Volume Title: Business Cycles, Indicators and Forecasting

Volume Author/Editor: James H. Stock and Mark W. Watson, editors

Volume Publisher: University of Chicago Press

Volume ISBN: 0-226-77488-0

Volume URL: http://www.nber.org/books/stoc93-1

Conference Date: May 3-4, 1991

Publication Date: January 1993

Chapter Title: A Procedure for Predicting Recessions with Leading Indicators: Econometric Issues and Recent Experience

Chapter Author: James H. Stock, Mark W. Watson

Chapter URL: http://www.nber.org/chapters/c7190

Chapter pages in book: (p. 95 - 156)

2 A Procedure for Predicting Recessions with Leading Indicators: Econometric Issues and Recent Experience

James H. Stock and Mark W. Watson

Since the pioneering work on leading indicators by Mitchell and Burns ([1938] 1961) and their collaborators at the NBER, the prediction of business-cycle turning points has been one of the core problems of business-cycle analysis. This paper describes one approach to forecasting the future state of the business cycle or, more simply, to predicting recessions. The paper has three objectives. The first is to provide the mathematical details of this approach to forecasting recessions. The second is to evaluate the empirical performance of the resulting recession probability forecasts. This evaluation focuses on the sharp economic downturn in the fall of 1990, which provided an opportunity to examine the performance of a range of leading economic indicators under the unusual conditions of a broadly weak economy facing the prospect of oil supply disruptions and war in the Persian Gulf. The third objective is to draw some general conclusions about the use of leading indicators for macroeconomic forecasting.

The methodology for estimating the probability that the economy will be in a recession at a future date is described in section 2.1. Rather than trying to forecast turning points (see, e.g., Kling 1987; Hymans 1973; Neftci 1982; Wecker 1979; and Zellner, Hong, and Gulati 1987), the scheme focuses on forecasting a 0/1 variable that indicates whether the economy will be in a recession in a given month. The basic idea is to define recessions and expansions as different patterns of economic activity in such a way that whether the

James H. Stock is professor of political economy at the Kennedy School of Government, Harvard University. Mark W. Watson is professor of economics at Northwestern University. Both are research associates of the National Bureau of Economic Research.

The authors thank Rob Engle, Kenneth Wallis, and Jeff Wooldridge for helpful comments on an earlier draft and Gustavo Gonzaga for research assistance. This work was supported in part by National Science Foundation grant SES-89-10601.

economy will be in a recession in, say, six months is equivalent to whether the path of overall economic activity six months hence falls in a recessionary or an expansionary pattern. With quantitative definitions for these two patterns, the probability that the economy is in a recession during a future month can then be computed by the stochastic simulation of a model that forecasts future economic activity.

The recession and growth forecasts examined here were produced by the model developed in Stock and Watson (1989). This model was estimated using data from January 1959 through September 1988. Since then, it has been used to produce three indexes of overall economic activity on a monthly basis: an experimental coincident index (the XCI); an experimental leading index (the XLI), which is a forecast of the growth in the XCI over the subsequent six months; and an experimental recession index (the XRI), which estimates the probability that the economy will be in a recession six months hence. The in-sample performance of the recession forecasts (the XRI) is examined in section 2.2. This investigation provides little evidence of misspecification in the recession definition, in the algorithm used to compute the recession probabilities, or in the linear structure of the forecasting model used to construct the XCI and the XLI.

The data since October 1988 provide true out-of-sample observations on the performance of the experimental indexes, including the recession index. Since May 1989, the XCI, XRI, and XLI have been publicly released on a monthly basis, with release dates approximately coinciding with the release of the Composite Index of Leading Indicators produced by the Department of Commerce (DOC). The performance of the experimental indexes over this period is studied in section 2.3. In brief, forecasts of growth rates through September 1990 performed quite well, with growth rate forecast errors half what they were in sample. However, the experimental indexes failed to forecast the sharp decline that began in October 1990.

Section 2.4 investigates a variety of possible sources for the poor performance of the indexes over the fall of 1990. The main conclusion is that the source of the large forecast errors and of the failure of the recession index to forecast the downturn is not the recession definition or the mathematical structure of the model but rather the choice of specific leading indicators used to construct the indexes. An analysis of a broad set of 45 coincident and leading indicators, including the seven in the experimental index, demonstrates that almost all performed quite poorly during this episode. Only a few, such as housing building permits, consumer expectations, a measure of business sentiment, oil prices, help wanted advertising, and stock prices, signaled that the economy would suffer a sharp contraction. It is of course easy to recognize that these particular indicators performed well ex post; the challenge is how they could have been identified ex ante. These and other conclusions are summarized in section 2.5.

2.1 Calculation of Recession Probabilities

This section outlines the procedure used to calculate the probability that the economy will be in a recession at time τ, conditional on leading and coincident economic indicators observed through time t. Let R_τ be an indicator variable that equals one if the economy is in a recession and zero otherwise. Throughout, x_s denotes a vector of coincident variables, and y_s denotes a vector of leading indicators that are useful in predicting future economic activity. It is assumed that x_s is stationary in first-differences and that the leading indicators have been transformed so that y_s is stationary.

The objective is to calculate the probability of being in a recession in month τ, given data on (x_s, y_s) through month t; this probability is denoted $P_{\tau|t}$. The approach to computing $P_{\tau|t}$ has three components: the specification of the conditional probability model for the state of the economy; the definition of the recession event R_τ in terms of the state of the economy; and the estimation of the model parameters. These three components are addressed in turn in the following subsections.

2.1.1 The Model

The probability model used to describe the evolution of $(\Delta x_t, y_t)$ is a dynamic single index model of the form proposed by Sargent and Sims (1977) and used, for example, by Geweke (1977) and Singleton (1980). This is discussed at length in Stock and Watson (1989, 1991) and is only sketched here. The comovements at all leads and lags among the coincident variables are modeled as arising from a single common source c_t, a scalar unobserved time series that can be thought of as the overall state of the economy. The idiosyncratic components of the growth of each of the coincident variables (the part not arising from leads and lags of c_t) is assumed to be stationary and uncorrelated with the idiosyncratic components of the other variables, but otherwise it can have a rich serial correlation structure. In particular,

$$\Delta x_t = \beta + \gamma(L)\Delta c_t + u_t, \tag{1}$$

$$D(L)u_t = \varepsilon_t, \tag{2}$$

$$\phi(L)\Delta c_t = \delta + \eta_t, \tag{3}$$

where (ε_t, η_t) are serially uncorrelated with a diagonal covariance matrix, and where $D(L) = \mathrm{diag}[d_{ii}(L)]$. To fix the timing of c_t, one of the elements of $\gamma(L)$, say, $\gamma_i(L)$, is set equal to γ_{i0} (in the empirical model, $\gamma_i[L] = \gamma_{i0}$ for three of the four coincident variables used).

Leading indicators are added to the model to help predict future values of c_t by replacing (3) with the autoregressive system,

(4) $$\Delta c_t = \mu_c + \lambda_{cc}(L)\Delta c_{t-1} + \lambda_{cy}(L)y_{t-1} + \nu_{ct},$$

(5) $$y_t = \mu_y + \lambda_{yc}(L)\Delta c_{t-1} + \lambda_{yy}(L)y_{t-1} + \nu_{yt},$$

where $\nu'_t = (\nu'_{ct}\, \nu'_{yt})$ is serially uncorrelated with mean zero and is independent of ε_t.

The model (1), (2), (4), and (5) can be solved to obtain linear minimum mean square error linear forecasts of future values of Δy_t and x_t or to estimate the unobserved state Δc_t or c_t. This is readily implemented using the Kalman filter, as described in Stock and Watson (1991). With the additional assumption, which is made throughout, that (ε_t, ν_t) are jointly normal with constant conditional covariances, these linear projections are also conditional expectations.

2.1.2 Definition of Recessions and Expansions

A key aspect of this analysis is obtaining a quantifiable definition of a recession. Burns and Mitchell (1946, 3) provide a somewhat vague but nonetheless useful description of a recession as a substantial prolonged decline in economic activity that occurs broadly across various sectors of the economy. More recent working definitions used by business-cycle analysts refine these ideas and emphasize the "three Ds": for a slowdown to be a recession, it should be sufficiently long (duration), it should involve a substantial decline in economic activity (depth), and it should involve multiple sectors or all the sectors of the economy rather than simply reflecting an isolated decline in a single sector or region (diffusion).

The generally accepted business-cycle chronology is maintained by the NBER's Business Cycle Dating Committee. In practice, each individual on the committee must trade off these various parts of the definition to decide whether a particular episode warrants classification as a recession. The committee eschews numerical rules; this would limit its flexibility in deeming a particular episode a recession when there are unforeseen extenuating circumstances that are not amenable to being incorporated in a formulaic definition.

The definition of a recession adopted here attempts to capture, in a simple way, the institutional process in which recessions are categorized. We define a recession in terms of the growth of the unobserved state of the economy, Δc_t; this embodies the requirement that the recession be economywide, not specific to only one or two individual series. We treat the problem of classifying a sequence $\{\Delta c_t\}$, were it observed, as a pattern recognition problem: if the sequence falls in a recessionary pattern, then it is classified as a recession; if it falls in an expansionary pattern, it is an expansion. The recessionary and expansionary patterns that are possible in a sequence $(\Delta c_{t-k+1}, \ldots, \Delta c_t)$ of length k constitute subsets of $\Re^k$; whether such a sequence is an expansion or a recession depends on which subset the sequence falls in.

We suppose there to be two elementary recessionary patterns. In the first, D_1, Δc_t falls below a threshold $b_{r,t}$ for six consecutive months; in the second,

D_2, Δc_t falls below $b_{r,t}$ for seven of nine consecutive months, including the first and last months. That is,

(6) $D_{1\tau} = \{\Delta c_s, s = \tau - 5, \ldots, \tau: \Delta c_s \leq b_{r,s}, s = \tau - 5, \ldots, \tau\}$,

(7) $D_{2\tau} = \{\Delta c_s, s = \tau - 8, \ldots, \tau: \Delta c_{\tau-8} \leq b_{r,\tau-8}, \Delta c_\tau \leq b_{r,\tau},$
$\#(\Delta c_s \leq b_{r,s}, s = \tau - 7, \ldots, \tau - 1) \geq 5\}$,

where $\#(\cdot)$ denotes the number of times that the event occurs. Given the thresholds $\{b_{r,t}\}$, the economy is in a recession in month t if and only if that month falls in a recessionary pattern. Since a recessionary pattern $D_{1\tau}$ can commence anytime between $t - 5$ and t for the month to be in a recession, the set of recessionary patterns constituting a recession at date t is

(8) $$D_t = \left(\bigcup_{\tau=t}^{t+5} D_{1\tau}\right) \cup \left(\bigcup_{\tau=t}^{t+8} D_{2\tau}\right) \in \Re^{17}.$$

Thus, the recession event R_t is

(9) $$R_t = 1[(\Delta c_{t-8}, \ldots, \Delta c_t, \ldots, \Delta c_{t+8}) \in D_t],$$

where $1(\cdot)$ is the indicator function. An expansion event is defined symmetrically. Specifically,

(10) $U_{1\tau} = \{\Delta c_s, s = \tau - 5, \ldots, \tau: \Delta c_s > b_{e,s}, s = \tau - 5, \ldots, \tau\}$,

(11) $U_{2\tau} = \{\Delta c_s, s = \tau - 8, \ldots, \tau: \Delta c_{\tau-8} > b_{e,\tau-8}, \Delta c_\tau > b_{e,\tau},$
$\#(\Delta c_s > b_{e,s}, s = \tau - 7, \ldots, \tau - 1) \geq 5\}$,

(12) $$U_t = \left((\bigcup_{\tau=t}^{t+5} U_{1\tau}\right) \cup \left(\bigcup_{\tau=t}^{t+8} U_{2\tau}\right) \in \Re^{17},$$

(13) $$E_t = 1[(\Delta c_{t-8}, \ldots, \Delta c_t, \ldots, \Delta c_{t+8}) \in U_t].$$

The complement of U_t and D_t in $\Re^{17}$ is nonempty; that is, these definitions leave room for indeterminant sequences. Because the recession/expansion classification is dichotomous, these indeterminant events are ruled out in computing the probability of a recession. Thus, the probability that the economy is in a recession in month τ, conditional on coincident and leading indicators observed through month t and the cutoff values, is

$$\Pr[R_\tau = 1 | (R_\tau = 1) \cup (E_\tau = 1), x_t, x_{t-1}, \ldots, x_1, y_t, y_{t-1}, \ldots, y_1; b_r, b_e),$$

where b_r and b_e are the collection of cutoff values.

This probability is conditional on the sequence of cutoffs $(b_{r,t}, b_{e,t})$. One approach is to treat these as unknown time-invariant parameters, which could then be estimated. There are, however, at least two arguments for treating these parameters as random. First, this definition is in terms of c_t, while the process of identifying actual recessions involves the examination of a broad set of indicators; one interpretation of this is that the cutoff used in the recession definition should itself depend on macroeconomic variables that are omitted from this analysis. Second and alternatively, the process by which the Business Cycle Dating Committee reaches a decision involves different assessments of what constitutes a recession among the different members of the committee; one model of this is that each committee member has in mind some pair $(b_{r,t}, b_{e,t})$ for month t but that these vary across committee members and indeed over time for each member. Both arguments suggest that $(b_{r,t}, b_{e,t})$ can usefully be treated as random, and this approach is adopted here. Specifically, $b_{r,t}$ and $b_{e,t}$ are modeled as

$$(14) \qquad b_{r,t} = \mu_r + \zeta_t, \quad b_{e,t} = \mu_e + \zeta_t, \quad \zeta_t \text{ i.i.d. } N(0, \sigma_\zeta),$$

where ζ_t is independent of (ε_t, ν_t).

The probability that the economy is in a recession in month τ, given information through month t, is thus

$$(15) \qquad \begin{aligned} P_{\tau|t} &= \int \Pr[R_\tau = 1 | (R_\tau = 1) \cup (E_\tau = 1), x_t, x_{t-1}, \ldots, x_1 \\ &\qquad y_t, y_{t-1}, \ldots, y_1 \,;\, b_e(\tilde{\zeta}_\tau), b_r(\tilde{\zeta}_\tau)] dF_{\tilde{\zeta}}(\tilde{\zeta}_\tau) \\ &= E[R_\tau | (R_\tau = 1) \cup (E_\tau = 1), x_t, x_{t-1}, \ldots, x_1\, y_t, y_{t-1}, \ldots, y_1], \end{aligned}$$

where $\tilde{\zeta}_\tau = (\zeta_{\tau-8}, \ldots, \zeta_{\tau+8})'$, and $F_{\tilde{\zeta}}(\cdot)$ is the c.d.f. of $\tilde{\zeta}_\tau$.

The conditional probability $P_{\tau|t}$ involves integrating over a thirty-four-dimensional Gaussian distribution (a seventeen-fold integration to compute the conditional probability

$$\Pr[R_\tau = 1 | (R_\tau = 1) \cup (E_\tau = 1), x_t, x_{t-1}, \ldots, y_t, y_{t-1}, \ldots \,;\, b_r(\tilde{\zeta}_\tau), b_e(\tilde{\zeta}_\tau)]$$

and an additional seventeen-fold integration over $\tilde{\zeta}_\tau$). In practice, the integration is performed by Monte Carlo simulation using the following algorithm:

i. Compute the conditional mean $m_{\tau|t}$ and covariance matrix $\Omega_{\tau|t}$ of $\tilde{c}_\tau(-8, 8)$, where $\tilde{c}_\tau(-k_1, k_2) = (\Delta c_{\tau-k_1}, \ldots, \Delta c_\tau, \ldots, \Delta c_{\tau+k_2})$, given data through month t. (In steady state, $\Omega_{\tau|t}$ is a function of $\tau - t$, not τ or t separately.)
ii. Draw a pseudorandom realization of $\tilde{c}_\tau(-8, 8)$ from the $N(m_{\tau|t}, \Omega_{\tau|t})$ conditional distribution of $\tilde{c}_\tau$.
iii. Draw a realization of $\tilde{b}_{r,\tau}$ and $\tilde{b}_{e,\tau}$, where $\tilde{b}_{i,\tau} = (b_{i,\tau-8}, \ldots, b_{i,\tau+8})$, as $(\tilde{b}_{r,\tau}, \tilde{b}_{e,\tau}) = (\mu_e + \tilde{\zeta}_\tau, \mu_r + \tilde{\zeta}_\tau)$ according to (14).
iv. For each realization of $[\tilde{c}_\tau(-8, 8), \tilde{b}_{r,\tau}, \tilde{b}_{e,\tau}]$, evaluate R_t and E_t according to (9) and (13), respectively.
v. Repeat ii–iv (in practice enough times to obtain a minimum of two thousand draws of E_t or R_t), and compute $P_{\tau|t}$ as $\#(R_t)/[\#(R_t) + \#(E_t)]$.

