= H_t'(H_tP_{t|t-1}H_t' + \Sigma_{ww,t})^{-1}(y_t^+ - \mu_t - H_t\hat{s}_{t|t-1}) + (I - H_t'B_t')F'r_{t+1} \end{aligned}$$

The algorithm starts from $r_{T+1} := 0$ and iterates for $t = T, \ldots, 1$,

$$\begin{aligned} r_t &= H_t'(H_tP_{t|t-1}H_t' + \Sigma_{ww,t})^{-1}(y_t^+ - \mu_t - H_t\hat{s}_{t|t-1}) + (I - H_t'B_t')F'r_{t+1} \\ \hat{s}_{t|T} &= \hat{s}_{t|t-1} + P_{t|t-1}r_t \end{aligned}$$

4. NEW COINCIDENT INDEX

4.1. Data

We apply the method to US coincident business cycle indicators to obtain a new coincident index of business cycles. The indicators are quarterly real GDP and the four monthly coincident indicators that currently make up the CI; see Table I for their descriptions. The sample period is from January 1959 to December 2000. To stationarize the series, we take the first difference of the natural log of each series and multiply it by 100, which is approximately equal to the quarterly or monthly percentage growth rate.

Table II summarizes descriptive statistics of the growth rate series. The mean monthly growth rate of EMP (0.18%) is lower than those of the others including GDP (0.28–0.29%), and the standard deviations of the growth rates of EMP and INC are smaller than those of IIP and SLS. The low mean and the small standard deviation of the growth rate of EMP strongly pull down

Table I. US coincident business cycle indicators

Indicator	Description
	Quarterly
GDP	Real GDP (billions of chained 1996 dollars, SA, AR)
	Monthly
EMP	Employees on non-agricultural payrolls (thousands, SA)
INC	Personal income less transfer payments (billions of chained 1996 dollars, SA, AR)
IIP	Index of industrial production (1992 = 100, SA)
SLS	Manufacturing and trade sales (millions of chained 1996 dollars, SA)

Note: SA means 'seasonally-adjusted' and AR means 'annual rate'.

Table II. Descriptive statistics of the indicators

Indicator	Mean	S.D.	Min.	Max.
		Quarterly		
$\Delta \ln \text{GDP}$	0.84	0.90	−2.06	3.78
		Monthly		
$\Delta \ln \text{EMP}$	0.18	0.23	−0.86	1.23
$\Delta \ln \text{INC}$	0.28	0.40	−1.10	1.61
$\Delta \ln \text{IIP}$	0.29	0.87	−4.25	6.00
$\Delta \ln \text{SLS}$	0.28	1.04	−3.21	3.54
$\Delta \ln \text{CI}$	0.24	0.36	−1.44	1.89

the mean growth rate of CI (0.24%), which is a weighted average of the growth rates of the four monthly indicators using weights proportional to the inverses of their standard deviations.

4.2. Estimation Result

Following Stock and Watson (1991), we take two shortcuts in estimating the dynamic one-factor model (4)–(7). First, to reduce the number of parameters, we demean the series and delete the constant term from (7). To identify the common factor as the common factor component in the growth rate of latent monthly real GDP, however, we do not want to standardize $\Delta \ln \text{GDP}$; hence, contrary to Stock and Watson (1991), we do not normalize the variance of each series to be 1. Second, we use the approximate ML estimator instead of the exact one regarding the initial state for the Kalman filter. These shortcuts are common in practice; for instance, we often estimate an AR model with normal innovations by applying OLS to the demeaned series.

These shortcuts are important for applications. Computation of the exact ML estimator involves inversion of a large matrix in (13). In our case, we must invert a 900×900 matrix ($N = 5$) in each iteration, which is time-consuming. Using Ox 3.10 by Doornik (2001) on a Pentium III (800 MHz) processor, estimation of the finally selected model ($p = 1$, $q = 2$) by a quasi-Newton method (Broyden–Fletcher–Goldfarb–Shanno algorithm) from an *ad hoc* initial guess takes about 9 hours for the exact ML estimator, while less than 20 minutes for the approximate one. It takes even longer without demeaning.

