NBER WORKING PAPER SERIES

A PROBABILITY MODEL OF THE COINCIDENT ECONOMIC INDICATORS

James H. Stock

Mark W. Watson

Working Paper No. 2772

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
November 1988

The authors thank Robert Hall, Geoffrey Moore, Lawrence Summers, John Taylor, and numerous colleagues for helpful advice. Myungsoo Park provided skillful research assistance. An earlier version of this paper was circulated under the title, "What Do the Leading Indicators Lead? A New Approach to Forecasting Swings in Economic Activity." Financial support for this project was provided by the National Science Foundation and the National Bureau of Economic Research. This research is part of NBER's research program in Economic Fluctuations. Any opinions expressed are those of the authors not those of the National Bureau of Economic Research.

NBER Working Paper #2772
November 1988

A PROBABILITY MODEL OF THE COINCIDENT ECONOMIC INDICATORS

ABSTRACT

The Index of Coincident Economic Indicators, currently compiled by the U.S. Department of Commerce, is designed to measure the state of overall economic activity. The index is constructed as a weighted average of four key macroeconomic time series, where the weights are obtained using rules that date to the early days of business cycle analysis. This paper presents an explicit time series model (formally, a dynamic factor analysis or "single-index" model) that implicitly defines a variable that can be thought of as the overall state of the economy. Upon estimating this model using data from 1959-1987, the estimate of this unobserved variable is found to be highly correlated with the official Commerce Department series, particularly over business cycle horizons. Thus this model provides a formal rationalization for the traditional methodology used to develop the Coincident Index. Initial exploratory exercises indicate that traditional leading variables can prove useful in forecasting the short-run growth in this series.

James H. Stock
Kennedy School of Government
Harvard University
Cambridge, MA 02138

Mark W. Watson
Department of Economics
Northwestern University
Evanston, IL 60208

1. Introduction

Since their initial development in 1938 by Wesley Mitchell, Arthur Burns, and their colleagues at the National Bureau of Economic Research, the Composite Indexes of Coincident and Leading Economic Indicators have played an important role in summarizing the state of macroeconomic activity. This paper reconsiders the problem of constructing an index of coincident indicators. We will use the techniques of modern time series analysis to develop an explicit probability model of the four coincident variables that comprise the Index of Coincident Economic Indicators (CEI) currently compiled by the Department of Commerce (DOC). This probability model provides a framework for computing an alternative coincident index. As it turns out, this alternative index is quantitatively similar to the DOC index. Thus this probability model provides a formal statistical rationalization for, and interpretation of, the construction of the DOC CEI. This alternative interpretation complements that provided by the methodology developed by Mitchell and Burns (1938) and applied by, for example, Zarnowitz and Boschan (1975).

The model adopted in this paper is based on the notion that the comovements in many macroeconomic variables have a common element that can be captured by a single underlying, unobserved variable. In the abstract, this variable represents the general "state of the economy." The problem is to estimate the current state of the economy, i.e. this common element in the fluctuations of key aggregate time series variables. This unobserved variable -- the "state of the economy" -- must be defined before any attempt can be made to estimate it. In technical terms, this requires formulating a probability model that provides a mathematical definition of the unobserved state of the economy. In nontechnical terms, this problem can be phrased as a question: What do the leading indicators lead?

Our proposed answer to this question is given in Section 2. This section presents a parametric "single-index" model in which the state of the economy -- referred to as C_t -- is an unobserved variable common to multiple aggregate time series. Because this model is linear in the unobserved variables, the Kalman Filter can be used to construct the Gaussian likelihood function and thereby to estimate the unknown parameters of the model by maximum likelihood. As a side benefit, the Kalman Filter automatically computes the minimum mean square error estimate of C_t using data through period t. This estimate, $C_{t|t}$, is the alternative index of coincident indicators computed using the single-index model.

The single-index model is estimated using data on industrial production, real personal income, real manufacturing and trade sales, and employment in nonagricultural establishments from 1959 to 1987. The results are reported in Section 3. Also in this section, the estimated alternative index $C_{t|t}$ is compared with the DOC series. The similarity between the two is striking, particularly over business-cycle horizons.

Section 4 presents an initial investigation into forecasting the growth of $C_{t|t}$ using a variety of leading or predictive macroeconomic variables. The main conclusion is that a parsimoniously parameterized time series model with $C_{t|t}$ and six leading variables can forecast approximately two-thirds of the variance of the growth in $C_{t|t}$ over the next six months.

A conceptually distinct forecasting problem is explored in Section 5. A traditional focus of business cycle analysis has of course been identifying expansions and contractions. Several recent forecasting exercises have focused on forecasting turning points; see, for example, Hymans (1973), Wecker (1979), Zarnowitz and Moore (1982), Kling (1987), and Zellner, Hong and Gulati (1987). Rather than focusing on turning points, the approach taken in Section 5 is to forecast directly the binary variable representing whether the economy is in a recession or expansion six months hence. The main conclusion is that, among the binary-response models considered, expansions can be forecasted fairly reliably, recessions less so. Section 6 concludes.

2. The Coincident Indicator Model: Specification and Estimation

One approach to studying aggregate fluctuations is to pick an important economic time series -- say employment or GNP -- as the object of interest for subsequent analysis and forecasting. This decided, life becomes relatively easy, since economists have decades of experience constructing models to analyze and to forecast observable time series variables. From the perspective of business cycle analysis, however, this approach is rather limited. Individual series measure more or less well-defined concepts, such as the value of all goods and services produced in a quarter or the total number of individuals working for pay. But these series measure only various facets of the overall state of economic activity; none measure the state of the economy (in Burns and Mitchell's (1946) terminology, the "reference cycle") directly. Moreover, even the concepts that the series purport to measure are measured with error.[1]

The formulation developed here is based instead on the assumption that there is a single unobserved variable common to many macroeconomic time series. This places Burns and Mitchell's (1946) reference cycle in a fully specified probability model. The proposed model is a parametric version of the "single-index" models discussed by Sargent and Sims (1977), in which the single unobserved index is common to multiple macroeconomic variables. Estimates of this unobserved index, constructed using variables that move contemporaneously with this index, provide an alternative index of coincident indicators. This index can then be forecasted using leading variables.

A. The Single-Index Model

Let X_t denote a nx1 vector of macroeconomic time series variables that are hypothesized to move contemporaneously with overall economic conditions. In the single-index model, X_t consists of two stochastic components: the common unobserved scalar time series variable, or "index", C_t, and a n-dimensional component that represents idiosyncratic movements in the series and measurement error, v_t. Both the unobserved index and the idiosyncratic component are modeled as having linear stochastic structures. In addition, C_t is assumed to enter each of the variables contemporaneously. This suggests the formulation:

(1) $$X_t = \beta + \gamma C_t + v_t$$

(2) $$\phi(L)C_t = \delta + \eta_t$$

(3) $$D(L)v_t = \epsilon_t$$

where L denotes the lag operator, $\tilde{\phi}(L)$ is a scalar lag polynomials, and $\tilde{D}(L)$ is a lag polynomial matrix. According to (1), C_t enters each of the n equations in (1), although with varying lags and weights.

As an empirical matter, many macroeconomic time series are well characterized as containing stochastic trends; see, for example, Nelson and Plosser (1982). A theoretical possibility is that these stochastic trends would enter through C_t; in this case, each element of X_t would contain a stochastic trend, but this trend would be common to each element. Thus X_t would be cointegrated of order k-1 as defined by Engle and Granger (1987). Looking ahead to the empirical results, however, this turns out not to be the case: while we cannot reject the hypothesis that the coincident series we consider individually contain a stochastic trend, neither can we reject the hypothesis that there is no cointegration among these variables.[2] The system (1)-(3) is therefore reformulated in terms of the changes (or, more precisely, the growth rates) of the variables. Specifically, assume that $\tilde{\phi}(L)$ and $\tilde{D}(L)$ can be factored so that $\tilde{\phi}(L)=\phi(L)\Delta$ and $\tilde{D}(L)=D(L)\Delta$, where $\Delta=1-L$. Let $Y_t=\Delta X_t$ and $u_t=\Delta v_t$, so that (1)-(3) become:

(4) $$Y_t = \beta + \gamma\Delta C_t + u_t$$

(5) $$\phi(L)\Delta C_t = \delta + \eta_t$$

(6) $$D(L)u_t = \epsilon_t \ .$$

In practice, X_t will be a vector of the logarithms of time series variables, so that Y_t is a vector of their growth rates. The lag polynomials $\phi(L)$ and $D(L)$ are assumed to have finite orders p and k, respectively.