It is worth emphasizing that this definition of a recession treats the identification of recessions (more generally, cycles) as a pattern recognition algorithm that could be applied to many series. This contrasts with approaches in which R_t is related to the time-series properties of the process, in which R_t is useful in predicting future c_t given its past. An example of the latter situation is Hamilton's (1989) model in which a discrete variable, empirically identified as a recession/expansion indicator, enters the conditional mean of the time series. One can usefully think of the latter situation as being one in which the definition of the recession event is intrinsic to the time-series model generating the data; a recession is then not well defined if the process is in fact linear and Gaussian. In contrast, the pattern recognition approach developed here can be applied whether the series is linear, Gaussian, or stationary.

2.1.3 Estimation of the Model Parameters

The estimation strategy is based on a partition of the joint density of the leading indicators, the coincident variables, and the recession indicator. Let $Y_t = (y_1, \ldots, y_t)$, $X_t = (x_1, \ldots, x_t)$, $S_t = (R_1, \ldots, R_t)$, and $C_t = (c_1, \ldots, c_t)$. The joint density of (Y_T, X_T, S_T) can be factored

$$(16) \qquad f(Y_T, X_T, S_T|\theta, \mu) = f_1(S_T|Y_T, X_T; \mu, \theta) f_2(Y_T, X_T|\theta).$$

This factorization is done without loss of generality and serves to define the parameter vector μ as the additional parameters introduced in the conditional density f_1. In terms of the model in section 2.1.1 and the definition of the recession variable R_t, θ is the vector of parameters given in (1), (2), (4), and (5), and μ is the vector of parameters describing the distribution of the recession threshold parameters, so $\mu = (\mu_e, \mu_r, \sigma_\zeta)$ as defined in (14).

In general, as long as θ appears in f_1, computing the maximum likelihood estimator (MLE) will entail maximization of the joint density $f(Y_T, X_T, S_T|\theta, \mu)$. The MLE simplifies to a two-stage process if θ does not appear in f_1, which would occur were R_t defined in terms of the observable variables (X_t, Y_t) and the parameters μ, for example, if Δc_t were replaced by Δx_{1t} in the definitions of D_{iT} and E_{iT} in section 2.1.2. However, because c_t is unobserved, θ enters f_1, and the MLE does not have a convenient simplification. Intuitively, because c_t is unobserved, R_t provides another dependent variable (in this case, discrete valued) that, in conjunction with the continuous variables, potentially provides useful information for estimation.

Unfortunately, because R_t is a discrete-valued time-series variable, the implementation of the MLE for (16) is numerically imposing. The parameters are therefore estimated in a two-stage process, estimating θ first, then μ. The estimation of θ is described at length in Stock and Watson (1989, 1991) and is not discussed here. In the second stage, μ is estimated conditional on the first-stage estimate of θ. While this simplifies the estimation of θ, maximization of the conditional likelihood $f_1(S_T|Y_T, X_T; \mu, \theta)$ remains numerically demanding. Estimation therefore proceeds by minimizing the mean square error $\Sigma_{t=t_1}^{T}(R_t - P_{t|T})^2$ (where $t_1 = t + 36$ so that the probabilities could be computed using the steady-state state covariance matrix Ω). The resulting estimators for μ_r, μ_e, and σ_ζ, computed by a grid search, are $\hat{\mu}_r = -1.5$, $\hat{\mu}_e = -0.25$, and $\hat{\sigma} = 0.8$.[1]

The estimated model and various in-sample specification tests are discussed in Stock and Watson (1989, 1991), to which the reader is referred for details.

2.1.4 Treatment of Data Irregularities

The form of the model in sections 2.1.1–2.1.3 used for monthly forecasting incorporates two modifications for data irregularities that arise when working with monthly data releases. Both involve conceptually straightforward (but computationally and notationally involved) modifications of the basic Kalman filter for the state space representation of the model (1), (2), (4), and (5). The general strategy for handling data irregularities is to make an appropriate modification of the state vector, the state transition equation, and the measurement equation. We now turn to the specifics.

One coincident indicator (manufacturing and trade sales) is reported by the Department of Commerce with a lag of an additional month. Let α_t denote the state vector in the state space representation of (1), (2), (3), and (4); let $\tilde{\alpha}_{t|t}$ denote the expected value of the state vector given observations on all variables except x_{it} through month t and on x_{it} through month $t - 1$, and let $\alpha_{t|t}$ denote the expected value of α_t given data on all variables through date t. Because complete data are available through $t - 1$, the Kalman filter can be applied to the unmodified model to form $\alpha_{t-1|t-1}$. At date t, the state space model is altered by modifying the measurement equation (1) to exclude the equation for the coincident variable in question. Alternatively, the equation could be included, an arbitrary finite observation used for the variable in question, and a measurement error term appended to (1) with infinite variance (in practice approximated by a large constant).

The second important modification of the standard Kalman filter is to handle revisions in many of the coincident and leading variables. Let z^j_{it} denote the value of z_{it} published at date $t + j$, where z_{it} indicates an element of the vector $z_t = (\Delta x_t', y_t')$. Thus, $j = 0$ corresponds to the initial release of z_{it}, $j = 1$ corresponds to the first monthly revision, etc. The revision error is $z_{it} - z^j_{it} = e^j_{it}$, and the model is modified to account for this additional error. The appropriate modification to the model depends on the covariance properties of e^j_{it}. We find it useful to consider two extreme assumptions concerning e^j_{it}, analogous to the "news" and "noise" assumptions of Mankiw and Shapiro (1986) (see also Mankiw, Runkle, and Shapiro 1984). The first assumption—noise—corresponds to the classical errors-in-variable model

1. The sensitivity to the choice of optimand was checked by recomputing the estimates using the pseudolikelihood obtained by treating R_t as an independent Bernoulli random variable with probability $P_{t|T}$. The point estimates for μ_r, μ_e, and σ were close for the two optimands, and, more important, the estimated probabilities $P_{t|T}$ were virtually indistinguishable. For both optimands, the surface of the objective function was rather flat in a neighborhood of the optimized values. Evidence of this insensitivity is given in sec. 2.4.1 below.

(17) $$z_{it}^{j} = z_{it} + e_{it}^{j},$$

where e_{it}^{j} is uncorrelated with z_{it}. Because this is a dynamic model, it is further assumed that e_{it}^{j} is uncorrelated with all values of the actual data, that is, $E(e_{it}^{j}z_{k\tau}) = 0$ for all j, i, k, t, and τ, and that measurement errors are uncorrelated across series, that is, $E(e_{it}^{j}e_{k\tau}^{n}) = 0$ for all j, n, t, and τ when $k \neq i$.

The second assumption—news—corresponds to the optimal forecasting model

(18) $$z_{it} = z_{it}^{j} + e_{it}^{j},$$

where e_{it}^{j} is uncorrelated with z_{it}^{j}. Thus, z_{it}^{j} is viewed as an unbiased forecast of z_{it}, and e_{it}^{j} contains information (news) about z_{it} not contained in z_{it}^{j}.

The modifications needed to incorporate a single "noise" variable in univariate models are discussed in Harvey et al. (1981). The modification to handle multiple noise variables in this application is a straightforward generalization of this single variable modification. The modifications necessary to incorporate a "news" variable into the model are simpler: if the preliminary variable is an optimal forecast of the final variable, and if (as is assumed) the data collection agency uses a superset of the information in (X_t, Y_t) to produce this optimal forecast of the final series, then optimal estimates and forecasts of α_t can be constructed by substituting the preliminary data in place of the actual data and running the Kalman filter on the unmodified model. However, while no modification is necessary to produce $\alpha_{t|t}$, it is necessary to modify its covariance matrix to reflect the increased uncertainty associated with the preliminary data. The details of the Kalman filter modifications for measurement error are provided in Stock and Watson (1988).[2]

2.1.5 Summary of the Estimated Indexes and Their Interpretation

Since its estimation in early 1989, this model has been used to produce three indexes on a monthly basis: an experimental coincident index (XCI), an experimental leading index (XLI), and an experimental recession index (XRI). The XCI is the estimate of the state at time t, that is, XCI $= c_{t|t}$. The XLI is the estimate of its growth over the subsequent six months, $c_{t+6|t} - c_{t|t}$ (because x_t is in logarithms, $c_{t+6} - c_t$ is the six-month growth in c_t; the XLI is reported at annual percentage growth rates, i.e., $200[c_{t+6|t} - c_{t|t}]$). And the XRI is the probability that the economy will be in a recession in six months (XRI $= P_{t+6|t}$). The coincident and leading variables used in the model, which were selected by a modified stepwise regression procedure (see Stock and Watson 1989), are listed in panels A and B, respectively, of table 2.1. Since mid-1990, we have also been tracking a second set of indexes (the XLI2 and the XRI2), based solely on nonfinancial indicators. The coincident indi-

2. The empirical implementation allows for a maximum of $j = 12$ revisions. The covariance matrices of $e_{it} = (e_{it}^{1}, \ldots, e_{it}^{12})'$ were estimated using data from 1981:1 through 1985:12.

Table 2.1 Coincident and Leading Indicators in the XRI and XRI2

Mnemonic	Transformation	Description
		A. Coincident Indicators
IP	Growth rates	Industrial production, total
GMYXP8	Growth rates	Personal income, total less transfer payments, 1982$
MT82	Growth rates	Manufacturing and trade sales, total, 1982$
LPMHU	Growth rates	Employee-hours in nonagricultural establishments
		B. Leading Indicators in the XLI
HSBP	Levels	Housing authorizations—new private housing
MDU82S	Growth rates	Manufacturers' unfilled orders: durable goods industries, 1982$, smoothed
EXNWT2S	Growth rates	Trade-weighted nominal exchange rate between the United States and the United Kingdom, West Germany, France, Italy, and Japan, smoothed
LHNAPSS	Growth rates	Part-time work in nonagricultural industries because of slack work (U.S. Department of Labor, The Employment Situation, Household Survey), smoothed
FYGT10S	Differences	Yield on constant-maturity portfolio of 10-year U.S. Treasury bonds, smoothed
CP6_GM6	Levels	Spread between interest rate on a 6-month commercial paper and the interest rate on 6-month U.S. Treasury bills (Federal Reserve Board)
G10_G1	Levels	Spread between the yield on constant-maturity portfolio of 10-year U.S. T-bonds and the yield on 1-year U.S. T-bonds (Federal Reserve Board)
		C. Leading Indicators in the XLI2
HSBP	Levels	Housing authorizations—new private housing
MDU82S	Growth rates	Manufacturers' unfilled orders: durable goods industries, 1982$, smoothed
EXNWT2S	Growth rates	Trade-weighted nominal exchange rate between the United States and the United Kingdom, West Germany, France, Italy, and Japan, smoothed
LPHRM	Levels	Average weekly hours of production workers in manufacturing
IPXMCA	Differences	Capacity utilization rate in manufacturing, total (Federal Reserve Board)
LHEL	Growth rates	Index of help wanted advertising in newspapers (The Conference Board)
IVPAC	Levels	Vendor performance: percentage of companies reporting slower deliveries

Note: The series described as "smoothed" were passed through the filter $(1 + 2L + 2L^2 + L^3)$. All variables except exchange rates and interest rates are seasonally adjusted.

cators entering the XLI2 are those in panel A, and the leading indicators entering the XLI2 are given in panel C of table 2.1.

Empirically, the XCI can be thought of as a monthly proxy for real GNP. Simple regression relations between the XCI produced by the estimated model, aggregated to a quarterly level, and real GNP are presented in table 2.2. The correlation between the six-month growth of the XCI and real GNP is large, approximately .88. Although the mean growth of the XCI and real GNP are approximately equal over this period, XCI growth is more volatile, and the regression coefficient of GNP growth onto XCI growth is .58. This implies that XCI growth of zero corresponds approximately to GNP growth of 1.3 percent.

2.2 In-Sample Analysis of Probabilities

This section examines the within-sample performance of the estimated recession probabilities. The analysis focuses on three types of potential misspecification: misspecification of the probability model, so that the information in the included leading and coincident indicators is not fully incorporated into the predicted probabilities; omission of alternative indicators that help predict recessions; and misspecification associated with the possible duration dependence in recessions and expansions, that is, with the possibility that the length of the current recession (expansion) might usefully predict when the next expansion (recession) will occur.

The probabilities examined here are based on the model outlined in section 2.1, estimated in early 1989 using data from 1959:1 to 1988:9. The seven leading indicators used in the XLI were selected from a "short list" of fifty-five series. Any such selection of a few variables from many exacerbates the usual risks of overfitting, so the in-sample analysis in this section provides only limited guidance in assessing the performance of the model. Still, rejection by these in-sample diagnostics would suggest specification problems in the way the probabilities are calculated.

Table 2.2 **Relation between the XCI and Real GNP: OLS Regressions of the Form $\ln(RGNP_t/RGNP_{t-k}) = \alpha + \beta \ln(XCI^{Q}_t/XCI^{Q}_{t-k}) + e_t$ (1962: I–1988:III, where XCI^{Q}_t is the XCI, aggregated to a quarterly level)**

k (quarters)	$\hat{\alpha}$	$\hat{\beta}$	R^2	SEE
1	1.286	.577	.65	2.37
	(.264)	(.042)		
2	1.296	.578	.78	1.50
	(.173)	(.030)		

Note: Autocorrelation-robust standard errors (computed using 6 lagged autocovariances with a Bartlett kernel) are reported in parentheses. Estimation used quarterly observations. The quarterly XCI^{Q} series was constructed by averaging the values of the XCI over the months in the quarter.

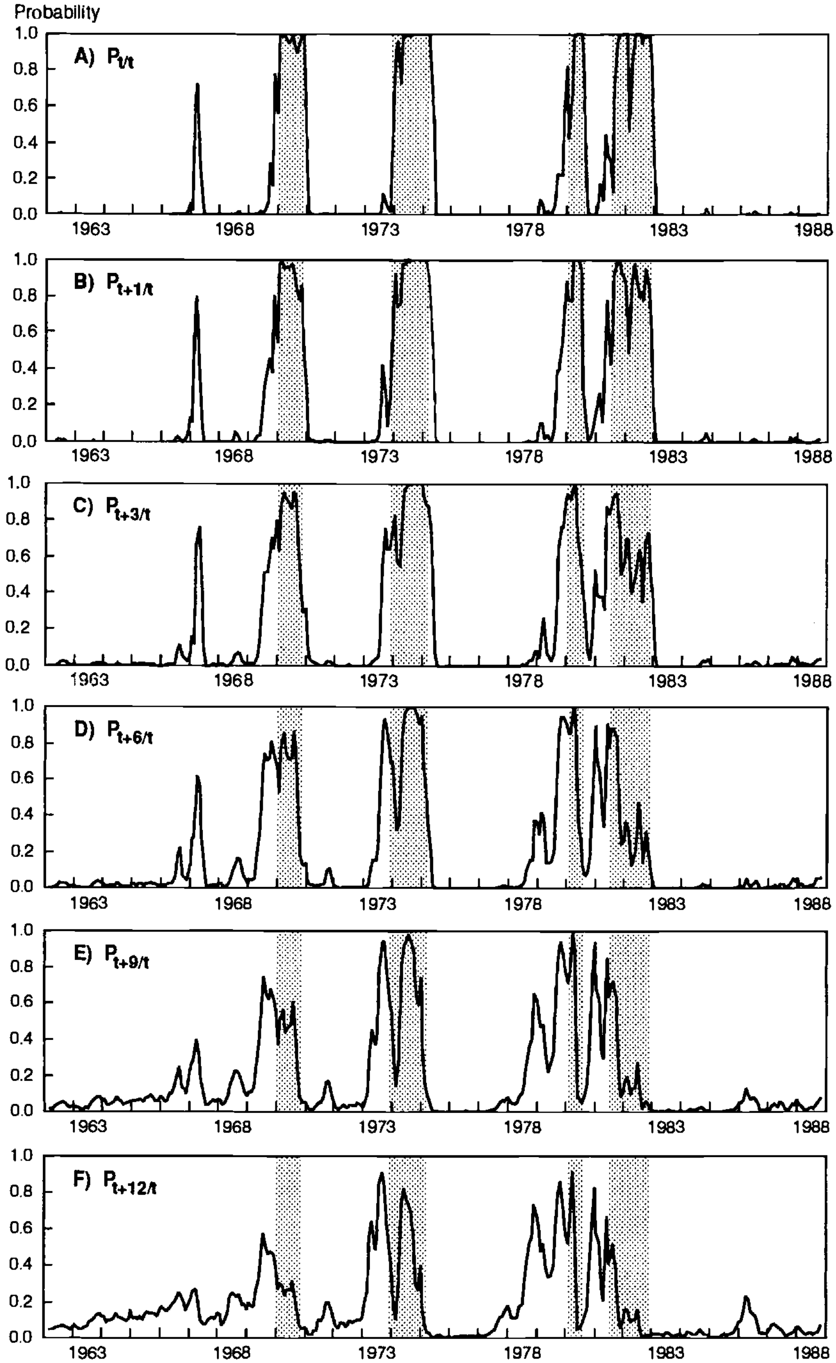

Fig. 2.1 Estimated recession probabilities, 1962:1–1988:9

Note: The dates on the horizontal axes denote t, the date through which the data are available for computing $P_{t+k/t}$. The figure is based on data revised through 1988:9.