Note that using a quasi-Newton method, we only obtain a local ML estimator; there is no guarantee that it coincides with the global ML estimator.

Before estimation, we must determine p and q, the orders of AR models for the common and specific factors respectively. One may use a model selection criterion, such as Akaike's information criterion (AIC) or Schwartz's Bayesian information criterion (SBIC), for that purpose. For our model,

$$\text{AIC} := -\frac{1}{T}\{\ln L(\hat{\theta}) - [(N-1) + p + 1 + N(q+1)]\}$$

$$\text{SBIC} := -\frac{1}{T}\left\{\ln L(\hat{\theta}) - \frac{\ln T}{2}[(N-1) + p + 1 + N(q+1)]\right\}$$

where $\hat{\theta}$ is the (approximate) ML estimator of θ. We find that AIC selects $(p, q) = (1, 3)$ while SBIC selects $(p, q) = (1, 2)$. We follow SBIC here, preferring the simpler model.

Table III. Estimation result for the one-factor model with real GDP

Parameter	$\Delta \ln \text{GDP}$	$\Delta \ln \text{EMP}$	$\Delta \ln \text{INC}$	$\Delta \ln \text{IIP}$	$\Delta \ln \text{SLS}$
β	1.00	0.49	0.81	2.14	1.74
		(0.04)	(0.06)	(0.13)	(0.11)
ϕ_f			0.56		
			(0.05)		
σ_1^2			0.08		
			(0.01)		
$\phi_{u,1}$	−0.04	0.10	−0.05	−0.05	−0.41
	(0.08)	(0.04)	(0.05)	(0.07)	(0.05)
$\phi_{u,2}$	−0.83	0.45	0.03	−0.06	−0.20
	(0.07)	(0.05)	(0.05)	(0.06)	(0.05)
Σ_{22}	0.19	0.02	0.09	0.25	0.61
	(0.04)	(0.00)	(0.01)	(0.02)	(0.04)

Note: Numbers in parentheses are asymptotic standard errors.

Table III summarizes the estimation result. Since the standard deviations of the growth rates of the indicators are different, we cannot compare the factor loadings directly; we must compare the factor loadings for the standardized series. Instead of re-estimating the model for the standardized series, we simply look at the factor loadings divided by the standard deviations of the growth rates of the indicators. This standardized factor loading is largest for $\Delta \ln \text{IIP}$ (2.14/0.87) and smallest for $\Delta \ln \text{SLS}$ (1.74/1.04), essentially the same result as that in Stock and Watson (1991, Table 4.1). Without knowing the standard deviation of the growth rate of latent monthly real GDP, we cannot compare the factor loading for $\Delta \ln \text{GDP}$ with others.

The common factor has positive autocorrelations. The specific factors have different time series properties: the AR coefficients are positive for $\Delta \ln \text{EMP}$, almost zero for $\Delta \ln \text{INC}$ and $\Delta \ln \text{IIP}$, and negative for $\Delta \ln \text{GDP}$ and $\Delta \ln \text{SLS}$. The negative AR coefficients may look odd at first sight. We find, however, that if we fit low-order ARMA models to the growth rate series of the indicators, then we often get negative MA coefficients. The MA part may correspond to the specific factor. In any case, the results for the monthly indicators are very close to that in Stock and Watson (1991, Table 4.1).

Given the estimates of the model parameters, we apply fixed-interval smoothing to obtain the smoothed estimates of the common factor, from which we construct our new coincident index as follows:

(1) Add the mean growth rate of latent monthly real GDP, i.e. the mean growth rate of quarterly real GDP divided by 3, to the smoothed estimates of the common factor and divide them by 100.
(2) Take the partial sums, and then take the exponentials.