The main identifying assumption in the model expresses the core notion of the single-index model that the comovements of the multiple time series arise from the single source C_t. This is made precise by assuming that $(u_{1t},\ldots,u_{nt},\Delta C_t)$ are mutually uncorrelated at all leads and lags. (When there are four or more variables this imposes testable overidentifying restrictions, which will be examined empirically below.) This is achieved by assuming that D(L) is diagonal and that the n+1 disturbances are mutually uncorrelated:

$$D(L) = \mathrm{diag}(d_1(L),\ \ldots,\ d_n(L))$$

and

$$E\begin{bmatrix}\eta_t\\ \epsilon_t\end{bmatrix}[\eta_t \quad \epsilon_t'] = \Sigma = \mathrm{Diag}(\sigma^2_\eta,\sigma^2_{\epsilon_1},\ldots,\sigma^2_{\epsilon_n}) \ .$$

In addition, the scale of ΔC_t is identified by setting $\mathrm{var}(\eta_t)=1$. (This is a normalization with no substantive implications.)

A final identifying assumption is required to estimate the mean growth rate for $C_{t|t}$. This mean is calculated here as a weighted average of the growth rates of the constituent series. The weights are those implicitly used to construct $\Delta C_{t|t}$ from the original data series. That is, in this model $\Delta C_{t|t}$ can be written,

(7) $$\delta = W(L)Y_t$$

where W(L) is a 1×n lag polynomial vector. The mean of $\Delta C_{t|t}$ equals $W(1)'\mu_Y$, where $W(1)=\sum_{i=0}^{\infty}W_i$ and μ_Y denotes the mean of Y_t. This implies,

$$\delta = \phi(1)W(1)'\mu_Y .$$

Taken together, these assumptions provide sufficient identifying restrictions to estimate the unknown parameters of the model and to extract estimates of C_t.

B. State Space Representation

The first step towards estimating the model (4)-(6) is to cast it into a state space form so that the Kalman Filter can be used to evaluate the likelihood function. This formulation has two parts, the state equation and the measurement equation. The state equation describes the evolution of the unobserved state vector, which consists of ΔC_t, u_t, and their lags. The measurement equation relates the observed variables to the elements of the state vector.

The transition equation obtains by combining (5) and (6). Because one objective is to estimate the level of C_t using information up to time t, it is convenient to augment these equations at this point by the identity $C_{t-1} = \Delta C_{t-1} + C_{t-2}$. The transition equation for the state is thus given by:

$$\text{(8)} \qquad \begin{bmatrix} C_t^* \\ u_t^* \\ C_{t-1} \end{bmatrix} = \begin{bmatrix} \Phi^* & 0 & 0 \\ 0 & D^* & 0 \\ Z_c & 0 & 1 \end{bmatrix} \begin{bmatrix} C_{t-1}^* \\ u_{t-1}^* \\ C_{t-2} \end{bmatrix} + \begin{bmatrix} Z_c & 0 \\ 0 & Z_u \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \eta_t \\ \epsilon_t \end{bmatrix}$$

where:

$$C_t^* = [\Delta C_t \ \Delta C_{t-1} \cdots \Delta C_{t-p+1}]'$$

$$u_t^* = [u_t' \quad u_{t-1}' \cdots u_{t-k+1}']'$$

$$\Phi^* = \begin{bmatrix} \phi_1 \cdots \phi_{p-1} & \phi_p \\ I_{p-1} & 0 \end{bmatrix}$$

$$D^* = \begin{bmatrix} D_1 \cdots D_{k-1} & D_k \\ I_{n(k-1)} & 0 \end{bmatrix}$$

$$Z_c = [1 \quad 0_{1\times(p-1)}]$$

$$Z_u = [I_n \; 0_{n\times n(k-1)}]$$

and where I_n denotes the nxn identity matrix, $0_{n\times k}$ denotes a n×k matrix of zeros, and $D_i = \text{diag}\,(d_{1i},\ldots,d_{ni})$, where $d_j(L)=1-\sum_{i=1}^{k} d_{ji}L^i$.

The measurement equation is obtained by writing (4) as a linear combination of the state vector:

(9) $$Y_t = \beta + [\gamma Z_c \quad Z_u \quad 0]\begin{bmatrix} C_t^* \\ u_t^* \\ C_{t-1} \end{bmatrix} .$$

The system (8) and (9) can be rewritten more compactly in the standard form,

(10) $$\alpha_t = T_t\alpha_{t-1} + R\zeta_t$$

(11) $$Y_t = \beta + Z\alpha_t + \xi_t$$

where

$$\alpha_t = (C_t^{*\prime} \quad u_t^{*\prime} \quad C_{t-1})'$$

$$\zeta_t = (\eta_t \quad \epsilon_t')'$$

and where the matrices T_t, R and Z respectively denote the transition matrix matrix in (8), the selection matrix in (8), and the selection matrix in (9). The covariance matrix of ζ_t is $E\zeta_t\zeta_t' = \Sigma$. For generality, a measurement error term ξ_t (assumed uncorrelated with ζ_t) has been added to the measurement equation (11), and the transition matrix T_t is allowed to vary over time. In the empirical work below, however, the measurement noise is set to zero and the time invariant transition matrix in (8) is used.[3]

C. Estimation

The Kalman Filter is a well-known way to compute the Gaussian likelihood function for a trial set of parameters; for a discussion, see Harvey (1981). The filter recursively constructs minimum mean square error (MMSE) estimates of the unobserved state vector, given observations on y_t. The filter consists of two sets of equations, the prediction and updating equations. Let $\alpha_{t|\tau}$ denote the estimate of α_t based on $(y_1,\ldots,y_\tau)$, let $E[\xi_t\xi_t']=H$, and recall that $E[\zeta_t\zeta_t']=\Sigma$. Also, let $P_{t|\tau}=E[(\alpha_{t|\tau}-\alpha_t)(\alpha_{t|\tau}-\alpha_t)']$. With this notation, the prediction equations of the Kalman filter are:

$$\alpha_{t|t-1} = T_t\alpha_{t-1|t-1} \tag{12}$$

$$P_{t|t-1} = T_t P_{t-1|t-1} T_t' + R\Sigma R' \quad . \tag{13}$$

The forecast of Y_t at time t-1 is $Y_{t|t-1}=\beta+Z\alpha_{t|t-1}$, and the forecast error is $\nu_t=Y_t-\beta-Z\alpha_{t|t-1}$. The updating equations of the filter are:

(14) $$\alpha_{t|t} = \alpha_{t|t-1} + P_{t|t-1}Z'F_t^{-1}\nu_t$$

(15) $$P_{t|t} = P_{t|t-1} - P_{t|t-1}Z'F_t^{-1}ZP_{t|t-1}$$

where $$F_t = E[\nu_t\nu_t'] = ZP_{t|t-1}Z' + H .$$

The Kalman filter equations (12)-(15) permit recursive calculation of the predicted state vector, $\alpha_{t|t-1}$, and of the covariance matrix of this estimate, $P_{t|t-1}$, given the assumed parameters in T_t, R, Σ, H, and Z, and given initial values for $\alpha_{t|t}$ and $P_{t|t}$. For exact maximum likelihood estimation, these initial values are taken to be the unconditional expectation of α_t and its covariance matrix, $E[(\alpha_t-E\alpha_t)(\alpha_t-E\alpha_t)']$; that is, $\alpha_{0|0}=0$ and $P_{0|0}=\sum_{j=0}^{\infty}T_{t-j}^{j}\Sigma T_{t-j}^{j\prime}$. Alternatively, one could set $P_{0|0}$ to an arbitrary constant matrix. In this case, the estimates are asymptotically equivalent to maximum likelihood.

The Gaussian log likelihood is then computed (up to an additive constant) as:

(16) $$\mathcal{L} = -\tfrac{1}{2}\sum_{t=1}^{T}\nu_t'F_t^{-1}\nu_t - \tfrac{1}{2}\sum_{t=1}^{T}\ln(\det(F_t)) .$$

The Gaussian maximum likelihood estimates of the parameters are found by maximizing $\mathcal{L}$ over the parameter space.