The recession probabilities $P_{t+k|t}$, as estimated through 1988:9, are plotted in figure 2.1 for various horizons. (The dating convention plots $P_{t+k|t}$ at time t.) At a horizon of one month, the probabilities are sharp; the forecasts become substantially less precise as the horizon increases beyond six months.

The performance of these predictions is investigated in tables 2.3–2.5. Panel A of table 2.3 presents, for each horizon, the average predicted probability, $\bar{P}$, the proportion of recession realizations $\bar{R}$, the R^2 of the predictions, and the RMSE (root mean square error) of the prediction. The table suggests a slight bias in the predictions: $\bar{P} > \bar{R}$ for all horizons. (Because the probabilities are nonlinear functions of μ_r, μ_e, and σ, minimizing the mean square

Table 2.3 Predictive Performance of Recession Probabilities $P_{t+k|t}$: Summary Statistics

A. Overall

Statistic[a]	Forecast Horizon (Months) 1	3	6	9	12
$\bar{P}$	.180	.186	.187	.185	.183
$\bar{R}$	.146	.147	.148	.150	.151
R^2	.789	.687	.577	.482	.367
RMSE	.176	.215	.251	.279	.309

B. Statistics by Cell

Cell	Statistic[a]	Forecast Horizon (Months) 1	3	6	9	12
$.0 \le P < .25$	$\bar{P}$	.015	.022	.042	.061	.086
	$\bar{R}$	.004	.000	.016	.024	.040
	N	261	250	251	249	248
	N_R	1	0	4	6	10
$.25 \le P < .50$	$\bar{P}$	.368	.362	.349	.349	.343
	$\bar{R}$	.067	.211	.308	.387	.525
	N	15	19	26	31	40
	N_R	1	4	8	12	21
$.50 \le P < .75$	$\bar{P}$	.627	.613	.639	.630	.613
	$\bar{R}$	.364	.577	.571	.643	.519
	N	11	26	21	28	27
	N_R	4	15	12	18	14
$.75 \le P < 1.0$	$\bar{P}$	.918	.896	.895	.889	.845
	$\bar{R}$	.896	.789	.781	.684	.444
	N	48	38	32	19	9
	N_R	43	30	25	13	4

[a]For definitions, see the text.

Table 2.4 In-Sample Regression Tests for Omitted Variables in $P_{t+k|t}$ (p-values of test statistics) Based on OLS Regressions, 1962:1–1988:9 − k

	Forecast Horizon (Months)				
Variable	0	1	3	6	9
Constant	.064	.058	.141	.327	.475
			Coincident Indicators		
IP	.740	.549	.961	.545	.293
GMYXP8	.914	.317	.624	.950	.089
MT82	.614	.655	.558	.746	.389
LPMHUADJ	.482	.249	.396	.594	.414
			Leading Indicators in the XRI		
HSBP	.770	.865	.978	.122	.349
MDU82S	.094	.126	.241	.497	.861
EXNWT2FS	.755	.422	.425	.425	.342
LHNAPSS	.837	.385	.143	.476	.544
FYGT10FS	.734	.750	.004	.209	.461
CP6_GM6F	.823	.894	.858	.246	.985
G10_GLF	.859	.830	.277	.221	.727
XLI	.311	.382	.499	.654	.647
			Leading Indicators in the XRI2		
LPHRM	.258	.275	.475	.847	.625
IPXMCA	.613	.478	.555	.458	.305
LHEL	.839	.567	.558	.755	.965
IVPAC	.176	.135	.012	.607	.216
			Financial Indicators		
FSPCOMF	.083	.156	.041	.442	.589
FM1D82	.734	.493	.304	.275	.951
FM2D82	.841	.289	.260	.560	.998
FMBASE	.561	.867	.475	.553	.743
CCI30M	.738	.991	.882	.502	.354
FCBCUCY	.202	.209	.263	.743	.263
FYFFF	.981	.785	.198	.332	.948
BAA_G10F	.588	.615	.024	.244	.785
YLD_DUMF	.401	.200	.882	.800	.703
			Employment Indicators		
LUINC	.977	.942	.896	.596	.530
LHU5	.932	.454	.282	.758	.533
LHELX	.983	.750	.938	.558	.751

Table 2.4 (continued)

	Forecast Horizon (Months)				
Variable	0	1	3	6	9
	Consumption and Retail Sales				
IPCD	.685	.262	.417	.661	.424
GMCD82	.814	.628	.421	.689	.827
RTR82	.498	.493	.249	.707	.905
	Inventories and Orders				
MPCON8	.363	.334	.659	.225	.921
MOCM82	.285	.243	.225	.498	.462
MDO82	.173	.079	.086	.880	.388
IVMT82	.353	.231	.199	.667	.668
IVM1D8	.640	.218	.139	.811	.295
IVM2D8	.490	.506	.559	.556	.438
IVM3D8	.919	.093	.146	.397	.524
	Additional Indicators				
DLBLNPAP	.506	.705	.686	.577	.648
PMI	.370	.400	.266	.835	.498
PMNO	.290	.365	.446	.966	.529
HHSNTN	.414	.153	.057	.626	.471
HHST	.092	.813	.072	.850	.211
PW561	.482	.864	.798	.782	.731
PW561R	.505	.722	.828	.816	.765
FTM333	.601	.440	.216	.231	.527
FTM333R	.624	.485	.349	.472	.573
	Composite Indexes and Measures of Duration				
DLEAD	.686	.589	.125	.721	.268
DL3D	.790	.585	.478	.080	.032
DL3U	.355	.441	.572	.431	.858
IP3D	.011	.060	.596	.465	.423
IP3U	.145	.303	.459	.605	.762
MTREC	.002	.128	.627	.541	.368
MTEXP	.595	.463	.518	.148	.079
MTTOT	.538	.282	.615	.511	.912

Note: The p-values refer to Wald tests of the hypothesis that the coefficients on $(z_t, \ldots, z_{t-5})$ in the regression of $R_{t+k} - P_{t+k|t}$ on a constant and $z_t, \ldots, z_{t-5}$ are zero, where k refers to the forecast horizon (months). The tests were computed using autocorrelation- and heteroskedasticity-robust covariance matrices, constructed as weighted averages of $k + 5$ autocovariances with Bartlett kernel weights. A p-value of .000 denotes a p-value $< .0005$. The regressions were estimated from 1962:1–1988:9 $- k$. The variables are defined in the appendix.

Table 2.5 In-Sample Regression Tests for Omitted Variables in $P_{t+k|t}$ (p-values of test statistics) Based on WLS Regressions, 1962:1–1988:9 − k

Variable	Forecast Horizon (Months) 0	1	3	6	9
Constant	.001	.000	.001	.123	.035
	Coincident Indicators				
IP	.670	.760	.841	.658	.503
GMYXP8	.587	.853	.559	.882	.639
MT82	.932	.731	.682	.571	.899
LPMHUADJ	.981	.762	.466	.955	.492
	Leading Indicators in the XRI				
HSBP	.741	.840	.889	.844	.656
MDU82S	.352	.437	.466	.700	.455
EXNWT2FS	.593	.481	.799	.479	.172
LHNAPSS	.625	.678	.801	.699	.584
FYGT10FS	.882	.898	.573	.806	.706
CP6_GM6F	.973	.779	.953	.782	.547
G10_GLF	.701	.565	.461	.877	.300
XLI	.499	.535	.743	.933	.330
	Leading Indicators in the XRI2				
LPHRM	.735	.409	.461	.596	.510
IPXMCA	.486	.564	.372	.150	.131
LHEL	.398	.369	.677	.875	.954
IVPAC	.757	.692	.468	.857	.743
	Financial Indicators				
FSPCOMF	.630	.364	.483	.934	.755
FM1D82	.690	.803	.678	.767	.837
FM2D82	.870	.618	.692	.691	.837
FMBASE	.830	.776	.559	.609	.879
CCI30M	.628	.555	.494	.602	.501
FCBCUCY	.737	.874	.302	.899	.684
FYFFF	.744	.772	.773	.843	.445
BAA_G10F	.941	.915	.258	.829	.486
YLD_DUMF	.719	.955	.865	.777	.422
	Employment Indicators				
LUINC	.969	.934	.842	.950	.841
LHU5	.827	.851	.438	.826	.973
LHELX	.663	.806	.527	.801	.540

Table 2.5 (continued)

	Forecast Horizon (Months)				
Variable	0	1	3	6	9
	Consumption and Retail Sales				
IPCD	.831	.710	.639	.853	.538
GMCD82	.993	.797	.860	.740	.979
RTR82	.974	.648	.517	.697	.879
	Inventories and Orders				
MPCON8	.758	.849	.899	.480	.894
MOCM82	.507	.293	.499	.768	.859
MDO82	.375	.528	.759	.772	.654
IVMT82	.400	.491	.594	.787	.624
IVM1D8	.903	.545	.737	.608	.670
IVM2D8	.893	.861	.396	.897	.552
IVM3D8	.984	.662	.422	.969	.823
	Additional Indicators				
DLBLNPAP	.779	.865	.855	.678	.711
PMI	.539	.665	.736	.390	.841
PMNO	.584	.408	.660	.427	.854
HHSNTN	.728	.882	.721	.956	.792
HHST	.235	.788	.634	.906	.794
PW561	.054	.897	.980	.701	.910
PW561R	.078	.812	.942	.772	.866
FTM333R	.651	.896	.332	.883	.675
	Composite Indexes and Measures of Duration				
DLEAD	.870	.755	.749	.918	.611
DL3D	.924	.877	.932	.890	.806
DL3U	.151	.705	.875	.861	.576
IP3D	.388	.435	.613	.670	.611
IP3U	.333	.564	.195	.676	.860
MTREC	.012	.005	.040	.917	.752
MTEXP	.378	.434	.114	.000	.134
MTTOT	.620	.575	.098	.084	.054

Note: Computed by weighted least squares regression as discussed in the text, with weights $w_t = \min\{P_{t+k|t}(1 - P_{t+k|t}), .01\}$. See the note to table 2.4.

error [MSE] need not result in unbiased forecasts: reducing the sample bias would increase the sample MSE.)

Panel B of table 2.3 takes a closer look at the predictions by partitioning the observations into cells based on the predicted value. In the table, N represents the number of observations in the cell, N_R represents the number of these observations that turned out to be periods of recession. For example, of the

251 times within sample that $P_{t+6|t}$ fell within (0, 0.25), only four of those turned out to be recessionary months; if a value of $P_{t+6|t}$ below .25 is interpreted as a signal of "no recession," this corresponds to a false negative rate (the probability of a recession given a forecast of no recession) of 1.6 percent. Similarly, if $P_{t+6|t} \geq 0.75$ is interpreted as a recession signal, then this signal had a within-sample false positive rate of 22 percent (7/32). This interpretation of false negative and positive rates corresponds to monthly forecasts of whether the economy will be in a recession, which is different than whether the economy will shift from an expansion to a recession, or vice versa, in the next six months. The latter concept is of practical interest, but, given the few turning points in the sample, it is one for which a false positive/negative rate cannot be computed as reliably.

Tables 2.4 and 2.5 present the primary within-sample evidence concerning possible misspecification in the probability model. From (15), $P_{t+k|t}$ is the conditional expectation of R_{t+k} given data through t. One way to test whether the estimated probabilities satisfy this condition is to ask whether the errors $R_{t+k} - P_{t+k|t}$ can be predicted as linear functions of the observable indicators. This is done using regressions of the form

$$R_{t+k} - P_{t+k|t} = \alpha + \beta(L)z_t + e_t, \tag{19}$$

where z_t denotes an indicator observable at time t, transformed to be stationary so that conventional asymptotic theory can be used to interpret the regression results. Under the null hypothesis that the model (1), (2), (4), and (5) and the algorithm in section 2.1.2 are correctly specified, α and $\beta(L)$ will equal zero. Because R_t is a probability, e_t in (19) will be heteroskedastic, having a conditional variance under the null of $P_{t+k|t}(1 - P_{t+k|t})$. In addition, the k-step-ahead forecast error will be serially correlated. Were R_t observable at t, under the null hypothesis e_t would be MA($k - 1$); however, because turning points are declared only with a delay (typically of six to eighteen months), the order of the dependence of e_t is presumably greater.

Results of specification tests based on (19) are presented in tables 2.4 and 2.5.[3] In table 2.4, the p-values are computed by estimating (19) by OLS and computing heteroskedasticity- and autocorrelation-robust standard errors. Because the errors are conditionally heteroskedastic, table 2.5 reports p-values based on weighted least squares (WLS) regressions, where the weights are based on the conditional variance under the null, $P_{t+k|t}(1 - P_{t+k|t})$, and p-values were computed using an autocorrelation-robust covariance matrix.[4]

3. The dates of the cyclical peaks and troughs used to construct R_t for the subsequent empirical analysis are the official dates of the NBER Business Cycle Dating Committee, with one exception: the committee dated the 1969 cyclical peak as 1969:12, while throughout we use 1969:10. According to the recession definition in sec. 2.1.2, the earlier date is more consistent with the rules used to define the other historical turning points, and 1969:10 was the date used to estimate the model and to produce the results in Stock and Watson (1989).

4. The p-values ignore complications associated with the correlation between sampling error in the estimated parameters of the model and the regressors in (19).

The tests in tables 2.4 and 2.5 are computed using the data as revised through October 1988, with the exception of the series labeled "additional indicators," for which the data as revised through 1991:2 are used. (See the appendix for definitions of and sources for the series.)

The first blocks of tables 2.4 and 2.5 examine the first type of misspecification, in which the coincident and leading variables in the model might have predictive content for $P_{t+k|t}$. Because these included variables have no predictive content for the errors from the linear part of the model (1), (2), (4), and (5) (Stock and Watson 1989), rejections here would suggest misspecification in the definition of a recession. The XLI is also included in this panel. Aside from the regression on a constant, which reflects the bias discussed in the context of table 2.3, neither the OLS nor the WLS results indicate rejections at the 5 percent level at any horizons.

The next several blocks examine whether alternative leading and coincident indicators, not included in the model, have predictive content for R_{t+k} given $P_{t+k|t}$. The variables LPHRM through IVM3D8 were included in the original short list of fifty-five variables from which the seven included indicators were selected. Because the selection was done in the context of linear predictions of $c_{t|t}$, evidence of predictive content here would be evidence that the candidate variable has marginal value in predicting recessions and expansions, even though it does not in predicting overall economic growth rates. The results provide no strong evidence that the in-sample performance of the recession probabilities could have been improved by incorporating these indicators into the XRI model. If anything, the p-values tend to be rather high, reflecting the use of these indicators in the preliminary analysis.

The "additional indicators" in tables 2.4 and 2.5 are series arguably related to the 1990 downturn but not on the original short list of fifty-five leading indicators. These indicators will be examined in more detail in section 2.4 below; the relevant point here is that, on the basis of the 1962:1–1988:9 sample, taken individually none provide a significant improvement in the performance of the recession probabilities.

The final block of results in tables 2.4 and 2.5 examines the marginal predictive content of the DOC Composite Index of Leading Indicators (DLEAD) and of various nonlinear cyclical measures. Like the variables that compose it, the DOC leading index makes an insignificant contribution. There is some evidence that a variable constructed using the "three consecutive declines" rule of thumb, in which a recession is signaled when the DOC leading index declines for three consecutive months, has some marginal predictive content for long horizons and that such a rule of thumb, applied to IP alone, has marginal predictive value for short horizons.