The new coincident index is an estimate of the common factor component in latent monthly real GDP. One may not want to take the exponentials in the last step, because this only gives the exponential of the expectation of the natural log of the common factor, which is not the expectation of the common factor. It does not matter for determining business cycle turning points anyway; monotone transformations do not affect them.

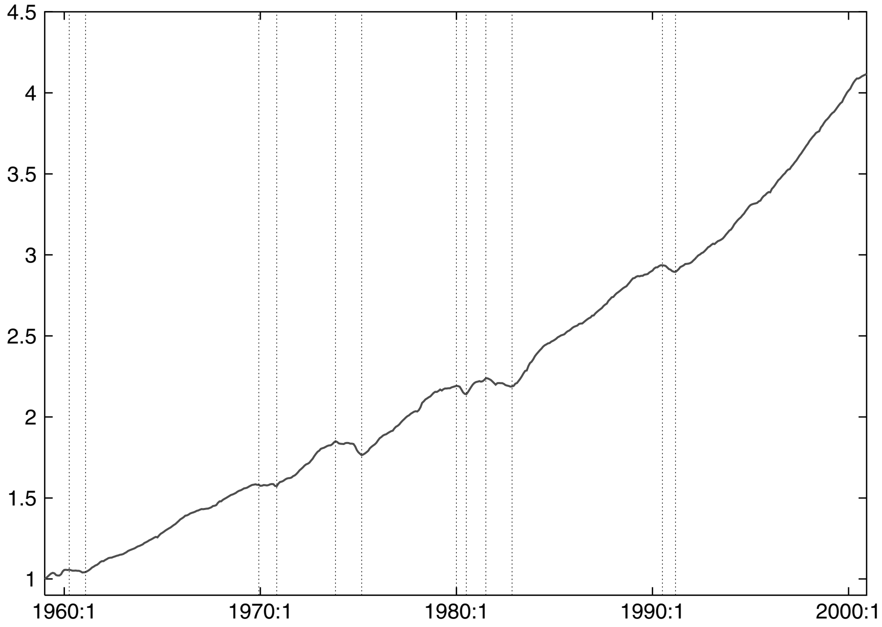

Figure 2. Historical plot of the new coincident index (1959 : 1=1). The vertical lines are the NBER business cycle reference dates

Figure 2 plots the new coincident index. It seems to capture the NBER business cycle reference dates very well.

4.3. Comparison with Other Indices

We compare our new coincident index with the CI and the S–W coincident index constructed from our data.

In the USA, The Conference Board calculates the CI of coincident indicators in the following five steps:

(1) Construct the monthly symmetric growth rate series of the indicators.
(2) Compute the standard deviation for each symmetric growth rate series, excluding outliers.
(3) Take a weighted average of the symmetric growth rate series, using weights proportional to the inverses of their standard deviations, to obtain the symmetric growth rate series of the CI.
(4) Convert the symmetric growth rate series to the level series.
(5) Rebase the level series to average 100 in the base year.

See the December 1996 issue of *Business Cycle Indicators* for details. For comparison with our new coincident index, we take the first difference of the natural log instead of the symmetric growth rate, and do not exclude outliers when computing the standard deviation.

Stock and Watson (1991) start from the static one-factor model (3)–(6) to construct the S–W coincident index. For identification of the model parameters, they assume that $\sigma_1^2 = 1$ instead of restricting β. They take the following shortcuts in estimation: standardize the series, delete the constant term from (3), and use the approximate ML estimator. They select $(p, q) = (2, 2)$ based on a likelihood-ratio test. We select $(p, q) = (1, 3)$ for our data, following AIC and SBIC. Table IV summarizes the estimation result. Although we do not have $\Delta \ln \text{GDP}$ here, the result is similar to that in Table III; the difference in the factor loading vectors comes from (i) different identification restrictions on the parameters, and (ii) different normalization of the series. The result is also similar to that in Stock and Watson (1991, Table 4.1).

As a by-product of ML estimation of the model parameters, the Kalman filter gives a sequence of the updated estimates of the common factor. From this, we construct the S–W coincident index as follows:

(1) Add the mean of the common factor defined below to the updated estimates of the common factor and divide them by 100.
(2) Take the partial sums, and then take the exponentials.