D. Construction of the Leading Index and Weights (W_i)

The alternative index of coincident indicators from the single-index model is the MMSE estimator of C_t constructed using data on the coincident variables available through time t. This is denoted by $C_{t|t}$. The Kalman Filter produces the MMSE estimator $\alpha_{t|t}$ of the state vector given $(Y_1,...,Y_t)$. In the notation of (8), the alternative index of coincident indicators is $C_{t|t}=(Z_c\ 0\ 0\ 1)\alpha_{t|t}$. The weights implicitly used by the Kalman Filter to construct $C_{t|t}$ from the coincident variables can be calculated by computing the response of $C_{t|t}$ to unit impulses in each of the observed coincident variables. The variance of $C_{t|t}$ is $(Z_c\ 0\ 0\ 1)P_{t|t}(Z_c\ 0\ 0\ 1)'$.

It is worth noting that this framework also permits the calculation of retrospective estimates of the state of the economy, $C_{t|T}$, and more generally $\alpha_{t|T}$. Estimates of α_t based on the entire sample are computed using the Kalman smoother (see Harvey [1981]).

The weighting polynomial W(L) in (7) can be obtained directly from the Kalman filter matrices. Because $\Delta C_{t|t}=e_1'\alpha_{t|t}$, where $e_1=(1\ 0\ \cdots\ 0)'$, the problem of finding the weights implicitly used to construct $\Delta C_{t|t}$ is a special case of the problem of finding the corresponding weights for $\alpha_{t|t}$. The estimate $\alpha_{t|t}$ computed by the Kalman filter is linear in current and past observations on Y_t. By substituting the relationship $\nu_t=Y_t-(\beta+Z\alpha_{t|t-1})$ into (14) and then using (12), one obtains:

$$(17)\qquad \alpha_{t|t} = (I-G_tZ)T_t\alpha_{t-1|t-1} + G_tY_t - G_t\beta$$

where $G_t=P_{t|t-1}Z'F_t^{-1}$ is the Kalman gain. When the data are expressed as deviations from their means (as is done in the empirical estimation below), β

is "concentrated out" of the likelihood. In addition, when T_t is time invariant (so that $T_t = T^*$), G_t converges nonstochastically to the steady-state Kalman gain, G^*. Under these conditions, (17) can be rewritten,

$$(18) \qquad (I-KL)\alpha_{t|t} = G^* Y_t$$

where $K=(I-G^*Z)T^*$. The weights W(L) are obtained by inverting (I-KL) in (18) and selecting the first row of the resulting infinite order moving average:

$$(19) \qquad \Delta C_{t|t} = e_1' \sum_{j=0}^{\infty} K^j G^* Y_{t-j} \ .$$

3. Empirical Results for the Coincident Model

This section presents estimates of the single-index model using four monthly time series for the United States from 1959:2 to 1987:12. The series are those used to construct the DOC coincident series: industrial production (IP), total personal income less transfer payments in 1982 dollars (GMYXP8), total manufacturing and trade sales in 1982 dollars (MT82), and employees on nonagricultural payrolls (LPNAG). The data were obtained from the August, 1988 release of Citibase. Throughout, we adopt the Citibase mnemonics for the variables when applicable.

A. Preliminary Data Analysis

The first step in specifying the model is to test for whether the series are integrated and, if they are, whether they are cointegrated. For each of the coincident indicators, Dickey and Fuller's (1979) test for a unit root (against the alternative that the series are stationary, perhaps around a linear time trend) was unable to reject (at the 10% level) the hypothesis that the series are integrated. The subsequent application of the Stock-Watson (1986) test of the null hypothesis that the four series are not cointegrated against the alternative of cointegration failed to reject at the 66% significance level. Thus these tests provided no evidence against the hypothesis that each series is integrated but they are not cointegrated. We therefore estimated the model (3)-(6) using for Y_t the first difference of the logarithm of each of the coincident series, standardized to have zero mean and unit variance.

The single-index model imposes the restriction that all the comovements in the series arise from a single source. Tests of this restriction, against the hypothesis that the coincident indicators have a spectral density matrix that is finite and nonsingular but otherwise unrestricted, were implemented by Sargent and Sims (1977). Their test examines the implication that the spectral density matrix of Y_t, averaged over any frequency bands, will have a factor structure in the sense of conventional factor analysis; thus the dynamic single-index restrictions may be tested by testing the "single factor" restrictions for a set of these bands and then aggregating the results. Specifically, since ΔC_t and Δu_t are by assumption uncorrelated at all leads and lags, (4) implies that $S_Y(\omega) = \gamma S_{\Delta C}(\omega)\gamma' + S_u(\omega)$, where $S_Y(\omega)$ denotes the spectral density matrix of Y_t at frequency ω, etc. Because $S_{\Delta C}(\omega)$ is a scalar and $S_u(\omega)$ is diagonal, this implies a testable restriction on $S_Y(\omega)$. Performing this test for the coincident indicator model over six equally-spaced bands constructed using Y_t (where the averaged matrix periodogram provides the unconstrained estimate of the spectrum) provides little evidence against the dynamic single-index structure: the χ^2_{30} test statistic is 19.8, having a p-value of 92%.

B. Maximum Likelihood Estimates

The parameters of two single-index models were estimated using IP, GMYXP8, MT82, and LPNAG over the period 1959:2-1987:12. In both models, a second order autoregressive specification was adopted for ΔC_t, so that p=2. In the first, the errors u_t are modeled as an AR(1) (k=1); in the second, they are modeled as an AR(2) (k=2). The log likelihood for the AR(1) model is 327.77, and for the AR(2) model is 341.38. A likelihood ratio test easily rejects the hypothesis that the additional four autoregressive parameters are zero. We therefore adopt the AR(2) specification henceforth.

The maximum likelihood estimates of the parameters of the single-index model are presented in Table 1. The negative estimates of d_{1i} for IP and MT82 indicate that the idiosyncratic component of these series exhibits negative serial correlation, although the idiosyncratic component of LPNAG exhibits substantial positive serial correlation. The estimated model for the unobserved component exhibits substantial first order -- but limited second-order -- dependence, with roots of (.60,-.05).

Statistics that examine the fit of the single-index model are presented in Table 2. The tests assess whether the disturbances in the observed variables are predictable: if the estimated model is correctly specified, they should be serially uncorrelated. The results suggest satisfactory specifications for the IP, GMYXP8, and MT82 equations. However, the disturbance in LPNAG is forecastable by each lagged disturbance and variable in the model, indicating

Table 1

Estimated Single-Index Model of Coincident Variables

Parameter	------- Coincident Variable -----			
	IP	GMYXP8	MT82	LPNAG
γ_i	0.717	0.521	0.470	0.602
	(0.037)	(0.044)	(0.030)	(0.041)
d_{1i}	-0.040	-0.087	-0.414	0.108
	(0.091)	(0.042)	(0.052)	(0.050)
d_{2i}	-0.137	0.154	-0.206	0.448
	(0.083)	(0.049)	(0.059)	(0.068)
σ_i ($\times 10^{-2}$)	0.488	0.769	0.735	0.540
	(0.035)	(0.026)	(0.030)	(0.030)

$$\Delta C_t = \underset{(0.062)}{0.545}\Delta C_{t-1} + \underset{(0.065)}{0.032}\Delta C_{t-2} + \eta_t$$

$\mathcal{L} = 341.38$

Notes: The estimation period is 1959:2-1983:12. The parameters were estimated by Gaussian maximum likelihood as described in the text. The parameters are $\gamma=(\gamma_1,\ldots,\gamma_4)$, $D(L)=\mathrm{diag}(d_1(L),\ldots,d_4(L))$ with $d_i(L)=1-d_{i1}L-d_{i2}L^2$, and $\Sigma=\mathrm{diag}(\sigma_1^2,\ldots,\sigma_4^2)$. Asymptotic standard errors (computed numerically) appear in parentheses.