The final three variables examine the possibility that the business cycle exhibits duration dependence. Cyclical duration dependence has been examined by Neftci (1982) and others, most recently including Diebold, Rudebusch, and Sichel (chap. 6 in this volume). The linear model, combined with the pattern recognition approach to identifying recessions used here, assumes that there is no duration dependence in expansions and recessions beyond that implied by the minimum six-month lengths of the events $D_{i\tau}$ and $U_{i\tau}$. This assumption can be checked by examining whether variables related to the duration of the current expansion/recession have additional predictive content. An obvious candidate variable is the duration M_t of the current expansion or recession. Although this is not known at time t with certainty because the dating committee identifies turning points only ex post, it can be estimated using the model of section 2.1. Let $M^r_{t|t}$ be the expectation of M_t in month t, conditional on being in a recession, that is, $M^r_{t|t} = E(M_t|R_t = 1, x_t, x_{t-1}, \ldots, y_t, y_{t-1}, \ldots)$; similarly define $M^e_{t|t}$ for expansions; and let $M_{t|t} = M^r_{t|t} P_{t|t} + M^e_{t|t} (1 - P_{t|t})$ be the expected length of the current spell whether or not it is a recession. The time series $M^r_{t|t}$, $M^e_{t|t}$, and $M_{t|t}$ were estimated using the model of section 2.1.[5]

The results for these duration dependence variables provide some evidence of this form of nonlinearity. Both the OLS and the WLS results suggest that $M^r_{t|t}$ (MTREC in tables 2.4 and 2.5) is a useful predictor for short forecast horizons; the WLS results indicate that $M^e_{t|t}$ (MTEXP) is a useful predictor for the six-month horizon as well. Thus, there appears to be some potential misspecification associated with the duration of recessions. There are several possible sources of this misspecification; for example, the linear model (1), (2), (4), and (5) might incorrectly ignore nonlinear feedback, perhaps from R_t, or the linear model might be correctly specified but the recession definition itself (i.e., the process by which recessionary patterns are identified in time series)

5. $M^r_{t|t}$ was constructed as follows. Using the algorithm in sec. 2.1.2, generate a historical realization of $\tilde{c}_t(-m - 8, 8)$, and, using adjacent septendecimtuples $\tilde{c}_\tau(-8, 8)$, classify each month for $\tau = t - m, \ldots, t$ as being in a recession or an expansion. This results in a vector of pseudorandom realizations of $(R_{t-m}, \ldots, R_t)$, constructed using data through t. In the computations, $m = 12$, and historical (true) values of R_t were appended for $\tau < t - 12$. Let L^r_{ti} be the length of the final string of 1s through time t (t = 1962:1, . . . , 1988:9) if $R_t = 1$, and let L^e_{ti} be the length of the final string of 0s if $R_t = 0$. Then $M^r_{t|t}$ is the average of L^r_{ti} over the $R_t = 1$ Monte Carlo draws and similarly for $M^e_{t|t}$. This construction provides an approximation to the joint conditional distribution of $(R_{t-m}, \ldots, R_t)$ or to the distribution of functions of these random variables such as M_t. This approximation, however, has two difficulties. First, because $b_{r,t}$ and $b_{e,t}$ are treated as random and varying over t, the event U_τ computed using $b_{r,t}$ differs from U_τ computed using $b_{r,t+1}$ (say). Second, even if $b_{r,t}$ and $b_{e,t}$ were constant ($\zeta_t = \zeta$), the marginal distribution of R_t constructed using this procedure will differ from that based on the algorithm in sec. 2.1.2: the marginals in sec. 2.1 are implicitly

$$\Pr[R_t|R_t \cup E_t] = \Pr\{R_t|R_t \cup E_t, R_\tau \cup E_\tau \cup [\Re^{17}/(R_\tau \cup E_\tau)], \tau \neq t\}$$

(where $\Re^{17}/[R_\tau \cup E_\tau]$ is the complement of $R_\tau \cup E_\tau$ in $\Re^{17}$), while those computed here are

$$\Pr[R_t|R_t \cup E_t] = \Pr[R_t|R_t \cup E_t, R_\tau \cup E_\tau, \tau \neq t].$$

It should be emphasized that this difficulty arises only when computing joint probabilities, not when computing sequences of marginal probabilities (e.g., $P_{t+k|t}$, $k = 0, 1, 2, \ldots$). Resolving this issue awaits further research.

have a temporal dependence that is not captured by the pattern recognition algorithm of section 2.1.2. Ascertaining which if either of these possibilities produces these rejections must await future research.

Taken together, these results suggest that there is little in-sample evidence of misspecification associated with the inefficient use of information in the included indicators or in candidate alternative leading indicators. Although there is some evidence of nonlinear misspecification related to duration dependence, the evidence is strongest at short forecasting horizons, and, in any event, this misspecification is not well proxied by any of the alternative leading indicators.

2.3 Out-of-Sample Performance

The XLI model was estimated using data through 1988:9. This section examines the performance of the XLI and the XRI over the period from 1988:10 through 1991:10, the month for which the most recent data were available to us. This provides thirty-seven months, including a cyclical peak in July 1990, with which to assess the performance of the indexes and to draw conclusions concerning the modification of the indexes.

2.3.1 Out-of-Sample Performance: An Overview

Forecasts of the growth in the XCI (annualized growth rates) made using the XLI model since 1988:1 are plotted in figure 2.2 for forecasting horizons of three, six, and nine months. The six-month-ahead forecast (panel B) is the XLI (i.e., $c_{t+6|t} - c_{t|t}$). The estimated recession probabilities $P_{t+k|t}$ for $k = -2, 0, 1, 2, 3, 6$, are presented in figure 2.3 and table 2.6.

As figure 2.2 makes plain, it is useful to consider the performance of the XLI over two episodes: prior to the summer of 1990 (approximately 1990:5) and subsequently. In the first episode, the performance of the XLI was very good, forecasting both the slowdown in the spring of 1989 and the growth that followed. During the fall of 1988 and the winter of 1989, interest rates rose substantially, by many reports in conjunction with an attempt by the Federal Reserve Board to control inflation; for example, the six-month Treasury-bill rate rose from 7.5 percent in October 1988 to 8.85 percent in March 1989. With the easing of interest rates in the spring of 1989, the financial market indicators in the XLI became more optimistic: by July, the commercial paper–Treasury-bill spread had fallen to 58 basis points, just above its postwar average and well below its March peak of 113 basis points. With this decline in interest rates and spreads, the XLI forecast increased growth: on the basis of unrevised data (i.e., as the XLI was originally computed), the XLI for March was -1.1 percent, while, by July, the XLI had risen to 0.7 percent. The behavior of the XLI forecasts was broadly consistent with the overall outlook at the time as reported in the economic and financial press, which was one of

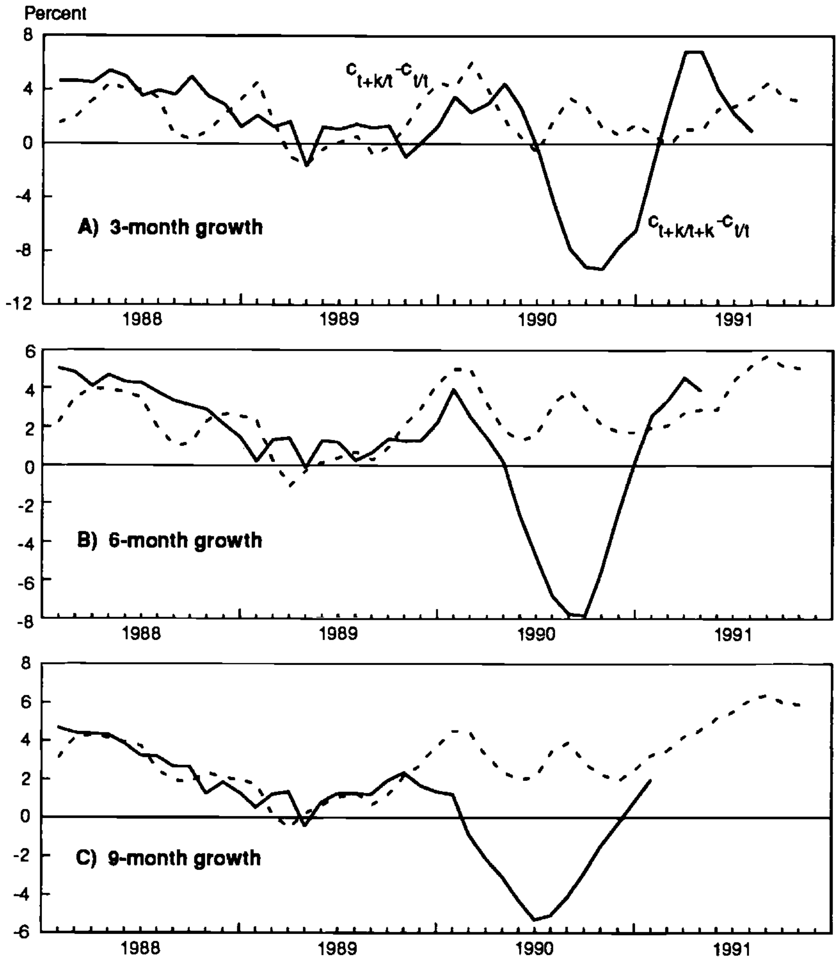

Fig. 2.2 Out-of-sample performance of the XLI model
Note: The series $c_{t+k/t} - c_{t/t}$ is based on preliminary data, and $c_{t+k/t+k} - c_{t/t}$ is based on revised data.

general concern over economic conditions in the early spring being replaced by cautious optimism in the late spring and early summer.[6]

As can be seen from figure 2.2, over this episode the XLI model provided very good forecasts of overall activity, not only at the six-month horizon for

6. For example, commenting on the 31 May 1989 release of the Department of Commerce's Leading Index in the *New York Times* (1 June 1989, C1), Michael P. Niemira of the Mitsubishi Bank stated, "The message is more strength still in the pipeline." The article later states, "Weakness in various measures of output and sales have signaled that economic growth is slowing and raised some concerns about a possible recession. The slowdown has also, however, raised hopes

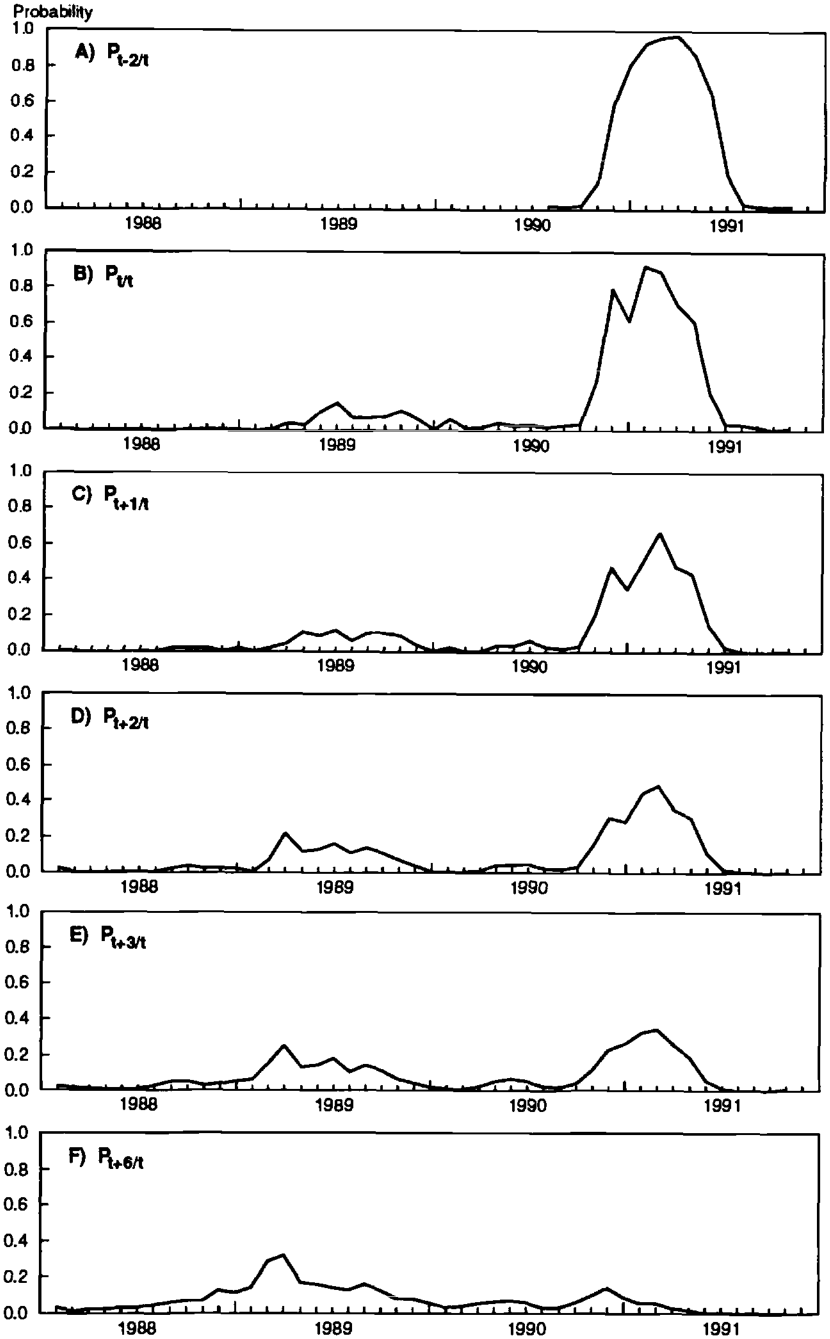

Fig. 2.3 Estimated recession probabilities

Note: The dates on the horizontal axes denote t, the date through which the data are available for computing $P_{t+k/t}$. The figure is based on data revised through 1988:9 and unrevised data since 1988:10. The series $P_{t-2/t}$ begins in 1990:7, and all other series begin in 1988:1.

Table 2.6 Estimated Recession Probabilities $P_{t+k|t}$, by Month

Date	$P_{t-2|t}$	$P_{t|t}$	$P_{t+1|t}$	$P_{t+2|t}$	$P_{t+3|t}$	$P_{t+6|t}$
1988:10	. . .	.01	.02	.02	.03	.07
1988:11	. . .	.01	.01	.02	.04	.13
1988:12	. . .	.01	.02	.02	.05	.11
1989:1	. . .	.00	.01	.01	.06	.14
1989:2	. . .	.01	.02	.07	.15	.29
1989:3	. . .	.04	.05	.22	.25	.32
1989:4	. . .	.03	.11	.12	.13	.17
1989:5	. . .	.10	.09	.13	.14	.16
1989:6	. . .	.15	.12	.16	.18	.14
1989:7	. . .	.07	.06	.11	.10	.13
1989:8	. . .	.07	.10	.14	.14	.16
1989:9	. . .	.08	.10	.11	.11	.13
1989:10	. . .	.11	.09	.07	.07	.09
1989:11	. . .	.07	.04	.04	.05	.09
1989:12	. . .	.01	.01	.01	.02	.06
1990:1	. . .	.06	.03	.01	.01	.03
1990:2	. . .	.02	.01	.00	.01	.05
1990:3	. . .	.02	.01	.01	.02	.05
1990:4	. . .	.04	.04	.05	.05	.06
1990:5	. . .	.03	.03	.05	.06	.07
1990:6	. . .	.04	.06	.05	.05	.05
1990:7	.03	.03	.03	.03	.02	.03
1990:8	.02	.04	.02	.02	.02	.03
1990:9	.03	.04	.03	.04	.05	.06
1990:10	.16	.28	.20	.16	.12	.10
1990:11	.59	.80	.48	.31	.23	.14
1990:12	.81	.62	.35	.28	.26	.09
1991:1	.93	.93	.52	.45	.33	.05
1991:2	.97	.87	.67	.49	.35	.04
1991:3	.98	.70	.48	.35	.26	.03
1991:4	.88	.61	.44	.31	.19	.02
1991:5	.65	.22	.15	.11	.05	.01
1991:6	.20	.04	.03	.02	.01	.01
1991:7	.04	.04	.01	.01	.01	.01
1991:8	.02	.02	.01	.01	.01	.01
1991:9	.02	.01	.01	.01	.01	.01
1991:10	.02	.01	.01	.01	.01	.01

Note: Recession probabilities were computed using unrevised (original) data.

which it had been optimized, but also at the three- and nine-month horizons. During this episode, the XRI indicated an increased probability of a recession: the XRI peaked at 32 percent in March 1989 and then quickly declined.

that the Federal Reserve might ease the tight grip it has kept on monetary policy for more than a year." In the *Wall Street Journal* that same day (p. A2), Gary Ciminero of Fleet/Norstar Financial Group was quoted as saying, "I think it [the DOC Leading Index] means that if we do encounter a

The second episode starts in the summer of 1990. On the basis of data through March 1990, the XLI was 3.1 percent, down from almost 5 percent in January and February 1990. In comparison, the XCI growth over the six months from March to September was 1.6 percent (annual rate), a forecast error of 1.5 percent, similar to previous out-of-sample performance and only approximately 1 percentage point in GNP units. The three-month-ahead forecast based on data through June 1990, for June–September, was -0.8 percent (annual rates); actual growth in the XCI over this period was -0.4 percent. However, the slowdown—correctly predicted over the next three months—was predicted to be short, to be followed by positive but slow growth. As a consequence, the recession probability—computed for each future month using data through June 1990—remained low, only 5 percent for each month from August through February 1991. This forecast of moderate growth in the fall of 1991 was, as it turned out, dramatically wrong: the XLI computed in August was 3.6 percent, while the actual growth of the XCI over this period was -7.3 percent, a forecast error of 10.9 percent (over 6 percent in GNP units at an annual rate). At the time of this writing (December 1991), the XLI appears to be back on track: the XCI increased by 3.9 percent at an annual rate between April and October 1991 (the most recent month for which data are available), and in April the XLI predicted that this growth would be 2.9 percent.