Kim and Nelson (1999, sec. 3.5) define the S–W coincident index in this way. As we note in the Introduction, the original S–W coincident index is essentially the updated estimate of the *cumulative* common factor, and not the cumulative sum of the updated estimates of the common factor. We do not pursue the effect of this distinction here, however.

Stock and Watson (1991) identify the mean of the common factor as follows. Combining the updating equation (19) and the prediction equation (21), and using the lag operator L,

$$\begin{aligned}\hat{s}_{t|t} &= F\hat{s}_{t-1|t-1} + B_t(y_t - \mu - HF\hat{s}_{t-1|t-1}) \\ &= (I - B_tH)FL\hat{s}_{t|t} + B_t(y_t - \mu) \\ &= [I - (I - B_tH)FL]^{-1}B_t(y_t - \mu)\end{aligned}$$

where I is the identity matrix (μ and H are time-independent without quarterly series). Let $W(L) := [I - (I - BH)F]^{-1}B$, where B is the steady-state Kalman gain matrix. In the steady state,

$$\hat{s}_{t|t} = W(L)(y_t - \mu)$$

Table IV. Estimation result for the one-factor model without real GDP

Parameter	Δ ln EMP	Δ ln INC	Δ ln IIP	Δ ln SLS
β	0.55	0.52	0.69	0.45
	(0.03)	(0.04)	(0.04)	(0.03)
ϕ_f		0.56		
		(0.05)		
$\phi_{u,1}$	−0.04	−0.02	−0.10	−0.38
	(0.05)	(0.05)	(0.08)	(0.05)
$\phi_{u,2}$	0.44	0.07	−0.13	−0.16
	(0.05)	(0.05)	(0.07)	(0.05)
$\phi_{u,3}$	0.26	0.10	−0.04	0.05
	(0.06)	(0.05)	(0.07)	(0.05)
Σ_{22}	0.31	0.57	0.29	0.58
	(0.03)	(0.04)	(0.03)	(0.04)

Note: Numbers in parentheses are asymptotic standard errors.

The first component of $\hat{s}_{t|t}$ is the updated estimate of the common factor, which is a linear combination of the current and past y_t. Stock and Watson (1991) use these weights, i.e. the first row of $W(L)$, to identify the mean of the common factor. Thus, given stationarity, they identify the mean of the common factor as the first component of $W(1)\mu$. Note that this identification is valid only for the updated estimates.

Table V summarizes descriptive statistics of the monthly growth rate series of alternative indices. To separate out the effect of including real GDP from that of smoothing, we look at both the updated and the smoothed versions of our new coincident index. The mean monthly growth rate of the S–W coincident index is high, and the standard deviation is large. We see that smoothing reduces volatility for our new coincident index.

Table VI shows correlations between the monthly growth rates of these indices. The CI has relatively low correlations with others. The S–W coincident index and the updated version of our new coincident index have an extremely high correlation. This is probably because the updated version of our new coincident index can use information in the current real GDP only in the third month of each quarter. Since the smoothed version of our new coincident index can use such information every month, the correlation between the S–W coincident index and the smoothed version of our new coincident index is not that high. Smoothing is crucial when one includes quarterly series.

Table VII compares business cycle turning points determined by each index with the NBER business cycle reference dates. The CI captures the NBER reference dates best, but our new coincident index also performs well. The S–W coincident index signals the peak in January 1980 too early, which is the same result as in Stock and Watson (1991, Figure 4.1).

Figure 3 plots the CI, the S–W coincident index, and our new coincident index from 1979 to 1983, during which there are two peaks and two troughs. We see that the three indices have notable differences. First, the S–W coincident index signals the peak in January 1980 too early.