Table 2

Marginal Significance Levels of
Diagnostic Tests for Single-Index Model

p-values

Regressor	------- Dependent Variable ------			
	e_{IP}	e_{GMYXP8}	e_{MT82}	e_{LPNAG}
e_{IP}	.625	.445	.063	.014
e_{GMYXP8}	.320	.986	.786	.034
e_{MT82}	.198	.952	.810	.004
e_{LPNAG}	.359	.790	.163	.000
$\Delta\ln(IP)$	.593	.464	.196	.002
$\Delta\ln(GMYXP8)$	.366	.986	.905	.008
$\Delta\ln(MT82)$	.219	.875	.628	.000
$\Delta\ln(LPNAG)$	.241	.556	.189	.000

Notes: The entries in the table are p-values from the regression of e_y against a constant and six lags of the indicated regressor; the p-values correspond to the usual F-test of the hypothesis that the coefficients on these six lags are zero. No attempt was made to correct the test statistic or the distribution for degrees of freedom, other than the usual correction for the number of regressors. The series e_y denotes the one-step ahead forecast errors from the single-index model. That is, e_y is $y_t-y_{t|t-1}$, where $y_{t|t-1}$ is computed using the Kalman filter applied to the estimated model reported in Table 1.

misspecification of the LPNAG equation. A possible source of this misspecification is that LPNAG is not an exactly coincident variable, but slightly lags the unobserved factor. This could be investigated by including lags of C_t in the equation for LPNAG in (4). Alternatively, one might change the dynamic specification for u_{4t}. We do not, however, pursue these options for two reasons. First, with the exception of the LPNAG equation, the overall fit of the model appears to be good. Second, and more importantly, our primary objective is to see whether a purely coincident single-index model can rationalize the DOC coincident index; adopting a mixed coincident/lagging single-index specification would defeat this purpose. We therefore proceed with the coincident single-index model of Table 1, but raise the apparent misspecification of the LPNAG equation as an issue for future research.

C. Comparison Between $C_{t|t}$ and the DOC Series

We now examine the relation between the contemporaneous estimate $C_{t|t}$ of C_t obtained using the coincident single-index model and the Index of Coincident Indicators published by the Department of Commerce, henceforth referred to as C_t^{DOC}. The summary statistics reported in Table 3 indicate that these series are highly correlated, but that the standard deviation of the growth rate of C_t^{DOC} exceeds that of $C_{t|t}$ by 80%. This is a consequence of how C_t^{DOC} is constructed: the weights on the deviations of the constituent series from their means sum to 1.8, while the weights implicitly used to construct $\Delta C_{t|t}$ sum to 1. This difference affects the graphical presentation of the series in levels, but of course not the correlations or other functions of centered moments estimated using their growth rates.

To facilitate a graphical comparison, we calculated a series $C^*_{t|t}$ by scaling $\Delta C_{t|t}$ to have the same variance as the growth in C_t^{DOC} and to have

Table 3

Comparison of Growth Rates of $C_{t|t}$ and C_t^{DOC}: Summary Statistics

Growth Rates (on an annual basis) of:	Sample Mean	Sample Standard Deviation
Commerce Series	3.53%	8.98%
$C_{t\|t}$	2.96%	4.80%

Contemporaneous correlation between C_t^{DOC} and $C_{t|t}$: 0.936

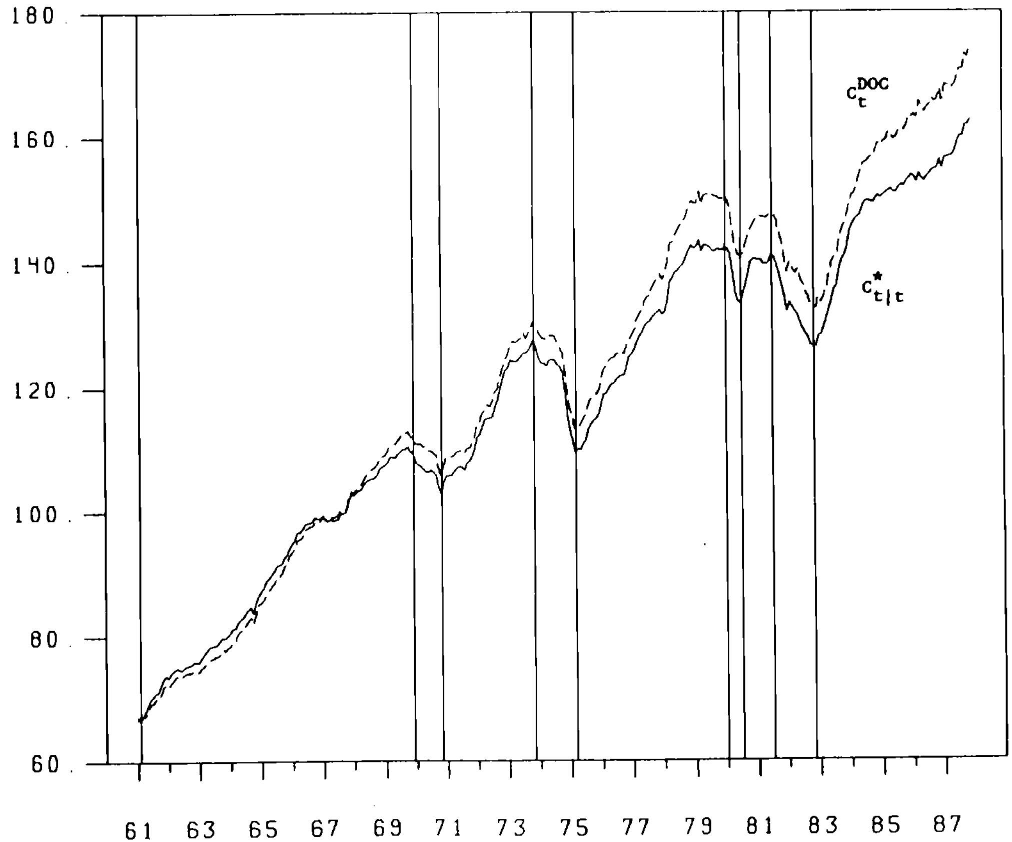

Figure 1. $C^*_{t|t}$ and C_t^{DOC}, 1961-1987.

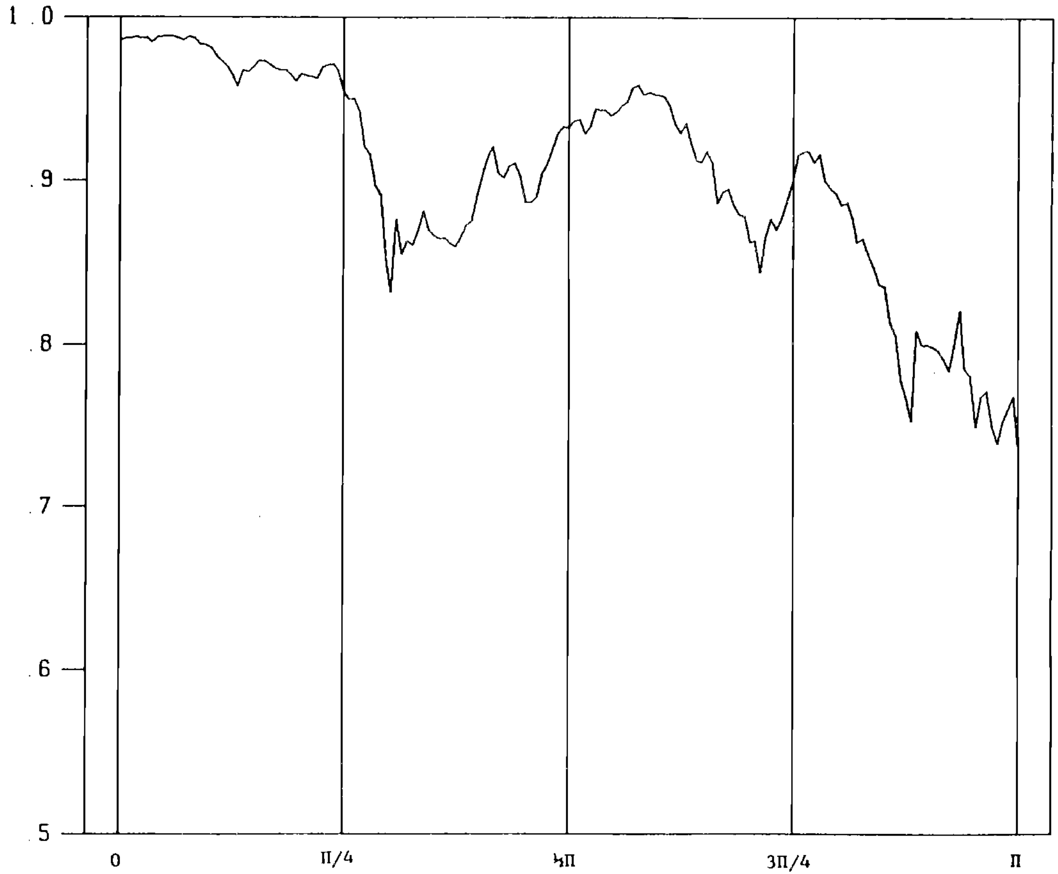

Figure 2. Coherence between $\Delta C_{t|t}$ and $\Delta \ln(C_t^{DOC})$

the same mean as $\Delta C_{t|t}$. This modified growth rates series was then used to construct the levels series $C^*_{t|t}$, indexed to equal 100 in January, 1967. The series C^{DOC}_t and $C^*_{t|t}$, are plotted in Figure 1. Although the mean growth rates differ, the two series exhibit a striking similarly. One of the few noticeable differences between the two series is the somewhat slower average growth of $C^*_{t|t}$ from 1984 to 1987.