The performance within sample and during these two out-of-sample episodes is summarized in table 2.7 in terms of the RMSE and mean absolute errors (MAEs) of the forecasts.[7] The table shows that, during the first episode (1988:10–1990:4), the out-of-sample performance of the XLI was noticeably better than expected on the basis of the in-sample experience, with RMSEs and MAEs half what they were in sample. During the second episode (1990:5–1991:4; 1991:4 is the final month for which $c_{t+6|t+6}$ has been observed), forecast errors were approximately two times as large as within sample.

As can be seen from figure 2.3, the XRI has continued to estimate a six-month-ahead recession probability of under 20 percent; the XRI missed the July 1990 peak. It should be emphasized, however, that shorter-run forecasts indicated an increased probability of a recession, although not until October or November. For example, $P_{t|t}$, computed using data through October, was 28 percent; computed using data through November, it was 80 percent. Even so, the probability of a recession declined sharply with the horizon; in November, the three-month-ahead recession probability was only 23 percent.

An initial possibility is that the XLI continued to be a good forecast of

more significant slowdown in the economy, it's not going to occur in the next few months. I think we'll encounter a recession at the start of next year."

7. Note that k-month-ahead forecasts made during $1988{:}10 - k, \ldots, 1988{:}9$, $k \geq 1$, are partly out of sample, even though they are not included in the span used to compute the out-of-sample summary statistics in table 2.7.

Table 2.7 **Performance of the XLI: Summary Statistics**

Sample Period	RMSE	MAE
	A. Forecasting Performance of the XLI	
1962:1–1988:9	2.89	2.32
1980:1–1988:9	3.50	2.94
1980:1–1990:8	3.54	2.83
1988:10–1990:8	3.72	2.32
1988:10–1990:4	1.35	1.13
1990:5–1991:4	6.26	4.96
	B. Relation between XCI Growth and GNP	
1962:I–1988:III	2.34	1.84
1988:IV–1990:IV	1.32	1.07

Note: Panel A: The root mean square error (RMSE) and mean absolute error (MAE) are computed for the difference between the XLI ($c_{t+6|t} - c_{t|t}$) and the 6-month growth in the XCI ($c_{t+6|t+6} - c_{t|t}$). The dates in the first column correspond to the date that the forecast was made, so 1991:4 corresponds to the last observation for which there are data on $c_{t+6|t+6}$. Panel B: The statistics are computed for the residual from a regression (estimated over 1962:I–1988:III) of the quarterly growth of real GNP at annual rates on the quarterly growth of the XCI, where the XCI growth is the quarter-to-quarter growth of the monthly XCI, averaged across the months in the quarter.

economic activity but that the relation between overall economic activity (say, real GNP) and the XCI had deteriorated since 1988:9. This possibility is, however, readily dismissed: the out-of-sample relation between the XCI and real GNP was, if anything, closer than it had been in sample. This is documented in the final rows of table 2.7. Although there are currently only two observations on the quarterly XCI and real GNP during the second episode, the relation between the two appears to have been stable: the residuals from the 1962:I–1988:III regression of quarterly GNP growth onto quarterly XCI growth in table 2.2 yields forecast errors of -1.2 percent and 0.9 percent for 1990:III and 1990:IV, respectively, less than either the in-sample RMSE or the in-sample MAE.

The remainder of this paper explores possible explanations for the failure of the XLI and the XRI to predict the 1990 recession, with the objective of drawing lessons from this experience to guide revisions of the index and, more generally, future research on leading indicators. This section first documents the contributions of the individual variables to the XLI over 1990, then examines the importance of data revisions during the second episode. Section 2.4 turns to more fundamental issues of specification, construction of the index, and the choice of leading indicators.

2.3.2 Contributions of Individual Indicators to the Overall Index

The XLI model as described in section 2.2 is linear in the data and in c_t; as a result, $c_{t+k|t}$ can be written as a linear projection onto current and past values of the observable series

$$c_{t+k|t} - c_{t|t} = \lambda k + \sum_{i=1}^{11} A_{ki}(L)z_{it}, \tag{20}$$

where $z_t = (\Delta x_t', y_t')$ denotes the vector of four coincident and seven leading indicators in the XLI, where $A_{ki}(L)$ are lag polynomial weights, and where λ is a trend growth rate. These weights $A_{ki}(L)$ are readily computed numerically and are plotted in Stock and Watson (1989). The weighted averages $A_{ki}(L)z_{it}$ constitute the contribution of each of the eleven indicators to the deviation of the k-step-ahead forecast from its mean λk. An examination of these contributions for $k = 6$ therefore shows how each of the variables influenced the performance of the XLI on a month-by-month basis.

These historical contributions to the index are plotted in figure 2.4 over 1988:1–1991:10 and are presented in table 2.8 for January 1990 hence. Through the summer of 1990, the coincident indicators made negligible contributions to the index, usually 0.3 percent or less. The largest contributions to the index typically were made by building permits, exchange rates, the public-private yield spread, and the Treasury-bond yield spread.

This pattern changed during the second half of 1990. Although the public-private spread variable made positive contributions during July and August, its contribution in September and October was approximately zero. This was consistent with the doubling of this spread over this period, from thirty-nine basis points in July to seventy-nine basis points in November. The largest negative contributions came from housing authorizations and part-time work due to slack work. Three variables typically made substantial, incorrectly positive contributions to the index: the Treasury-bill yield curve, exchange rates, and, since December 1990, industrial production. The positive contribution of industrial production during this period reflects a mean reverting component in the model after the sharply negative values of IP in October and November. The yield curve and exchange rate contributions suggest that these variables might be partly to blame for the poor performance of the index. However, even if the contributions of these two variables are eliminated, then the XLI in September would still have been 1.7, while in fact the XCI declined by over 7 percent over this period.

In the winter and early spring of 1991, housing permits, unfilled orders, and part-time employment continued to provide negative signals. The appreciation of the dollar during the first quarter led exchange rates to provide a negative contribution to the XLI. The yield curve continued to steepen as short-term interest rates fell more quickly than long-term rates, leading to an

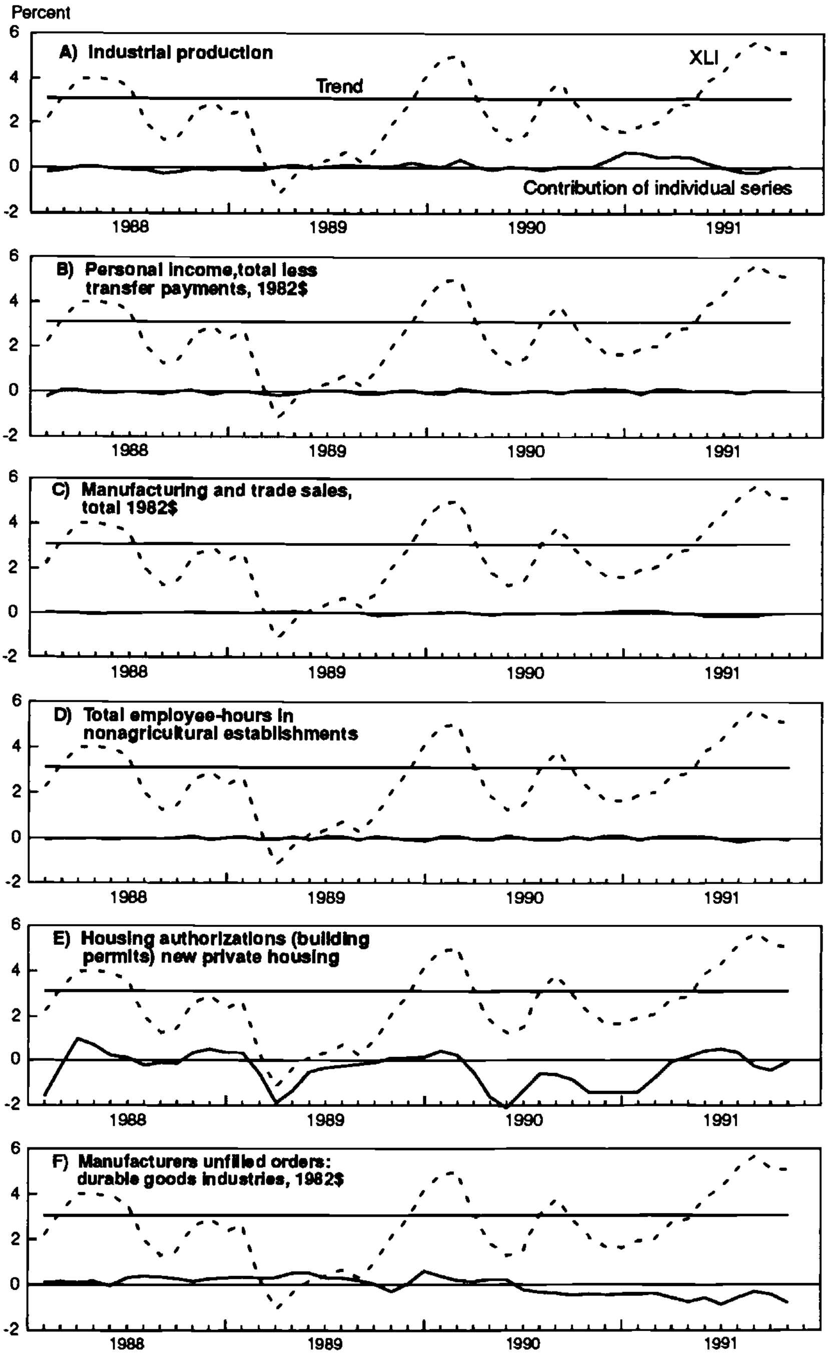

Fig. 2.4 Historical decomposition of the XLI

Note: The XLI and the historical contributions are based on preliminary data.

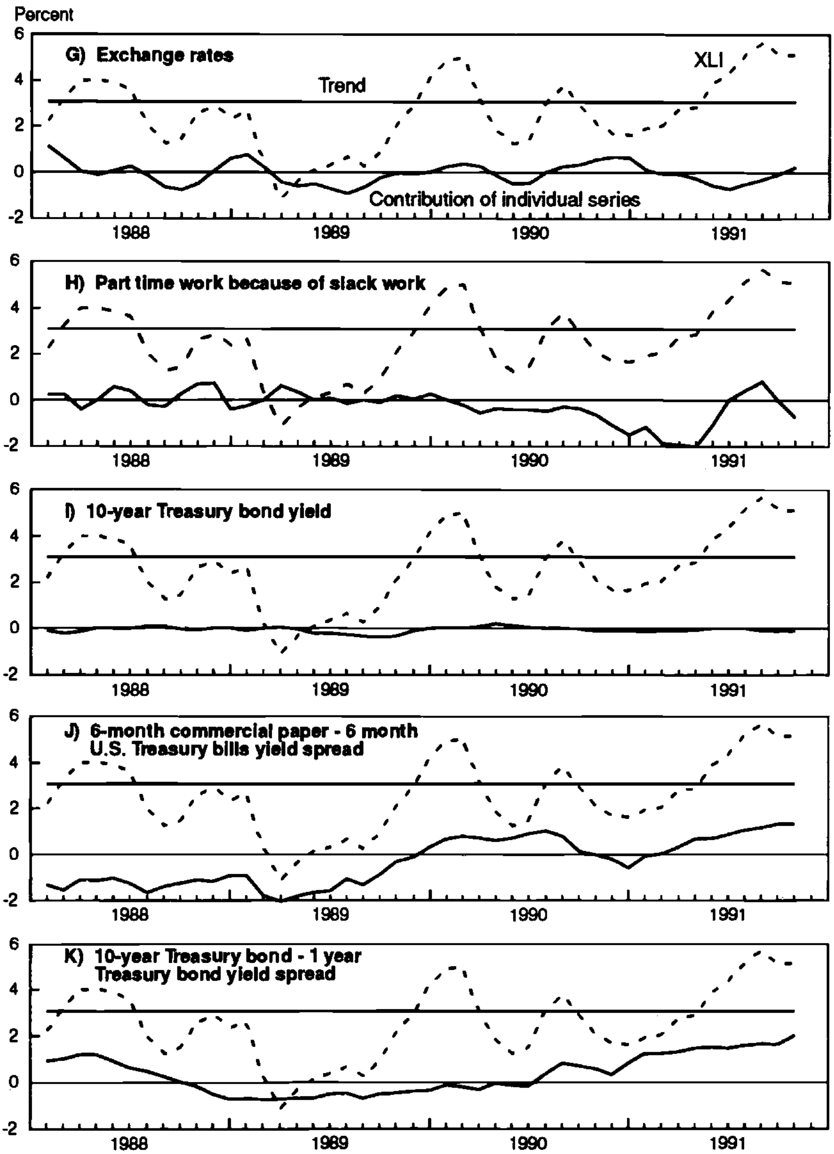

Fig. 2.4 (continued)

even more positive contribution from the yield curve spread. Finally, the commercial paper–Treasury-bill spread fell below its historical average, providing a positive contribution to the XLI. By the middle of 1991, all the indicators except exchange rates and unfilled orders were providing positive signals and suggesting stronger-than-average short-term growth in the XCI.

Table 2.8 Historical Contributions to the XLI, January 1990–February 1991

Month	IP	GMYXP8	MT82	LPMHU	HSBP	MDU82S	EXNWT2FS	LHNAPSS	FYGT10FS	CP6_GM6F	G10_G1F	XLI	$c_{t+6\|t+6} - c_{t\|t}$
1990:1	.0	−.0	.0	.0	.5	.4	.3	.0	.0	.7	−.1	5.0	4.0
1990:2	.4	.1	.0	.0	.3	.2	.4	−.2	.0	.8	−.2	5.0	2.5
1990:3	.0	.0	−.0	−.0	−.5	.1	.3	−.5	.1	.7	−.3	3.1	1.4
1990:4	−.0	−.0	−.0	−.0	−1.6	.2	−.1	−.4	.2	.6	−.0	1.8	.1
1990:5	.0	−.0	.0	.0	−2.1	.2	−.5	−.4	.2	.8	−.1	1.3	−2.6
1990:6	.0	.0	.0	.0	−1.4	−.3	−.4	−.4	.0	.9	−.1	1.5	−4.7
1990:7	−.0	.0	−.0	−.0	−.6	−.3	.0	−.5	.0	1.0	.4	3.1	−6.8
1990:8	.0	−.0	−.0	−.0	−.6	−.4	.3	−.3	.0	.8	.8	3.8	−7.7
1990:9	.0	.0	−.0	.0	−.9	−.5	.4	−.4	−.0	.2	.8	2.9	−7.8
1990:10	.0	.0	.0	−.0	−1.4	−.4	.6	−.6	−.0	.0	.6	2.1	−5.5
1990:11	.3	.2	.0	.1	−1.4	−.5	.7	−1.1	−.1	−.2	.4	1.7	−2.4
1990:12	.7	.1	.0	.1	−1.4	−.4	.7	−1.5	−.0	−.6	.8	1.6	.2
1991:1	.7	−.0	.1	−.0	−1.4	−.4	.2	−1.2	−.1	−.0	1.3	1.9	2.6
1991:2	.5	.1	.0	.0	−.7	−.4	−.0	−1.8	−.1	.0	1.3	2.1	3.4
1991:3	.5	.1	−.0	.0	−.0	−.6	−.0	−1.9	−.1	.4	1.3	2.8	4.6
1991:4	.5	.0	−.0	.0	.2	−.8	−.2	−2.0	−.0	.7	1.5	2.9	3.9
1991:5	.2	.0	−.0	.0	.5	−.6	−.5	−1.1	.0	.7	1.5	3.9	. . .
1991:6	.0	.0	−.1	−.0	.5	−.9	−.7	.0	.0	.9	1.4	4.4	. . .
1991:7	−.1	−.0	−.0	−.1	.4	−.6	−.5	.4	.0	1.1	1.6	5.2	. . .
1991:8	−.2	.0	−.0	−.0	−.2	−.3	−.3	.8	−.0	1.2	1.7	5.7	. . .
1991:9	.0	.0	−.0	.0	−.4	−.4	−.0	−.0	−.1	1.4	1.6	5.2	. . .
1991:10	.0	.0	−.0	−.0	−.0	−.8	.2	−.7	−.1	1.4	2.0	5.1	. . .