Table V. Descriptive statistics of alternative indices

Index	Mean	S.D.	Min.	Max.
$\Delta \ln \mathrm{CI}$	0.24	0.36	−1.44	1.89
$\Delta \ln \mathrm{SW}$	0.41	1.16	−5.75	6.30
$\Delta \ln \mathrm{New}^u$	0.28	0.32	−1.36	1.83
$\Delta \ln \mathrm{New}^s$	0.28	0.31	−1.25	1.74

Note: 'SW' and 'New' denote the S–W coincident index and the new coincident index respectively. The superscripts u and s denote the updated and smoothed estimates respectively.

Table VI. Correlations between alternative indices

	$\Delta \ln \mathrm{CI}$	$\Delta \ln \mathrm{SW}$	$\Delta \ln \mathrm{New}^u$	$\Delta \ln \mathrm{New}^s$
$\Delta \ln \mathrm{CI}$	1.000			
$\Delta \ln \mathrm{SW}$	0.964	1.000		
$\Delta \ln \mathrm{New}^u$	0.972	0.997	1.000	
$\Delta \ln \mathrm{New}^s$	0.960	0.984	0.986	1.000

Note: See the note to Table V.

Table VII. Business cycle turning points determined by alternative indices

NBER	CI	SW	New^u	New^s
		Peaks		
1960/4	0	−2	0	0
1969/12	−2	−2	−2	−2
1973/11	0	0	0	0
1980/1	0	−10	+1	0
1981/7	0	0	+1	0
1990/7	−1	−4	+1	0
		Troughs		
1961/2	0	0	0	−2
1970/11	0	0	0	0
1975/3	0	0	0	0
1980/7	0	0	0	0
1982/11	+1	+1	0	−1
1991/3	0	0	0	0

Note: See the note to Table V. The numbers are lags from the NBER business cycle reference dates.

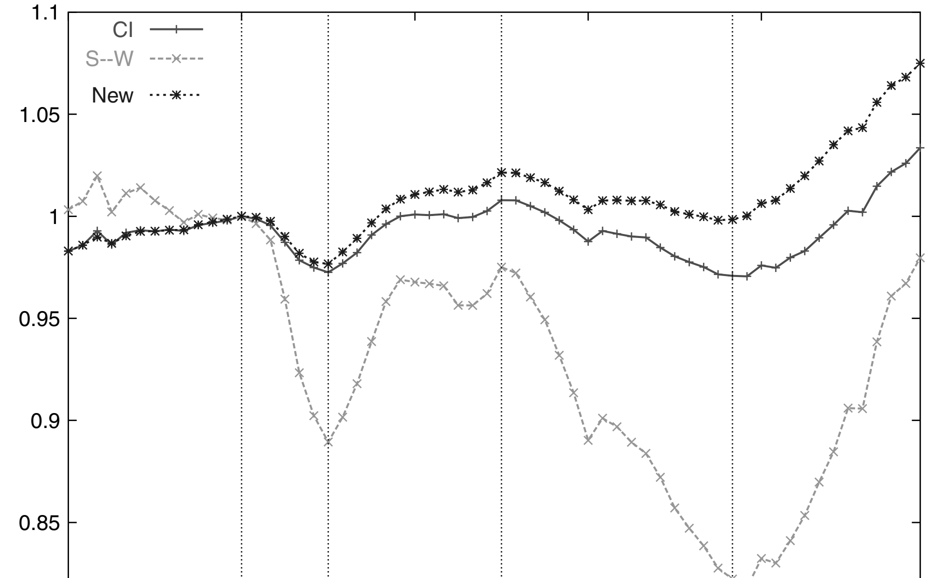

Figure 3. Comparison of alternative indices from 1979 to 1983 (1980 : 1=1). The vertical lines are the NBER business cycle reference dates

Second, our new coincident index signals the trough in November 1982 one month early, while the other two signal it one month late. Third, the heights of the peaks and the depth of the troughs indicated by the three indices are very different. For the peaks, the CI and our new coincident index indicate that the peak in July 1981 is higher than that in January 1980, while the S–W coincident index indicates the opposite. For the troughs, the CI and the S–W coincident index indicate that the trough in November 1982 is deeper than that in July 1980, while our new coincident index indicates the opposite.