A comparison of the weights used in the construction of the growth rates of C^{DOC}_t and $C^*_{t|t}$ provides an additional indication of their similarity. These weights, normalized to add to one, are presented in Table 4. The two sets of weights are generally similar. For example, contemporaneous values receive most of the weight in $C_{t|t}$ (and 100% in C^{DOC}_t). The major exception is that a second lag of LPNAG receives a substantial negative weight.

A final measure of the relation between $C_{t|t}$ and C^{DOC}_t is the coherence between their growth rates; this is plotted in Figure 2. The very high estimated coherence at low frequencies (in excess of 95% for periods over two years) indicates that the two series are very similar at horizons associated with the business cycle.

In summary, these measures all suggest that the coincident index $C_{t|t}$ estimated using single-index model agrees closely with the current DOC CEI, especially at business cycle frequencies.

4. Forecasts of the Coincident Index Using Leading Variables

This section explores one approach to forecasting the coincident series $C_{t|t}$. Rather than considering $C_{t|t}$ directly, we focus on forecasting the

Table 4

Weights on coincident variables used in constructing C_t^{DOC} and $C_{t|t}$

A. C_t^{DOC}

Lag	IP	GMYXP8	MT82	LPNAG
0	.156	.227	.138	.479
1	.000	.000	.000	.000
2	.000	.000	.000	.000
3	.000	.000	.000	.000
4	.000	.000	.000	.000

B. $C_{t|t}$

Lag	IP	GMYXP8	MT82	LPNAG
0	.271	.163	.072	.668
1	.031	.026	.036	-.029
2	.052	-.016	.021	-.271
3	.005	-.000	.003	-.019
4	.003	-.001	.001	-.014
>4	.000	.000	.000	-.002
Total:	.361	.173	.134	.332

Notes: The weights are normalized so that they add to one. The weights implicitly used to construct the C_t^{DOC} computed by regressing the growth rate of the Commerce series on contemporaneous values of the growth rates of the four series used in its construction: industrial production (IP), real personal income (GMYXP8), real manufacturing and trade sales (MT82), and employment at nonagricultural establishments (LPNAG). Thus the weights for all lags other than the contemporaneous variables are zero by construction. The R^2 of this regression, estimated from 61:1 to 87:12, is .937. Reasons why the R^2 would be less than one include rounding error and minor differences between our series and those used by the DOC to construct the index. The weights implicitly used to construct $C_{t|t}$ were obtained by computing the responses of $C_{t|t}$ to unit impulses to each of the four series using the Kalman filter.

growth of $C_{t|t}$ over six months, denoted by $f_t(6)$. The strategy is to construct a parsimoniously parameterized system of autoregressive equations, where the parameterization is suggested by the data. The resulting "base" model is then used to assess the marginal predictive content of additional candidate leading variables.

A. Autoregressive Systems to Forecast $f_t(6)$

Summary statistics for several autoregressive systems are presented in Table 5. In addition to $\Delta C_{t|t}$, each system includes five key leading variables: manufacturing and trade inventories (IVMT82), manufacturers' unfilled orders (MDU82), housing starts (HSBP), the yield on a constant-maturity portfolio of 10-year treasury bonds, and the spread between the 10 year bond yield and the interest rate on 90 day T-bills. Note that $\Delta C_{t|t}$, not $f_t(6)$, enters the autoregressions. The six-month forecast of $C_{t|t}$, $\hat{f}_t(6)$, is computed from the estimated systems.

The classical VAR(12), while delivering the highest R^2 between $f_t(6)$ and $\hat{f}_t(6)$, has many parameters and thus might be expected to be unstable. The Bayesian VAR's reported in Table 5 impose "smoothness priors", in effect imposing restrictions that might mitigate this problem. Unfortunately, the predictive performance of the Bayesian VAR's deteriorates substantially when the tightness of the prior is increased. As an alternative, Panel D reports results for a set of models that are relatively parsimoniously parameterized, but that do not result in a marked deterioration of six-step ahead forecasting performance. In these, the equations for IVMT82, MDU82, HSBP, FYGT10, and FYGT10-FYGM3 each include one lag of each of the six variables, while the $\Delta C_{t|t}$ equation includes the indicated number of lags (p_{1j}) of the six

Table 5

Performance of Various Systems for Forecasting $f_t(6)$

Variables: $C_{t|t}$, IVMT82, MDU82, HSBP, FYGT10, FYGT10-FYGM3

Model	R^2 between $\hat{f}_t(6)$ and $f_t(6)$ 61:1-87:6	61:1-79:9	79:10-87:6
A. Classical VAR's			
1. VAR(6)	.578	.585	.565
2. VAR(12)	.654	.633	.673
B. Bayesian VAR's			
3. BVAR(12), γ=.2	.620	.614	.626
4. BVAR(12), γ=.1	.577	.585	.567
C. VAR's with ΔFYGT10			
5. VAR(12)	.538	.534	.632
6. BVAR(12), γ=.1	.473	.475	.542
D. Mixed systems:			
$[p_{ii}, i>1; p_{ij}, i\neq j>1; p_{11}; (p_{12}, p_{13}, p_{14}); (p_{15}, p_{16})]$			
7. (1,1,12,12,12)	.633	.630	.636
8. (1,1,4,12,12)	.615	.601	.634
9. (1,1,4,12,8)	.607	.590	.634
10. (1,1,4,8,8)	.587	.597	.573

Notes: $f_t(6)=C_{t+6|t+6}-C_{t|t}$, where $C_{t|t}$ is estimated from the single-index model reported in Table 1; $\hat{f}_t(6)$ denotes the six-month ahead forecast of $C_{t|t}$ computed using the relevant estimated system, i.e. $\hat{f}_t(6)=\hat{C}_{t+6|t+6}-C_{t|t}$. All systems were estimated in RATS using the six variables listed. The γ parameter in the Bayesian VAR's represent different degrees of tightness in the Bayesian prior, with the lower value corresponding to a more tight prior. The estimation period was 1961:1-1987:12; the R^2 measures were computed for different subsamples using the $\hat{f}_t(6)$ generated by the autoregressive model estimated over the entire period. The mixed systems impose many zero restrictions; the p_{ij} notation in the specification of the mixed systems refers to the number of lags of the j-th variable in the i-th equation in the autoregressive system. $C_{t|t}$, IVMT82, and MDU82 appear in growth rates and HSBP and FYGT10-FYGM3 appear in levels. In panels A, B, and D, FYGT10 appears in levels, while in panel C it appears in first differences.

variables. Of these, model 9 was selected as the "base" model for subsequent analysis.

B. An Examination of Additional Variables

This base model provides a framework for assessing the marginal predictive content of alternative leading variables. This is done by including trial variables, one at a time, as additional leading variables in the mixed autoregressive system. The results for selected variables are reported in Table 6. (The variable definitions are provided in Appendix A.)

The strongest evidence of additional predictive content comes from yields on various financial instruments. Theory suggests that bond and stock prices will incorporate expectations of various aspects of future economic performance and so should be useful leading indicators. Yet it is perhaps surprising that these variables provide additional useful information, because the base model already includes two interest rates (FYGT10 and FYGM3) and an interest-sensitive series, HSBP. These high R^2 values should be interpreted cautiously, however. For the interest rates, at least, the two largest R^2's drop in the 1980's, indicating potential instability of these series as useful leading indicators. It is also interesting to note that stock prices help to forecast $f_t(6)$ not by predicting $\Delta C_{t|t}$ directly, but rather by helping to forecast the other variables in the system which in turn are useful in forecasting $\Delta C_{t|t}$.