Note: The decompositions and the XLI are based on unrevised (original) data. The contributions are deviations from trend; the sum of the contributions, plus a trend of 3.1 percent, equals the XLI. For a discussion of the decompositions, see the text.

To understand the behavior of the XLI during the second half of 1990, it is useful to contrast the behavior of the financial variables around the cyclical peak of July 1990 to their behavior just before the cyclical peaks in November 1973, January 1980, and July 1981. In each of these three periods, the yield curve was strongly inverted: the spread between ten- and one-year Treasury-bond yields (G10_G1) was, respectively, -0.61, -1.59, and -1.39 in the month before each of these cyclical peaks, while its July 1990 value was 0.38. Similarly, the commercial paper–Treasury-bill yield spread (CP6_GM6) was 1.60, 0.96, and 1.13 in the month before these peaks but only 0.43 in June 1990 (approximately its postwar average value). Thus, the strong negative signals given by these variables prior to the previous recessions were replaced by neutral or slightly positive signals during the summer of 1990. Although the corporate paper–Treasury-bill spread had increased to 0.79 by November 1990, this increase occurred only after the general slowdown in September and October had become apparent.

In summary, this analysis of the historical contributions during the onset of the 1990 recession suggests two observations. First, the financial variables—in particular, the yield curve spread and, to a lesser extent, exchange rates—gave optimistic signals throughout this episode, even in the late fall, when the general public perception was that a recession was inevitable. Second, although none of the other variables gave strong positive signals, only three—part-time work, building permits, and unfilled orders—gave negative signals, and these negative contributions were still moderate, particularly in the second and third quarters of 1990.

Whether there were other variables that would have predicted this recession had they been incorporated into the index is the topic of the next section of the paper. First, however, we turn briefly to a discussion of data revisions during this episode.

2.3.3 Revisions to the Coincident Indicators

Some of the revisions to the data on the coincident indicators during the fall of 1990 were large. Table 2.9 presents these data as they were released over this period, in terms of monthly growth at annual percentage rates. On the basis of recent revisions, we see that industrial production growth was positive through September; estimates of the decline in IP in October ranged from almost 11 percent (annual rate) in the 1990:11 data release to only 7.6 percent in the 1991:1 release. An examination of table 2.9 reveals comparable revisions in the other coincident indicators. Even though measurement error models are explicitly incorporated into the XLI model as described in section 2.1, large revisions in the coincident indicators can nonetheless result in substantial changes in $c_{t+k|t}$ for k small (say, $k \leq 3$). This raises the possibility that the poor performance of the XLI and the XRI was, at least in part, due to their reliance on these substantially revised preliminary data.

This possibility is examined in figure 2.5, which presents the XLI com-

Table 2.9 Coincident Indicator Data and Revisions, 1990:5–1991:2: Monthly Growth at Annual Rates

	Data Released During the Month Following:					
Data for:	1990:9	1990:10	1990:11	1990:12	1991:1	1991:2
	A. Industrial Production					
1990:5	6.60	6.60	6.60	6.60	6.60	6.60
1990:6	7.65	7.65	7.65	7.65	7.65	7.65
1990:7	2.18	3.27	3.27	3.27	3.27	3.27
1990:8	1.09	.00	1.09	1.09	1.09	1.09
1990:9	3.26	1.09	−1.09	1.09	1.09	1.09
1990:10	. . .	−9.81	−10.92	−8.71	−7.62	−7.62
1990:11	. . .	. . .	−21.02	−22.06	−18.71	−17.60
1990:12	. . .	. . .	. . .	−7.82	−13.38	−12.25
1991:1	. . .	. . .	. . .	. . .	−5.62	−6.74
1991:2	. . .	. . .	. . .	. . .	. . .	−10.18
	B. Personal Income, Total, Less Transfer Payments, 1982$					
1990:5	.17	.17	.17	.17	.17	.17
1990:6	1.36	1.36	1.36	1.36	1.36	1.36
1990:7	3.05	3.25	2.92	2.92	2.92	2.92
1990:8	−5.93	−4.90	−5.36	−5.36	−5.36	−5.36
1990:9	−3.14	−4.55	−4.34	−4.34	−4.34	−4.34
1990:10	. . .	−8.23	−11.32	−11.78	−12.96	−12.75
1990:11	. . .	. . .	−.75	.75	1.05	2.30
1990:12	. . .	. . .	. . .	6.51	7.05	5.09
1991:1	. . .	. . .	. . .	. . .	−17.60	−17.06
1991:2	. . .	. . .	. . .	. . .	. . .	−.21
	C. Manufacturing and Trade Sales, Total, 1982$					
1990:5	11.24	11.24	11.24	11.24	11.24	11.24
1990:6	7.43	7.43	7.43	7.43	7.43	7.43
1990:7	−4.81	−7.52	−7.53	−7.53	−7.53	−7.53
1990:8	11.13	14.91	14.07	14.09	14.09	14.09
1990:9	. . .	−22.84	−18.78	−22.19	−22.19	−22.19
1990:10	. . .	. . .	−1.38	−1.68	−2.24	−2.18
1990:11	. . .	. . .	. . .	−19.96	−16.31	−20.41
1990:12	. . .	. . .	. . .	. . .	−23.54	−23.38
1991:1	. . .	. . .	. . .	. . .	. . .	−14.94
1991:2	. . .	. . .	. . .	. . .	. . .	. . .
	D. Total Employee-Hours in Nonagricultural Establishments					
1990:5	9.84	9.84	9.84	10.01	10.01	10.01
1990:6	5.65	5.65	5.65	9.23	9.23	9.23
1990:7	−3.32	−3.32	−3.32	−3.56	−3.56	−3.56
1990:8	−4.06	−4.71	−4.71	−4.74	−4.74	−4.74
1990:9	5.20	5.73	5.93	5.96	5.96	5.96

Table 2.9 (continued)

	Data Released During the Month Following:					
Data for:	1990:9	1990:10	1990:11	1990:12	1991:1	1991:2
1990:10	. . .	−15.24	−16.93	−17.39	−17.39	−17.39
1990:11	. . .	. . .	3.85	4.13	4.13	4.13
1990:12	. . .	. . .	. . .	6.01	6.00	4.36
1991:1	. . .	. . .	. . .	. . .	−18.88	−15.77
1991:2	. . .	. . .	. . .	. . .	. . .	3.71

Source: Department of Commerce Electronic Bulletin Board, various releases.
Note: Data on manufacturing and trade sales are available with a two-month lag.

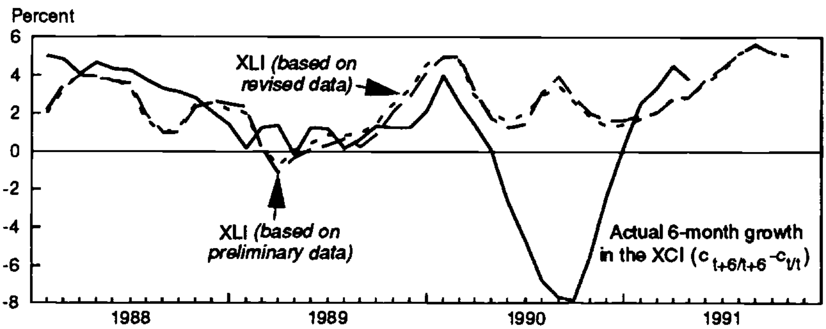

Fig. 2.5 Original and revised XLI
Note: The XLI and the growth in the SCI are at annual rates.

puted using the unrevised and the revised data as well as the actual growth of the XCI based on the most recently available revised data. Although the revisions in the coincident indicator data are large, these revisions had scant effect on the XLI: the change in the XLI based on the revised data is typically less than 0.4 percentage points. One explanation for this is that the XLI relies in large part on leading variables not subject to revision, in particular, the financial variables; another is that the predictive role of the coincident indicators, although substantial for very short horizons, diminishes markedly for longer horizons. In any event, the large revisions to the coincident indicators over the summer and fall of 1990 do not seem to be the source of the breakdown in the XLI and XRI forecasts.

2.4 Alternative Specifications: Recent Performance

This section investigates the failure of the XLI to predict the downturn in the fall of 1990. The analysis centers around two main possibilities. The first is that, given the list of indicators selected, the model was incorrectly speci-

Table 2.10 **Sensitivity of $P_{t+6|t}$ to μ_r, μ_e, and σ_ζ: RMSE for $R_{t+6} - P_{t+6|t}$, in Sample (1962:1–1988:9) and out of Sample (1988:10–1990:8)**

$\mu_r \backslash \mu_e$	.000	−.250	−.500	$\mu_r \backslash \mu_e$	.000	−.250	−.500
	A. σ_ζ = .6				C. σ_ζ = 1.0		
−1.250	.250	.248	.250	−1.250	.250	.247	.250
	.486	.483	.484		.485	.484	.482
−1.500	.249	.250	.250	−1.500	.249	.250	.250
	.486	.486	.476		.486	.486	.486
−1.750	.250	.251	.251	−1.750	.249	.251	.253
	.485	.486	.484		.483	.486	.485
	B. σ_ζ = .8				D. Additional RMSEs		
−1.250	.250	.249	.250		$\mu_r = 0$, $\mu_e = 0$,		.258
	.486	.483	.482		$\sigma_\zeta = .8$:		.447
−1.500	.249	.250	.249		$\mu_r = 0$, $\mu_e = 0$,		.258
	.485	.486	.485		$\sigma_\zeta = 1.0$:		.489
−1.750	.250	.250	.250				
	.484	.487	.485				

Note: The upper entry in each (μ_r, μ_e, σ_ζ) cell is the in-sample RMSE; the lower entry is the out-of-sample MSE. For these calculations, the most recent turning point was the cyclical peak of 1990:7.

fied or "overfit," in the sense that it was too heavily parameterized. The second is that the list of indicators was flawed: had other indicators been included, would the XRI have predicted a recession?

These two questions are addressed in several steps. We first consider the effect of possible modifications of the definition of a recession on both the in-sample and the out-of-sample performance of the XRI. Because the main source of the failure in the XRI stems from the overly optimistic forecasts embodied in the XLI in August and September, it is not surprising that tuning the recession definition does not substantially improve the performance of the XRI. Next, we consider the possibility that the model is overparameterized, resulting in "overfitting" the in-sample data. This is examined by studying an alternative set of indexes, based on the seven leading indicators in the XLI, in which the number of estimated parameters is reduced and the structure of the index is simplified substantially. The performance of these indexes is comparable to the XLI, both in and out of sample, which leads us to conclude that overparameterization or model misspecification does not account for the poor performance of the XLI. The analysis therefore turns to the performance of the individual indicators constituting the index and of alternative leading indicators and indexes.

2.4.1 Changes in the Definitions of Recession/Expansion Events

The first possibility investigated is that the XRI would have performed better had it used a different definition of recessions. This is investigated, first, by considering the effect of changing the parameters (μ_r, μ_e, σ_ζ) that enter the definition of recession and expansion events and, second, by redefining recession events so that a recession can last only four months rather than six.

The results of these changes are summarized in table 2.10 and 2.11. Table 2.10 presents the RMSEs of the XRI forecast errors (i.e., $R_{t+6} - P_{t+6|t}$), both in and out of sample. The much larger RMSEs out of sample than in sample largely reflect the failure of the XLI to predict the downturn in the fall of 1990. In general, the RMSE, both in and out of sample, is insensitive to changes of $\pm$ .25 for μ_r and μ_e and of $\pm$ .2 for σ_ζ; minor modifications in the recession definition would not have resulted in more accurate recession probabilities.

Six-month-ahead recession probabilities for 1990, computed using alternative parameter values, are summarized in table 2.11. Although modest changes make negligible differences, increasing the recession cutoffs to μ_r =

Table 2.11 The Effect of Changing the Definition of a Recession on the XRI Probabilities ($P_{t+6|t}$) for January–December 1990

	(μ_r, μ_e, σ_ζ) =					
	(−1.5, −.25, .8)	(−1.75, −.25, .8)	(−1.5, .0, .8)	(−1.5, −.25, .6)	(.0, .0, 1.0)	4-Month Recessions
			A. Recession Probabilities			
Jan.	.021	.023	.016	.010	.041	.023
Feb.	.026	.026	.036	.044	.062	.031
Mar.	.047	.023	.035	.041	.098	.066
Apr.	.056	.040	.038	.040	.102	.077
May	.048	.030	.057	.037	.098	.076
June	.030	.021	.040	.034	.120	.082
July	.023	.024	.023	.020	.082	.045
Aug.	.016	.022	.041	.016	.061	.037
Sep.	.029	.037	.048	.034	.130	.065
Oct.	.092	.102	.070	.088	.182	.100
Nov.	.134	.105	.136	.129	.238	.125
Dec.	.095	.073	.097	.097	.243	.116
			B. RMSEs			
1962:7–1988:9	.250	.251	.249	.250	.258	.253
1988:10–1990:8	.486	.486	.485	.486	.489	.485

Note: Panel A: The entries are the values of the XRI ($P_{t+6|t}$) that would have been computed for each month during 1990, had the indicated values of (μ_r, μ_e, σ_ζ) been used. To facilitate comparisons, the first column reports the value of the XRI; the next four columns report probabilities computed with alternative parameter values. The final column reports probabilities computed using the parameters used in the model, but with the minimum length of a recession taken to be 4 rather than 6 months. Panel B: The entries are the RMSEs of $R_{t+k} - P_{t+k|t}$ over the indicated (in-sample and out-of-sample) ranges.

$\mu_e = 0$ would have produced slightly higher probabilities; the XRI would have registered 13 percent in September rather than 3 percent. However, cutoffs this high are implausible: the pre-1988:10 RMSEs for these parameters exceed those for the chosen parameter values, and, more important, because the XCI is more volatile than GNP, recession cutoffs of $\mu_r = \mu_e = 0$ approximately correspond to a recession occurring if real GNP growth drops below 1.3 percent for two consecutive quarters. In the past, the Business Cycle Dating Committee has dated recessions as if the appropriate cutoff is approximately zero growth in real GNP.

The final modification considered here is reducing the length of the shortest recession from six to four months, with the result that (6) is replaced by

$$(6') \quad D_{1\tau} = \{\Delta c_s, s = \tau - 3, \ldots, \tau: \Delta c_s \leq b_{r,s}, s = \tau - 3, \ldots, \tau\},$$

and similarly for $U_{1\tau}$. As seen in the final column of table 5.11, this change has little effect on the XLI, in terms of either the in-sample RMSE or the probabilities over 1990.

In short, modifications of the recession definition—even major ones, such as permitting recessions with a duration as short as four months or as shallow as growth dipping below 1.3 percent in the units of annual GNP growth—have little effect on the recession probabilities computed over this episode.

2.4.2 Possible Overparameterization and Overfitting

Because the linear model outlined in section 2.1.1 has a large number of parameters, it is possible that the poor performance during the fall of 1990 can be attributed to "overfitting" in the sense of having too few observations per parameter. This possibility is investigated by constructing some simple alternative indexes that entail fitting considerably fewer parameters. These indexes are of the form

$$(21) \qquad \hat{y}_t = \sum_{i=1}^{n} \alpha_i \hat{y}_{it},$$

where $\hat{y}_{it}$ are the indexes (forecasts) constructed from each of the n individual leading indicators entering the index. The individual forecasts are computed as the projection of the growth of the XCI over the next six months onto current and lagged values of each candidate leading indicator y_{it} and onto current and lagged values of XCI growth; that is, $\hat{y}_{it}$ is the fitted value from the regression

$$(22) \qquad c_{t+6|T} - c_{t|T} = \omega + \beta_i(L)y_{it} + \gamma_i(L)\Delta c_{t|t} + v_t.$$

An index constructed according to (21) and (22), like the XLI, is linear in the leading indicators and has a representation analogous to (20).[8]

8. For related work that approaches the construction of indexes of leading indicators as a forecasting problem, see Auerbach (1982), Stekler and Schepsman (1973), and Vaccara and Zarnowitz (1978).