In conclusion, we find significant differences between the three indices. Recall that our new coincident index is an extension of the S–W coincident index that has an economic interpretation as the common factor component in latent monthly real GDP. As such, the level of our new coincident index resembles that of observable quarterly real GDP. Such an index should be more appealing than those without economic interpretation.

5. CONCLUSION

Here we summarize our concern about the S–W coincident index described in this paper:

(1) The S–W coincident index is essentially the updated estimate of the cumulative common factor among coincident indicators, but some people use the cumulative sum of the updated estimates of the common factor instead. The two are not identical. This distinction is irrelevant for the smoothed estimate.
(2) The identification method for the mean growth rate of the S–W coincident index is valid only for the updated estimate. Its extension to the smoothed estimate is not immediate.
(3) The updated estimate cannot use information in quarterly indicators in the first two months of each quarter, because such information is not available yet. Hence smoothing is crucial when one includes quarterly indicators.
(4) Without economic interpretation, the level of the S–W coincident index may not coincide with the levels of important indicators such as real GDP. This may cause a false signal; for instance, compare March 1979 and January 1980 in Figure 3.

Our concern also applies to the recent extensions of the S–W coincident index that introduce Markov regime-switching, e.g. Kim and Yoo (1995), Chauvet (1998), and Kim and Nelson (1998). Including quarterly indicators in these extensions seems straightforward theoretically, but may be cumbersome numerically.

ACKNOWLEDGEMENTS

We thank Konstantin Kholodilin, Mitsuru Nakagawa, and Haruhisa Nishino for useful comments.

REFERENCES

Brockwell PJ, Davis RA. 1991. *Time Series: Theory and Methods*, 2nd edn. Springer-Verlag: New York.
Brockwell PJ, Davis RA, Salehi H. 1991. A state-space approach to transfer-function modeling. In *Statistical Inference in Stochastic Processes*, Prabhu NU, Basawa IV (eds). Marcel Dekker: New York.
Chauvet M. 1998. An econometric characterization of business cycle dynamics with factor structure and regime switching. *International Economic Review* **39**: 969–996.
de Jong P. 1988. A cross-validation filter for time series models. *Biometrika* **75**: 594–600.
de Jong P. 1989. Smoothing and interpolation with the state-space model. *Journal of the American Statistical Association* **84**: 1085–1088.

Doornik JA. 2001. *Ox: An Object-Oriented Matrix Language*, 4th edn. Timberlake Consultants: New York.
Hamilton JD. 1994. *Time Series Analysis*. Princeton University Press: Princeton, NJ.
Kim CJ, Nelson CR. 1998. Business cycle turning points, a new coincident index, and tests of duration dependence based on a dynamic factor model with regime switching. *Review of Economics and Statistics* **80**: 188–201.
Kim CJ, Nelson CR. 1999. *State-Space Models with Regime Switching*. The MIT Press: Cambridge, MA.
Kim MJ, Yoo JS. 1995. New index of coincident indicators: A multivariate Markov switching factor model approach. *Journal of Monetary Economics* **36**: 607–630.
Koopman SJ. 1998. Kalman filtering and smoothing. In *Encyclopedia of Biostatistics*, Peter A, Theodore C (eds). John Wiley: New York.
Shumway RH, Stoffer DS. 1982. An approach to time series smoothing and forecasting using the EM algorithm. *Journal of Time Series Analysis* **3**: 253–265.
Shumway RH, Stoffer DS. 2000. *Time Series Analysis and Its Applications*. Springer-Verlag: New York.
Stock JH, Watson MW. 1989. New indexes of coincident and leading economic indicators. *NBER Macroeconomics Annual* **4**: 351–409.
Stock JH, Watson MW. 1991. A probability model of the coincident economic indicators. In *Leading Economic Indicators*, Lahiri K, Moore GH (eds). Cambridge University Press: New York.