Among selected BCD leading indicators (panel A), only IVPAC makes a statistically strong contribution to the $\Delta C_{t|t}$ equation; but including IVPAC actually reduces the six-step ahead forecasting performance. Measures of productivity, prices, exchange rates, and services make only modest

Table 6

Tests of the marginal predictive value of alternative leading variables

Series	Transformation	F-tests: p-values lags 1-12	F-tests: p-values lags 7-12	R^2 between $\hat{f}_t(6)$ and $f_t(6)$ 61:1-87:12	R^2 between $\hat{f}_t(6)$ and $f_t(6)$ 61:1-79:9	R^2 between $\hat{f}_t(6)$ and $f_t(6)$ 79:10-87:12
Base model		--	--	.607	.590	.634
A. Selected Leading Indicators						
bus	Δln(x)	.093	.699	.624	.622	.628
lhu5	Δln(x)	.091	.389	.624	.609	.643
condo9	Δln(x)	.642	.406	.618	.603	.641
hsfr	Δln(x)	.719	.502	.617	.612	.626
ipi	Δln(x)	.384	.569	.621	.604	.650
ipmfg	Δln(x)	.345	.279	.626	.623	.640
ipcd	Δln(x)	.204	.648	.642	.642	.636
gmcd82	Δln(x)	.612	.430	.615	.592	.655
ivm28	Δln(x)	.693	.945	.645	.648	.634
mocm82	Δln(x)	.868	.876	.615	.594	.649
mpcon8	Δln(x)	.831	.544	.603	.602	.606
ivpac	Δln(x)	.001	.212	.579	.568	.596
rcar6d	Δln(x)	.949	.667	.610	.594	.634
B. Productivity						
loutbm	Δln(x)	.192	.367	.599	.558	.633
lboutum	Δln(x)	.945	.722	.604	.591	.631
C. Prices						
pw	Δln(x)	.689	.904	.629	.628	.619
pwfc	Δln(x)	.712	.637	.611	.595	.636
ppsmc	none	.216	.057	.632	.632	.637
punew	Δln(x)	.104	.136	.605	.600	.624
pzunew	Δln(x)	.180	.344	.610	.607	.620
D. Exchange Rates and trade						
exrwt1	Δln(x)	.746	.575	.613	.598	.636
exrwt2	Δln(x)	.639	.426	.616	.599	.642
exnwt1	Δln(x)	.694	.461	.616	.605	.640
exnwt2	Δln(x)	.656	.359	.618	.604	.643
tb1	none	.459	.207	.624	.599	.662
D. Services						
gmws	Δln(x)	.371	.192	.619	.607	.657
lpsp	Δln(x)	.398	.647	.611	.613	.621
lpgov	Δln(x)	.166	.119	.621	.615	.641
lpspng	Δln(x)	.884	.934	.612	.608	.627

Table 6 (continued)

Series	Transformation	F-tests: p-values lags 1-12	F-tests: p-values lags 7-12	R^2 between $\hat{f}_t(6)$ and $f_t(6)$ 61:1-87:12	R^2 between $\hat{f}_t(6)$ and $f_t(6)$ 61:1-79:9	R^2 between $\hat{f}_t(6)$ and $f_t(6)$ 79:10-87:12
F. Money and Credit						
fmld82	$\Delta\ln(x)$	1.000	.997	.607	.593	.632
fm2d82	$\Delta\ln(x)$	.873	.977	.624	.613	.629
fmbase	$\Delta\ln(x)$	.110	.045	.641	.637	.687
fcln82	$\Delta\ln(x)$	.203	.815	.608	.575	.674
fcbcuc	$\Delta\ln(x)$	.701	.963	.648	.650	.622
fclbmc	$\Delta\ln(x)$	.125	.037	.633	.772	.238
cci30m	$\Delta\ln(x)$	.567	.537	.642	.640	.641
G. Stock prices and volume						
fspcom	$\Delta\ln(x)$	.317	.393	.673	.649	.696
fsdj	$\Delta\ln(x)$	.420	.384	.645	.630	.659
fspin	$\Delta\ln(x)$	.300	.308	.672	.647	.697
fsvol	$\Delta\ln(x)$	.171	.112	.639	.614	.671
H. Nominal and ex-post real interest rates						
fyff	none	.077	.986	.680	.703	.616
fygtl	none	.436	.668	.636	.594	.699
fycp	none	.007	.756	.739	.751	.691
fyffr	none	.037	.982	.614	.595	.644
fygm3r	none	.188	.923	.623	.618	.626
fygtlr	none	.086	.972	.615	.599	.639
fygt10r	none	.571	.641	.612	.585	.664

Notes: The base model in the first line of this table is the mixed VAR, model 9 in Table 5. For each of the variables, the base model was augmented by the trial variable, with 12 lags of the trial variable in the $C_{t|t}$ equation and 1 lag of the trial variable entering each of the remaining six equations. The p-values reported in the third and fourth columns refer to the F-tests of the hypothesis that the coefficients on the indicated lags of the trial variable in the $C_{t|t}$ equation are zero. The R^2 measures in the final two columns are described in the notes to Table 5.

contributions. Two of the measures of money and credit -- FMBASE and FCBCUC -- made substantial improvements in the R^2. However, money and credit variables must be viewed with suspicion as stable leading indicators because of recent changes in financial markets (see Friedman [1988] for a discussion). This is highlighted by the very poor performance of the six-month ahead forecasts based on FCLBMC in the 1980's.

In summary, the leading variables in the base model, along with some of the variables in Table 6, indicate that approximately two-thirds of the variance of the six-month ahead growth in $C_{t|t}$ can be forecasted using restricted autoregressive systems. The next step, left for further research, is to assess the stability of these systems.

5. Forecasts of Recessions Using Leading Variables

This section presents an initial exploration of the possibility of predicting whether six months hence the economy will be in a NBER-dated recession. The approach taken here is to estimate several binary logit models. The dependent variable is R_{t+6}, where $R_t=1$ if the economy is in a NBER-dated recession in month t, and $R_t=0$ if not. The predictive variables are the forecasts $\hat{f}_t(1),\ldots,\hat{f}_t(6)$ calculated from an eight-variable mixed autoregressive system with $C_{t|t}$, IVMT82, MDU82, HSBP, FYGT10, FYGT10-FYGM3, FYCP, and IPCD. In addition, individual leading variables were included, both singly and as a group.

The logit models were estimated by maximum likelihood. It is important to recognize that this likelihood is misspecified: because the dependent variable refers to an event six months hence, the errors in the latent variable equation (and the six-step ahead prediction errors) will in general have a moving average structure. Thus this procedure should be viewed simply as providing a convenient functional form for exploring the predictability of R_{t+6}.

The results are reported in Table 7. The first noteworthy result is that the one- through six-step ahead forecasts of $C_{t|t}$, produced by the eight-variable mixed autoregression, by themselves produce a substantial reduction in the unexplained variance of R_t. In addition, these results suggest that expansions can be forecasted accurately, but that the probability of correctly forecasting a recession six months hence is the range of two thirds. Considerable gains can be made by incorporating variables in addition to $\hat{f}_t(1),\ldots,\hat{f}_t(6)$; financial variables seem particularly useful in this regard.

The predictions $R_{t+6|t}$ of the six-month ahead NBER forecasts, computed using model 15 in Table 7, are plotted in Figure 3. The dating convention in this figure is that $R_{t|t-6}$ is plotted; for example, in the figure the July, 1982 probability is the forecast of the binary July, 1982 recession/expansion variable, made using data through January 1982. The actual NBER-dated recessions are indicated by vertical lines. This plot bears out the false positive and false negative rates presented in the final two columns of Table 7. On the one hand, false recession forecasts occur relatively rarely, with the largest false recession forecasts occurring during 1967 and throughout 1979. On the other hand, the six-month-ahead recession probabilities rarely approach one when in fact a recession does occur. For example, the six-month ahead probability of a recession incorrectly drops below 50% during the middle of the 1970 contraction.