The parameters ω, $\beta_i(L)$, and $\gamma_i(L)$ in (22) were estimated by running the regression in (22) with contemporaneous values and five lags of y_{it} and with contemporaneous values and two lags of $\Delta c_{t|t}$. If, in fact, $\Delta c_{t|t}$ and y_{it} follow a VAR(p), this regression would be inefficient relative to estimating a VAR(p), but this procedure has two advantages: first, it produces conditionally unbiased projections (conditional on lags of y_{it}) if the true projection is linear (with the specified lag length); second, it reduces the number of parameters that need to be estimated for multistep forecasting.

Because the individual "indexes" $\hat{y}_{it}$ are each forecasts of $c_{t+6|T} - c_{t|T}$, the problem of constructing a composite index—that is, computing the weights $\{\alpha_i\}$ in (21) given the set of individual indexes to be included—reduces to the well-studied problem of the combination of forecasts. Two simple approaches are used here. The first is analogous to the weighting scheme effectively used to construct the DOC Leading Index; that is, all the weights are set to $1/n$.[9] The second approach is to produce the minimal mean squared error linear combination of forecasts, which is implemented by estimating α_i by OLS with $c_{t+6|T} - c_{t|T}$ as the dependent variable and $\hat{y}_{it}$ as the independent variable. The net effect of constructing indexes using this simple structure is to reduce substantially the number of parameters to be estimated, relative to the model of section 2.1.1.

Composite indexes based on (21) and (22) were estimated using data over the period 1959:1–1988:9. The RMSEs of the resulting indexes are summarized in table 2.12, along with the corresponding RMSEs for the XLI. (These and subsequent RMSEs for indexes of the form [21] and [22] are computed relative to the growth in the smoothed XCI, $c_{t+6|T} - c_{t|T}$.) Relative to the OLS-weighted index with the same seven leading indicators, the XLI performs better both in and out of sample, but it performs slightly worse than the equally weighted index out of sample. Overall, the performance is comparable across these three indexes for all subsamples.

The same exercise was repeated for the leading indicators composing the XLI2 (the alternative nonfinancial leading index). As with the indexes based on the XLI indicators, the performance is similar across indexes in all subsamples. The implication is that the fitting of many parameters in the XLI (or XLI2) model does not appear to be a key factor in the breakdown in the fall of 1990.

A second conclusion emerging from table 2.12 is that, although the XLI2 does markedly worse than the XLI in sample, it noticeably outperforms the XLI out of sample, with a reduction of almost one-quarter in the RMSE for forecasts into the fall of 1990. This, along with the findings of the previous section concerning the insensitivity of the XLI to the recession definition, suggests that the problems with the XLI resulted from omitting leading indi-

9. For discussions of the DOC weighting schemes, see Zarnowitz and Boschan (1975a, 1975b) and Zarnowitz and Moore (1982).

Table 2.12 **Forecasting Performance of Alternative Composite Indexes: Six-Month-Ahead Forecast Horizon**

	RMSE Computed over:				MAE Computed over:			
Series	62:1–88:9	88:10–90:8	88:10–90:4	90:5–90:8	62:1–88:9	88:10–90:8	88:10–90:4	90:5–90:8
Constant	4.38	3.55	1.75	7.61	3.12	2.45	1.39	7.53
XLI	2.76	3.42	1.39	7.62	2.20	2.15	1.08	7.24
XLI-equal	3.57	3.36	1.31	7.54	2.61	2.16	1.06	7.39
XLI-OLS	2.78	3.76	1.54	8.36	2.25	2.36	1.14	8.14
XLI2	3.80	2.61	1.57	5.24	2.83	1.93	1.35	4.67
XLI2-equal	3.94	3.20	1.38	7.07	2.83	2.10	1.08	6.93
XLI2-OLS	3.38	3.03	1.72	6.23	2.64	2.15	1.40	5.73

Note: "Constant" indicates forecasting $c_{t+6} - c_t$ by a constant only. The alternative composite indexes XLI-equal, XLI-OLS, XLI2-equal, and XLI2-OLS are constructed according to eqs. (21) and (22) in the text. In the "equally weighted" indexes, the weights in (21) are $\alpha_i = 1/n$, while, in the "OLS-weighted" indexes, $\{\alpha_i\}$ are estimated by ordinary least squares. XLI indicators: HSBP, MDU82S, EXNWT2FS, LHNAPSS, FYGT10FS, CP6–GM6F, G10–GLF; XLI2 indicators: HSBP, MDU82S, EXNWT2FS, LPHRM, IPXMCA, LHEL, IVPAC. All indexes were computed using the revised data as of 1991:2. RMSEs were computed relative to $(c_{t+6|T} - c_{t|T})$. All growth rates are in percentages, at annual rates.

cators that turned out to have important predictive content for the 1990 downturn and including others that did not.

2.4.3 Alternative Indicators: Diagnostic Tests

The tests based on the regressions (19) provide a useful framework for checking whether alternative leading or coincident indicators would have been useful in predicting R_t over 1988:10–1991:2. The p-values for the resulting Wald tests are reported in table 2.13. Because of the limited number of out-of-sample observations, only current values and two lags of each candidate indicator are used in the regressions. Also, in this out-of-sample period, there are only three nonoverlapping nine-month periods, so the most distant horizon considered is six months. Because there are fewer than five nonoverlapping observations on $P_{t+6|t}$, the six-month horizon results need to be interpreted cautiously as well.

The results provide strong evidence of misspecification during this period. The weights given to the included indicators—in particular, housing starts, exchange rates, and the ten-year bond rate—were ex post incorrect. This accords with the message of figure 2.4 above: exchange rates and the interest rate indicators yielded overoptimistic predictions during the summer and fall of 1990, and the XLI (and the XRI) would have performed better had building permits been given more weight.

Table 2.13 also demonstrates that alternative indicators would have been useful in predicting the XRI forecast errors, in particular, help wanted advertising, stock prices, money and credit supply measures, and, at longer horizons, oil prices and some measures of investment, orders, consumption, and consumer expectations. These results demonstrate that, over this episode, the XRI forecast errors could have been reduced by using additional indicators and by placing different weights on those that were included. Because the forecast errors in the XRI were large, however, partially explaining these errors does not in the end seem to be a very demanding task. More challenging is seeing whether the alternative indicators might have provided satisfactory forecasts of this and earlier recessions, either alone or as part of an alternative index.

2.4.4 Alternative Indicators: Performance of Single-Indicator Indexes

This subsection presents some initial results analyzing the performance of alternative indicators and indexes since 1988:10. The indexes are of the form (21) and (22). They exploit the cross-covariance among the candidate leading indicators in only a limited way, and they do not readily produce a recession index such as the XRI. However, for the purposes of this section, indexes based on (21) and (22) have two practical advantages over those based on the framework in section 2.1: they are faster to compute, permitting an examination of a much richer initial list of indicators, and, when used in their equally weighted form, the contribution to the composite index of each of the candidate indicators is transparent.

Table 2.13 **Out-of-Sample Regression Tests for Omitted Variables in $P_{t+k|t}$ (p-values of test statistics): OLS Regressions, 1988:10–1991:2 − k**

	Forecast Horizon (Months)			
Variable	0	1	3	6
Constant	.096	.022	.000	.000
	Coincident Indicators			
IP	.098	.318	.294	.718
GMYXP8	.214	.100	.135	.369
MT82	.223	.910	.917	.269
LPMHUADJ	.346	.703	.390	.219
	Leading Indicators in the XRI			
HSBP	.000	.000	.009	.016
MDU82S	.154	.039	.580	.003
EXNWT2FS	.214	.001	.000	.000
LHNAPSS	.121	.500	.048	.978
FYGT10FS	.190	.111	.005	.000
CP6_GM6F	.102	.021	.011	.007
G10_GLF	.444	.291	.142	.561
XLI	.207	.014	.007	.000
	Leading Indicators in the XRI2			
LPHRM	.111	.615	.872	.103
IPXMCA	.289	.442	.954	.352
LHEL	.435	.507	.005	.000
IVPAC	.033	.389	.057	.575
	Financial Indicators			
FSPCOMF	.174	.006	.000	.000
FM1D82	.008	.088	.012	.000
FM2D82	.235	.535	.327	.033
FMBASE	.183	.381	.627	.131
FYFFF	.179	.142	.353	.002
BAA_G10F	.184	.313	.219	.000
YLD_DUMF	.115	.001	.025	.000
	Employment Indicators			
LUINC	.462	.217	.979	.550
LHU5	.620	.798	.093	.950
LHELX	.283	.234	.358	.082
	Consumption and Retail Sales			
IPCD	.056	.143	.004	.022
GMCD82	.672	.693	.236	.835
RTR82	.240	.276	.239	.022

Table 2.13 (continued)

	Forecast Horizon (Months)			
Variable	0	1	3	6
	Inventories and Orders			
MPCON8	.579	.366	.633	.528
MOCM82	.641	.667	.764	.012
MDO82	.577	.227	.691	.421
IVMT82	.403	.186	.133	.027
IVM1D8	.495	.018	.114	.845
IVM2D8	.315	.174	.495	.000
IVM3D8	.499	.255	.303	.883
	Additional Indicators			
DLBLNPAP	.095	.000	.027	.010
PMI	.235	.130	.370	.907
PMNO	.297	.440	.467	.492
HHSNTN	.453	.181	.817	.243
HHST	.493	.102	.265	.000
PW561	.829	.851	.209	.002
PW561R	.798	.837	.257	.001
FTM333	.439	.534	.017	.246
FTM333R	.423	.546	.028	.131

Note: The p-values refer to Wald tests of the hypothesis that the coefficients on (z_t, z_{t-1}, z_{t-2}) in the regression of $R_{t+k} - P_{t+k|t}$ on a constant and z_t, z_{t-1}, z_{t-2}, are zero, where k refers to the forecast horizon (months). See the notes to table 2.4. All results were computed using the most recently available data through 1991:2.

Table 2.14 presents RMSEs and MAEs for the "univariate indexes" $\hat{y}_{it}$, which are the forecasts produced by the regression (22) estimated over 1962:1–1988:9. The results over 1988:10–1991:2 provide out-of-sample evidence on each candidate indicator when considered one at a time. The most striking feature of table 2.14 is that, even though the XLI had an MAE of 7.2 percent over the period 1990:5–1990:8, with few exceptions this large forecast error is typical of those for the univariate forecasts. For example, although the stock market declined in August 1990 in anticipation of the economic downturn, this one correct signal was insufficient to provide the XLI with improved forecasting power: its MAE over the final episode was 7.2 percent. Forecasts based on inventories and orders all had larger MAEs than the XLI, and forecasts based on retail sales had MAEs of well over 6 percent. Certain financial variables—a yield curve dummy that performed well in sample and two indicators in the XLI (the public-private spread and the slope of the yield curve)—performed particularly poorly, relative to the other variables. Only five of the indicators in table 2.14—housing starts, help wanted advertising, real M2, the quarterly ratio of the volume of commercial paper to

Table 2.14 Performance of Single-Indicator Indexes: Six-Month-Ahead Forecast Horizon

	RMSE Computed over:				MAE Computed over:			
Series	62:1–88:9	88:10–90:8	88:10–90:4	90:5–90:8	62:1–88:9	88:10–90:8	88:10–90:4	90:5–90:8
XLI	2.76	3.42	1.39	7.62	2.20	2.15	1.08	7.24
XLI2	3.80	2.61	1.57	5.24	2.83	1.93	1.35	4.67
	Coincident Indicators							
IP	4.31	3.36	1.79	7.04	3.11	2.36	1.39	6.97
GMYXP8	4.27	3.46	1.87	7.24	3.11	2.55	1.57	7.17
MT82	4.27	3.45	1.47	7.63	3.08	2.27	1.17	7.51
LPMHUADJ	4.33	3.47	1.82	7.31	3.11	2.48	1.49	7.18
	Leading Indicators in the XLI							
HSBP	3.81	2.40	1.23	5.09	2.87	1.62	.97	4.71
MDU82S	4.25	3.32	1.83	6.89	3.01	2.42	1.48	6.86
EXNWT2FS	4.17	4.15	1.71	9.23	3.08	2.71	1.41	8.89
LHNAPSS	4.22	3.48	1.93	7.21	3.01	2.59	1.64	7.11
FYGT10FS	4.06	3.58	2.07	7.31	2.97	2.67	1.73	7.18
CP6_GM6F	2.97	4.20	1.79	9.30	2.33	2.72	1.35	9.21
G10_GLF	3.68	3.49	1.15	7.99	2.81	2.06	.86	7.79
	Leading Indicators in the XLI2							
LPHRM	4.33	3.36	1.59	7.28	3.12	2.26	1.22	7.20
IPXMCA	4.31	3.39	1.71	7.23	3.11	2.34	1.33	7.12
LHEL	3.95	2.59	1.60	5.14	2.83	1.78	1.12	4.90
IVPAC	4.09	4.16	2.10	8.87	3.02	2.93	1.69	8.81
	Financial Indicators							
FSPCOMF	3.92	3.99	2.87	7.22	2.92	3.26	2.43	7.16
FM1D82	3.96	3.10	1.57	6.59	2.91	2.14	1.23	6.47
FM2D82	3.54	2.46	1.61	4.74	2.61	1.89	1.32	4.60
FMBASE	4.33	3.50	1.64	7.59	3.08	2.26	1.17	7.45
FCBCUCY	4.33	3.74	1.99	7.84	3.10	2.74	1.68	7.78
FYFFF	3.60	3.18	1.08	7.24	2.49	1.87	.77	7.12
BAA_G10F	4.14	3.60	2.04	7.41	2.94	2.61	1.63	7.29
YLD_DUMF	3.44	4.77	2.65	9.86	2.57	3.55	2.24	9.77
	Employment Indicators							
LUINC	4.20	3.57	1.64	7.78	3.01	2.42	1.30	7.71
LHU5	4.27	3.78	1.95	8.00	3.04	2.71	1.61	7.95
LHELX	3.96	2.98	1.35	6.52	2.85	2.01	1.09	6.41

Table 2.14 (continued)

	RMSE Computed over:				MAE Computed over:			
Series	62:1–88:9	88:10–90:8	88:10–90:4	90:5–90:8	62:1–88:9	88:10–90:8	88:10–90:4	90:5–90:8
	Consumption and Retail Sales							
IPCD	4.29	3.79	1.73	8.28	3.08	2.55	1.38	8.11
GMCD82	4.29	3.25	1.61	6.97	3.06	2.22	1.24	6.88
RTR82	4.26	3.09	1.44	6.72	3.04	2.11	1.17	6.58
	Inventories and Orders							
MPCON8	4.34	3.56	1.81	7.58	3.09	2.50	1.46	7.48
MOCM82	4.29	4.00	2.08	8.44	3.05	2.80	1.63	8.32
MDO82	4.30	3.76	2.16	7.68	3.05	2.68	1.67	7.46
IVMT82	4.32	3.99	1.94	8.60	3.08	2.70	1.49	8.47
IVM1D8	4.22	4.28	2.46	8.73	3.10	3.18	2.02	8.68
IVM2D8	4.32	3.80	1.82	8.20	3.13	2.58	1.42	8.08
IVM3D8	4.35	3.62	1.83	7.72	3.12	2.50	1.42	7.64
	Additional Indicators							
DLBLNPAP	4.27	2.80	1.35	6.05	3.08	1.95	1.11	5.98
PMI	4.21	3.46	1.60	7.54	3.08	2.30	1.21	7.49
PMNO	4.05	3.03	1.70	6.24	2.94	2.29	1.48	6.17
HHSNTN	4.00	3.19	2.21	5.95	2.87	2.52	1.86	5.67
HHST	4.08	3.82	2.63	7.14	2.90	3.17	2.36	6.98
PW561	4.22	4.07	2.19	8.51	3.12	2.93	1.82	8.24
PW561R	4.26	4.02	2.19	8.37	3.13	2.84	1.75	8.01
FTM333	4.30	3.45	1.23	7.84	3.05	2.13	.95	7.72
FTM333R	4.28	3.59	1.38	8.06	3.05	2.27	1.07	7.94

Note: All results were computed using the data as revised through 1991:2. See the notes to table 2.12.

bank loans, and the Michigan consumer expectations index—had MAEs less than 6 percent.