Table 7

Logit models to predict NBER-dated recessions at a six month horizon

Variables in Base Model: constant, $\hat{f}_t(1)$, $\hat{f}_t(2)$, ..., $\hat{f}_t(6)$

	- Additional Predictive Variables - Series	no. lags	R^2	False Positive Rate	False Negative Rate
1.	--	-	.448	.090	.452
2.	MDU82	5	.456	.089	.447
3.	IVMT82	5	.491	.084	.423
4.	HSBP	5	.513	.080	.403
5.	FYGT10	5	.486	.084	.423
6.	FSPRD	5	.551	.074	.372
7.	FYCP	5	.479	.084	.420
8.	IPCD	5	.530	.076	.380
9.	EXNWT2	5	.453	.089	.448
10.	LHU5	5	.482	.084	.422
11.	LBOUTU	5	.469	.087	.436
12.	FSPCOM	5	.497	.083	.416
13.	MDU82, IVMT82, HSBP, FYGT10, FSPRD, FYCP, IPCD	1	.597	.067	.334
14.	FYGT10, FSPRD, FYCP, FSPCOM	1	.551	.073	.366
15.	FYGT10, FSPRD, FYCP, FSPCOM	3	.613	.062	.309

Notes: The estimation period is 1961:1-1987:12. The dependent variable in the logit regressions has a value of 1 if there is a NBER-dated recession six months in the future, and has a value of 0 if six months hence there is a NBER-dated expansion. The mean value of this variable over the estimation period is .167. Model 1 is the "base model", with only $(1,\hat{f}_t(1),\hat{f}_t(2),\ldots,\hat{f}_t(6))$ as regressors. Contemporaneous value (and the number of lags indicated in the second column) of the additional predictive variables were included as well in the remaining models. The R^2 is computed using the actual (0/1) recession probabilities and their six-month ahead forecast computed by the logit model. The second-to-last column ("false positive rate") presents the average fraction of times that a recession is forecast when in fact no recession occurs, and the final column ("false negative rate") presents the average fraction of times that no recession is forecasted when in fact a recession occurs.

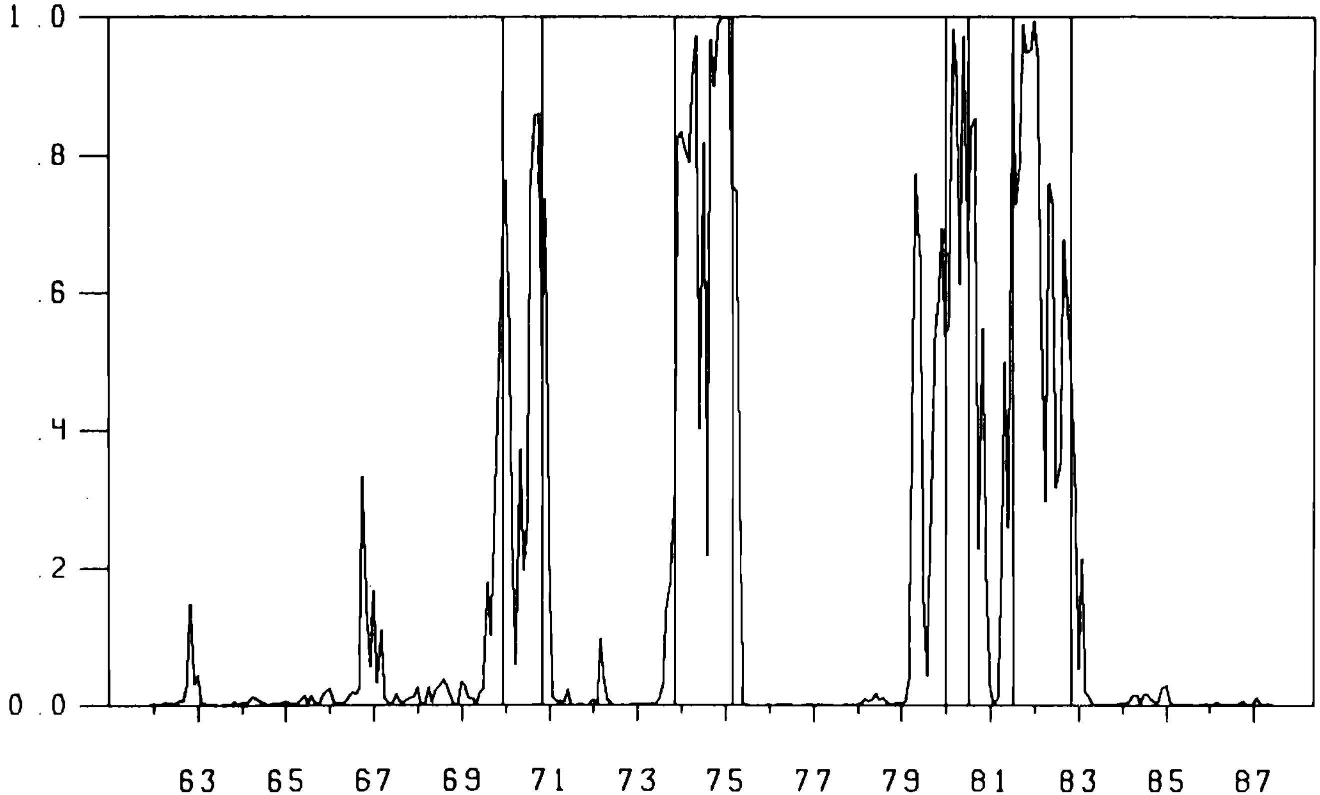

Figure 3. NBER-dated recessions and forecasts of their probability of occuring, made six months prior using model 15 of Table 7.

6. Conclusions

The single-index model provides an explicit probability model for the definition and estimation of an alternative index of coincident indicators. The empirical model produces a coincident index that is strikingly similar to the index currently computed by the Department of Commerce, particularly at the low frequencies associated with business cycles. The main evidence of misspecification in the empirical model appears in the equation for the employment series, LPNAG. One possible explanation, to be explored in future research, is that employment is a lagging rather than an exactly coincident series.

The forecasting exercises of Sections 4 and 5 indicate that time series models which incorporate leading variables can provide useful forecasts of the growth of the coincident index, and of a variable that indicates whether the economy is in a recession or expansion six months hence. This is unsurprising in the sense that these leading variables are so categorized because of their tendency to move in advance of the coincident index. The main contribution of these empirical investigations is rather to provide some specific models with which to make these forecasts, and some statistical measures of the within-sample forecast quality. A noteworthy finding from this investigation is that financial prices and yields appear to have greater predictive value than do measures of real output, real inputs, or prices of foreign or domestic goods.

Footnotes

1. Most modern research on the forecasting potential of the index of leading indicators has focused on its ability to forecast not the reference cycle, but some observable series such as industrial production or unemployment (e.g. Stekler and Schepsman (1973), Vaccara and Zarnowitz [1978], Sargent [1979], Auerbach [1982], and Koch and Raasche [1988]). Our perspective is closer to that underlying the work of Diebold and Rudebusch (1987) and Hamilton (1987). For a historical review of the development of the leading indicators, see Moore (1979).

2. A different way to make this point is that we are examining comovements among the first differences of the coincident variables at frequencies other than zero. Were the common factor ΔC_t the only source of power at frequency zero in the spectra of the first differences, the spectral density matrix of the first differences would be singular at frequency zero and the series would be cointegrated. Harvey, Fernandez-Macho, and Stock (1987) discuss modeling strategies for unobserved- component models with cointegrated variables.

3. The state space representation (8) and (9) is not unique. In practice, it is computationally more efficient to work with a lower-dimensional state vector. This can be achieved by filtering Y_t, $\gamma\Delta C_t$, and u_t in (4) by D(L) and treating ϵ_t as a measurement error. The resulting state vector has dimension p+1.

References

Auerbach, A.J. (1982), "The Index of Leading Indicators: 'Measurement Without Theory,' Thirty-five Years Later," Review of Economics and Statistics 64, no. 4, 589-595.

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

Dickey, D.A., and W.A. Fuller (1979), "Distribution of the Estimators for Autoregressive Time Series With a Unit Root," Journal of the American Statistical Society 74, no. 366, 427-431.

Diebold, F.X., and G.D. Rudebusch (1987), "Scoring the Leading Indicators," manuscript, Division of Research and Statistics, Federal Reserve Board.

Engle, R.F., and C.W.J. Granger (1987), "Co-Integration and Error Correction: Representation, Estimation and Testing," Econometrica 55, 251-276.

Friedman, B.M. (1988), "Monetary Policy Without Quantity Variables," NBER Working Paper no. 2552.

Hamilton, J.D. (1987), "A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle," manuscript, University of Virginia.

Harvey, A.C. (1981), Time Series Models. Oxford: Philip Allan.

Harvey, A.C., F.J. Fernandez-Macho, and J.H. Stock (1987), "Forecasting and Interpolation Using Vector Autoregressions with Common Trends," Annales d'Economie et de Statistique, #6-7, 279-288.

Hymans, S., "On the Use of Leading Indicators to Predict Cyclical Turning Points," Brookings Papers on Economic Activity, Vol. 2 (1973), 339-384.

Kling, J.L. (1987), "Predicting the Turning Points of Business and Economic Time Series," Journal of Business 60, no. 2, 201-238.

Koch, P.D., and R.H. Raasche (1988), "An Examination of the Commerce Department Leading-Indicator Approach," Journal of Business and Economic Statistics 6, no. 2, 167-187.

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

Moore, G.H. (1979) "The Forty-Second Anniversary of the Leading Indicators," in William Fellner, ed., Contemporary Economic Problems, 1979. Washington, D.C.: American Enterprise Institute, 1979; reprinted in Moore,G.H., Business Cycles, Inflation, and Forecasting, second edition, Cambridge, Mass: NBER, 1983.

Neftci, S.N. (1982), "Optimal Prediction of Cyclical Downturns," Journal of Economic Dynamics and Control 4, 225-241.

Nelson, C.R., and C.I. Plosser (1982), "Trends and Random Walks in Macroeconomic Time Series," Journal of Monetary Economics, 129-162.

Sargent, T.J., Macroeconomic Theory. New York: Academic Press, 1979.

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., New Methods in Business Cycle Research, Minneapolis: Federal Reserve Bank of Minneapolis.

Stekler, H.O., and M. Schepsman (1973), "Forecasting with an Index of Leading Series," Journal of the American Statistical Association 68, no. 342, 291-296.

Stock, J.H. and M.W. Watson (1986), "Testing for Common Trends," manuscript, Harvard University; forthcoming, Journal of the American Statistical Association.

Vaccara, B.N., and V. Zarnowitz (1978), "Forecasting with the Index of Leading Indicators," NBER Working Paper No. 244.

Wecker, W.E. (1979), "Predicting the Turning Points of a Time Series," Journal of Business 52, no. 1, 35-50.

Zarnowitz, V. and C. Boschan (1975), "Cyclical Indicators: An Evaluation and New Leading Indexes," in U.S. Department of Commerce, Bureau of Economic Analysis, Business Conditions Digest, May,1975, reprinted in Handbook of Cyclical Indicators.

Zarnowitz, V. and G.H. Moore (1982), "Sequential Signals of Recession and Recovery," Journal of Business 55, no. 1, 57-85.

Zellner, A., C. Hong, and G.M. Gulati (1987), "Turning Points in Economic Time Series, Loss Structures and Bayesian Forecasting," manuscript, Graduate School of Business, University of Chicago.

Appendix A

Variable Definitions

Unless otherwise noted, the data were obtained from the August, 1988 release of CITIBASE. The variable definitions are those in CITIBASE.

Coincident Variables

IP	INDUSTRIAL PRODUCTION: TOTAL INDEX (1977=100,SA)
GMYXP8	PERSONAL INCOME:TOTAL LESS TRANSFER PAYMENTS,82$(BIL$,SAAR)
MT82	MFG & TRADE SALES: TOTAL, 1982$(MIL$,SA)(BCD57)! 2 3
LPNAG	EMPLOYEES ON NONAG. PAYROLLS: TOTAL (THOUS.,SA)

Additional Variables (in alphabetical order)

BUS	INDEX OF NET BUSINESS FORMATION, (1967=100;SA)
CCI30M	CONSUMER INSTAL.LOANS: DELINQUENCY RATE,30 DAYS & OVER, (%,SA)
CONDO9	CONSTRUCT.CONTRACTS: COMM'L & INDUS.BLDGS(MIL.SQ.FT.FLOOR SP.;SA)
EXNWT1	Trade-weighted nominal exchange rate, U.S. vs Canada, France, Italy Japan, U.K., W. Germany (authors' calculation)
EXNWT2	Trade-weighted nominal exchange rate, U.S. vs France, Italy, Japan, U.K., W. Germany (authors' calculation)
EXRWT1	Trade-weighted real exchange rate, U.S. vs Canada, France, Italy, Japan, U.K., W. Germany; real rates based on CPI's (authors' calculation)
EXRWT2	Trade-weighted real exchange rate, U.S. vs France, Italy, Japan, U.K., W. Germany; real rates based on CPI's (authors' calculation)
FCBCUC	CHANGE IN BUS AND CONSUMER CREDIT OUTSTAND.(PERCENT,SAAR)(BCD111)
FCLBMC	WKLY RP LG COM'L BANKS:NET CHANGE COM'L & INDUS LOANS(BIL$,SAAR)
FCLN82	COMMERCIAL & INDUSTRIAL LOANS: OUTSTANDING,82$(MIL$,SA)! 2 3
FM1D82	MONEY STOCK: M-1 IN 1982$ (BIL$,SA)(BCD 105)
FM2D82	MONEY STOCK: M-2 IN 1982$(BIL$,SA)(BCD 106)
FMBASE	MONETARY BASE, ADJ FOR RESERVE REQ CHGS(FRB OF ST.LOUIS)(BIL$,SA)
FSDJ	COMMON STOCK PRICES: DOW JONES INDUSTRIAL AVERAGE
FSPCOM	S&P'S COMMON STOCK PRICE INDEX: COMPOSITE (1941-43=10)
FSPIN	S&P'S COMMON STOCK PRICE INDEX: INDUSTRIALS (1941-43=10)

FSVOL	STOCK MRKT: NYSE REPORTED SHARE VOLUME (MIL.OF SHARES;NSA)
FYCP	INTEREST RATE: COMMERCIAL PAPER, 6-MONTH (% PER ANNUM,NSA)
FYFF	INTEREST RATE: FEDERAL FUNDS (EFFECTIVE) (% PER ANNUM,NSA)
FYFFR	FYFF less 12-month CPI inflation rate (authors' calculation)
FYGM3R	FYGM3 less 12-month CPI inflation rate (authors' calculation)
FYGT1	INTEREST RATE: U.S.TREASURY CONST MATURITIES,1-YR.(% PER ANN,NSA)
FYGT10R	FYGT10 less 12-month CPI inflation rate (authors' calculation)
FYGT1R	FYGT1 less 12-month CPI inflation rate (authors' calculation)
GMCD82	PERSONAL CONSUMPTION EXPENDITURES:DURABLE GOODS,82$
GMWS	PERSONAL INCOME: WAGE & SALARY, SERVICE INDUSTRIES (BIL$,SAAR)
HSFR	HOUSING STARTS:NONFARM(1947-58);TOTAL FARM&NONFARM(1959-)(THOUS.,SA
IPCD	INDUSTRIAL PRODUCTION: DURABLE CONSUMER GDS (1977=100,SA)
IPI	INDUSTRIAL PRODUCTION: INTERMEDIATE PROD (1977=100,SA)
IPMFG	INDUSTRIAL PRODUCTION: MANUFACTURING (1977=100,SA)
IVM28	MFG INVENTORIES: CHEMICALS & ALLIED PRODUCTS (MIL$,SA)
IVPAC	VENDOR PERFORMANCE: % OF CO'S REPORTING SLOWER DELIVERIES(%,NSA)
LBOUTUM	Quarterly output per hour of all persons, business sector, distributed evenly across months in quarter, lagged two months (authors' calculation, based on CITIBASE series LBOUTU).
LHU5	UNEMPLOY.BY DURATION: PERSONS UNEMPL.LESS THAN 5 WKS (THOUS.,SA)
LOUTBM	Quarterly output per hour of all persons, nonfinancial corporate sector, distributed evenly across months in quarter, lagged two months (authors' calculation, based on CITIBASE series LOUTB).
LPGOV	EMPLOYEES ON NONAG. PAYROLLS: GOVERNMENT (THOUS.,SA)
LPSP	EMPLOYEES ON NONAG. PAYROLLS: SERVICE-PRODUCING (THOUS.,SA)
LPSPNG	LPSP less LPGOV (authors' calculation)
MOCM82	2 MFG NEW ORDERS: CONSUMER GOODS & MATERIAL,82$(BIL$,SA)! 2 3
MPCON8	CONTRACTS & ORDERS FOR PLANT & EQUIPMENT IN 82$(BIL$,SA)! 2 3
PPSMC	CHANGE IN SENSITIVE CRUDE AND INTERM PRODUCERS' PRICES(%)BCD98
PUNEW	CPI-U: ALL ITEMS (SA)
PW	PRODUCER PRICE INDEX: ALL COMMODITIES (NSA)
PWFC	PRODUCER PRICE INDEX: FINISHED CONSUMER GOODS (NSA)
PZUNEW	CPI-U: ALL ITEMS (NSA)
RCAR6D	RETAIL SALES: NEW PASSENGER CARS, DOMESTIC (NO.IN THOUS.;NSA)

TB1 Trade Balance as percent of personal income: total merchandise and trade exports less imports (current dollars), divided by current dollars personal income (authors' calculation, based on CITIBASE series FTM, FTE, F6TED, F6TMD, and GMPY)