Table 2.15 presents results for univariate forecasts of three-month growth, that is, indexes constructed with $c_{t+3|T} - c_{t|T}$ as the dependent variable in (22). The results are broadly similar to those in table 2.15. Each of the univariate indexes substantially misforecast three-month growth during the final episode, typically with forecast errors in the range of 5–7 percent (corresponding to one-quarter-ahead forecast errors of 3–4 percent in the units of annual GNP growth). The four indicators with the smallest MAEs at the six-month horizon have MAEs under 4 percent at the three-month horizon, as do the new orders index (PMNO) and consumer expectations.

In summary, these results suggest that, taken individually, only a handful of indicators would have been useful in predicting the 1990 downturn. During this contraction, consumer expectations, building permits, business expecta-

Table 2.15 Performance of Single-Variable Indexes: Three-Month-Ahead Forecast Horizon

	RMSE Computed over:				MAE Computed over:			
Series	62:1–88:9	88:10–90:8	88:10–90:4	90:5–90:8	62:1–88:9	88:10–90:8	88:10–90:4	90:5–90:8
	Coincident Indicators							
IP	4.53	3.06	1.96	4.93	3.28	2.29	1.41	4.67
GMYXP8	4.57	3.39	1.93	5.70	3.30	2.43	1.34	5.38
MT82	4.56	3.21	1.50	5.68	3.27	2.21	1.08	5.30
LPMHUADJ	4.55	3.37	1.84	5.74	3.29	2.40	1.33	5.30
	Leading Indicators in the XLI							
HSBP	4.15	2.53	1.56	4.13	3.07	1.87	1.19	3.72
MDU82S	4.48	3.25	2.01	5.31	3.20	2.28	1.32	4.90
EXNWT2FS	4.42	4.44	2.10	7.83	3.23	3.18	1.74	7.11
LHNAPSS	4.35	3.29	2.03	5.38	3.14	2.57	1.66	5.02
FYGT10FS	4.36	3.40	2.18	5.47	3.17	2.52	1.63	4.95
CP6_GM6F	3.66	4.35	1.51	8.00	2.80	2.98	1.24	7.73
G10_GLF	4.12	3.64	1.38	6.63	3.08	2.41	1.03	6.15
	Leading Indicators in the XLI2							
LPHRM	4.54	3.12	1.88	5.16	3.27	2.23	1.26	4.85
IPXMCA	4.55	3.08	1.86	5.08	3.28	2.23	1.29	4.77
LHEL	4.05	2.68	2.52	3.07	2.89	2.24	2.03	2.82
IVPAC	4.40	3.71	1.73	6.57	3.23	2.51	1.12	6.30
	Financial Indicators							
FSPCOMF	4.30	3.31	2.61	4.72	3.22	2.67	2.03	4.41
FM1D82	4.28	3.10	1.92	5.06	3.09	2.32	1.48	4.62
FM2D82	4.08	2.53	2.02	3.57	2.97	2.06	1.67	3.10
FMBASE	4.51	3.62	1.68	6.40	3.19	2.42	1.22	5.70
FCBCUCY	4.57	3.43	2.10	5.64	3.26	2.67	1.67	5.39
FYFFF	4.09	3.17	1.39	5.67	2.94	2.15	.98	5.30
BAA_G10F	4.25	3.19	2.13	5.05	3.04	2.47	1.69	4.60
YLD_DUMF	4.07	4.52	2.27	7.86	2.97	3.40	1.86	7.58
	Employment Indicators							
LUINC	4.46	3.18	1.83	5.34	3.21	2.28	1.29	4.99
LHU5	4.55	3.57	2.11	5.95	3.25	2.63	1.52	5.65
LHELX	4.29	2.95	1.84	4.81	3.10	2.22	1.40	4.44
	Consumption and Retail Sales							
IPCD	4.57	3.41	1.94	5.74	3.29	2.44	1.34	5.43
GMCD82	4.53	3.26	1.96	5.38	3.26	2.32	1.30	5.06
RTR82	4.53	3.15	1.82	5.27	3.27	2.22	1.25	4.86

Table 2.15 (continued)

	RMSE Computed over:				MAE Computed over:			
Series	62:1–88:9	88:10–90:8	88:10–90:4	90:5–90:8	62:1–88:9	88:10–90:8	88:10–90:4	90:5–90:8
	Inventories and Orders							
MPCON8	4.53	3.55	2.05	5.94	3.22	2.60	1.54	5.50
MOCM82	4.53	3.58	2.02	6.04	3.23	2.63	1.49	5.72
MDO82	4.50	3.64	2.33	5.87	3.17	2.59	1.62	5.24
IVMT82	4.55	3.81	2.37	6.21	3.25	2.78	1.70	5.72
IVM1D8	4.54	3.83	2.17	6.46	3.32	2.82	1.57	6.21
IVM2D8	4.57	3.53	1.93	6.02	3.34	2.56	1.40	5.71
IVM3D8	4.56	3.54	2.04	5.94	3.28	2.57	1.43	5.64
	Additional Indicators							
DLBLNPAP	4.51	2.80	1.71	4.60	3.26	2.17	1.33	4.44
PMI	4.42	3.04	1.84	5.01	3.22	2.30	1.40	4.75
PMNO	4.28	2.68	1.97	4.02	3.08	2.13	1.56	3.70
HHSNTN	4.27	2.32	2.01	3.00	3.10	1.90	1.72	2.41
HHST	4.34	2.75	2.21	3.86	3.12	2.21	1.74	3.48
PW561	4.56	3.94	2.04	6.82	3.29	2.77	1.60	5.95
PW561R	4.58	4.10	2.08	7.13	3.31	2.84	1.60	6.23
FTM333	4.58	2.88	1.61	4.89	3.29	2.05	1.15	4.50
FTM333R	4.57	3.04	1.67	5.18	3.29	2.19	1.20	4.88

Note: All results were computed using the data as revised through 1991:2. See the notes to table 2.12.

tions, help wanted advertising, and oil prices moved in advance of overall economic activity. With the exception of stock prices, indicators of financial market conditions—the slope of the Treasury yield curve, the public-private spreads, exchange rates, and interest rates—exhibited different patterns than they did during the recessions in the 1970s and 1980s. One interpretation of these observations is that, at least since 1969, recessions have been associated with contractionary monetary policy; this was captured by the interest rate indicators, accounting for their strong in-sample performance. The downturn of 1990, however, occurred in the face of monetary policy that, if not expansionary, was far less contractionary than it had been during the recessions of the 1970s and early 1980s. Instead, the contraction was associated with sharp drops in consumer expectations, business expectations, and uncertainty over a possible war in the Gulf.

2.4.5 Construction of Alternative Indexes

The results of the previous section suggest that the key problem in model specification and forecasting over this period was the ability to select those few leading indicators that forecast the 1990 downturn. This indicator selection problem is studied empirically here by constructing composite indexes that forecast six-month growth in the XCI from two shortened lists of the series in tables 2.14 and 2.15. The first list consists of the eleven leading indicators in the XLI and the XLI2, augmented by stock prices, the new orders index, consumer expectations, and oil prices. These four variables were intentionally chosen because of their good performance in the second episode. The second list eliminates from the first exchange rates and all interest rate indicators (EXNWT2FS, G10_G1F, FYGT10FS, CP6_GM6F), which are replaced by measures of sales, orders, new unemployment insurance claims, and manufacturing and trade inventories (RTR82, MPCON8, LUINC, IVMT82).[10]

For each list, all possible indexes of the form (21) and (22) were constructed, subject only to the restriction that the indexes included no more than seven leading indicators. The weights $\{\alpha_i\}$ for each index were estimated by OLS. For each list, this produced 16,383 indexes, which were then ranked by their Schwarz information criterion (evaluated over 1962:1–1988:9), where the number of parameters equaled the number of univariate indexes included in the composite trial indicator. All parameters were estimated over 1962:1–1988:9 (with earlier observations for initial lags) using the most recently available revised data.

The performance of the top fifteen indexes based on the financial list is summarized in panel A of table 2.16. (The individual indexes $\hat{y}_{it}$ were constructed as in table 2.14.) Not surprisingly, the in-sample RMSEs of these fifteen indexes are almost identical and slightly surpass that of the XLI. The out-of-sample RMSEs vary and are somewhat better than the XLI's, reflecting improvements made by the additional included variables—in particular, stock prices, the purchasing managers' index, and help wanted advertising. However, all but one of these top fifteen indexes still have out-of-sample RMSEs exceeding 6 percent, well above the RMSEs achieved by some of the individual indicators in table 2.14.

Indexes constructed using the fifteen non–interest rate and non–exchange rate indicators are examined in panel B of table 2.16. Although the in-sample RMSEs are substantially worse than in panel A (3.1 percent rather than 2.6 percent), the RMSEs in the 1990 episode are cut almost in half. All indexes have out-of-sample RMSEs near—in one case, less than—their in-sample RMSE. The variables that appear in most of the indexes are housing starts, weekly employee hours, help wanted advertising, stock prices, and, to a lesser extent, consumer sentiment. The price of oil appears in only one of these top fifteen indexes, and, significantly, this index exhibits out-of-sample performance typical of the indexes that exclude oil prices.

10. The results in tables 2.14 and 2.15 suggest that real money growth is another plausible indicator to be included. However, we find persuasive Friedman and Kuttner's (in press) evidence that the relation between real money growth and output has been unstable historically—particularly over the 1980s—and therefore excluded monetary aggregates as candidate indicators from these lists.

Table 2.16 **Performance of Alternative Composite Indexes: Six-Month Ahead Forecast Horizon**

Rank	Included Leading Indicators	RMSE Computed over: 62:1–88:9	88:10–90:8	88:10–90:4	90:5–90:8	MAE Computed over: 62:1–88:9	88:10–90:8	88:10–90:4	90:5–90:8
	XLI	2.76	3.42	1.39	7.62	2.20	2.15	1.08	7.24
	XLI2	3.80	2.61	1.57	5.24	2.83	1.93	1.35	4.67
		A. List Including XLI, XLI2, and Selected Alternative Indicators							
1	HSBP, CP6–GM6F, G10–G1F, LPHRM, PMNO, HHSNTN, PW561	2.57	3.04	1.49	6.52	2.07	1.95	.99	6.49
2	HSBP, CP6–GM6F, LPHRM, LHEL, PMNO, HHSNTN, PW561	2.58	2.78	1.39	5.94	2.04	1.86	1.01	5.89
3	HSBP, CP6–GM6F, LPHRM, FSPCOMF, PMNO, HHSNTN, PW561	2.58	3.00	1.52	6.37	2.04	2.02	1.12	6.34
4	HSBP, CP6–GM6F, G10–G1F, LPHRM, LHEL, PMNO, PW561	2.58	3.04	1.30	6.72	2.05	1.93	.94	6.64
5	HSBP, CP6–GM6F, G10–G1F, LPHRM, FSPCOMF, PMNO, PW561	2.59	3.25	1.38	7.20	2.06	2.01	.92	7.16
6	HSBP, CP6–GM6F, G10–G1F, LPHRM, LHEL, FSPCOMF, PMNO	2.59	3.13	1.14	7.09	2.05	1.95	.89	6.99
7	HSBP, FYGT10FS, CP6–GM6F, LPHRM, PMNO, HHSNTN, PW561	2.59	2.97	1.53	6.30	2.06	2.02	1.13	6.26
8	HSBP, CP6–GM6F, G10–G1F, LPHRM, LHEL, PMNO, HHSNTN	2.59	3.04	1.28	6.74	2.06	2.01	1.04	6.63

(continued)

Table 2.16 (continued)

		RMSE Computed over:				MAE Computed over:			
Rank	Included Leading Indicators	62:1–88:9	88:10–90:8	88:10–90:4	90:5–90:8	62:1–88:9	88:10–90:8	88:10–90:4	90:5–90:8
9	HSBP, CP6–GM6F, LPHRM, IVPAC, PMNO, HHSNTN, PW561	2.60	2.93	1.50	6.22	2.06	1.91	1.02	6.18
10	HSBP, MDU82S, CP6–GM6F, LPHRM, PMNO, HHSNTN, PW561	2.60	2.97	1.53	6.28	2.06	1.96	1.06	6.25
11	HSBP, EXNWT2FS, CP6–GM6F, LPHRM, PMNO, HHSNTN, PW561	2.60	2.98	1.52	6.34	2.06	1.96	1.05	6.30
12	HSBP, CP6–GM6F, LPHRM, IPXMCA, PMNO, HHSNTN, PW561	2.60	2.97	1.52	6.29	2.06	1.95	1.05	6.26
13	HSBP, LHNAPSS, CP6–GM6F, LPHRM, PMNO, HHSNTN, PW561	2.60	2.97	1.52	6.29	2.06	1.95	1.05	6.26
14	HSBP, CP6–GM6F, G10–G1F, LPHRM, FSPCOMF, PMNO, HHSNTN	2.60	3.26	1.36	7.24	2.07	2.06	.99	7.18
15	HSBP, CP6–GM6F, LPHRM, LHEL, FSPCOMF, PMNO, HHSNTN	2.60	2.99	1.31	6.57	2.05	2.00	1.05	6.49
		B. List Excluding Exchange Rates and All Interest Rate Indicators							
1	HSBP, LPHRM, LHEL, IVPAC, FSPCOMF, PMNO, HHSNTN	3.14	2.43	1.75	4.41	2.38	1.98	1.50	4.27
2	HSBP, LPHRM, LHEL, IVPAC, RTR82, PMNO, HHSNTN	3.14	2.45	1.72	4.52	2.40	1.99	1.49	4.37
3	HSBP, LPHRM, LEHL, IVPAC, MPCON8, PMNO, HHSNTN	3.15	2.17	1.52	4.02	2.39	1.74	1.30	3.86

4	HSBP, LPHRM, LHEL, IVPAC, IVMT82, PMNO, HHSNTN	3.15	2.11	1.60	3.69	2.40	1.68	1.29	3.53
5	HSBP, LPHRM, LHEL, FSPCOMF, MPCON8, PMNO, HHSNTN	3.17	2.00	1.62	3.25	2.41	1.63	1.33	3.09
6	HSBP, LPHRM, LHEL, IVPAC, LUINC, PMNO, HHSNTN	3.17	2.33	1.66	4.26	2.40	1.85	1.38	4.07
7	HSBP, MDU82S, LPHRM, LHEL, IVPAC, PMNO, HHSNTN	3.18	2.36	1.59	4.49	2.41	1.83	1.32	4.28
8	HSBP, LPHRM, LHEL, IVPAC, PMNO, HHSNTN, PW561	3.18	2.26	1.57	4.22	2.40	1.79	1.32	4.04
9	HSBP, LHNAPSS, LPHRM, LHEL, IVPAC, PMNO, HHSNTN	3.18	2.32	1.59	4.34	2.41	1.83	1.34	4.15
10	HSBP, LPHRM, IPXMCA, LHEL, IVPAC, PMNO, HHSNTN	3.18	2.30	1.58	4.31	2.41	1.82	1.34	4.12
11	HSBP, LPHRM, LHEL, FSPCOMF, RTR82, PMNO, HHSNTN	3.19	2.16	1.70	3.61	2.42	1.73	1.38	3.44
12	HSBP, LPHRM, LHEL, FSPCOMF, IVMT82, PMNO, HHSNTN	3.19	2.02	1.70	3.13	2.42	1.70	1.43	2.97
13	HSBP, LPHRM, LHEL, IVPAC, FSPCOMF, MPCON8, HHSNTN	3.20	2.60	2.02	4.42	2.46	2.19	1.77	4.18
14	HSBP, LPHRM, IPXMCA, LHEL, FSPCOMF, PMNO, HHSNTN	3.20	2.12	1.68	3.53	2.41	1.72	1.38	3.35
15	HSBP, LPHRM, LHEL, IVPAC, FSPCOMF, IVMT82, HHSNTN	3.20	2.54	2.05	4.14	2.46	2.12	1.75	3.90

Note: The indexes were selected form all possible indexes of the form (21) and (22) that include at most 7 indicators, selected from lists of 15 indicators. The lists are as follows: panel A: HSBP, MDU82S, EXNWT2FS, LHNAPSS, FYGT10FS, CP6_GM6F, G10_G1F, LPHRM, IPXMCA, LHEL, IVPAC, FSPCOMF, PMNO, HHSNTN, PW561; panel B: HSBP, MDU82S, LHNAPSS, LPHRM, IPXMCA, LHEL, IVPAC, FSPCOMF, LUINC, RTR82, MPCON8, IVMT82, PMNO, HHSNTN, PW561. All results were computed using the data as revised through 1991:2. See the notes to table 2.12.

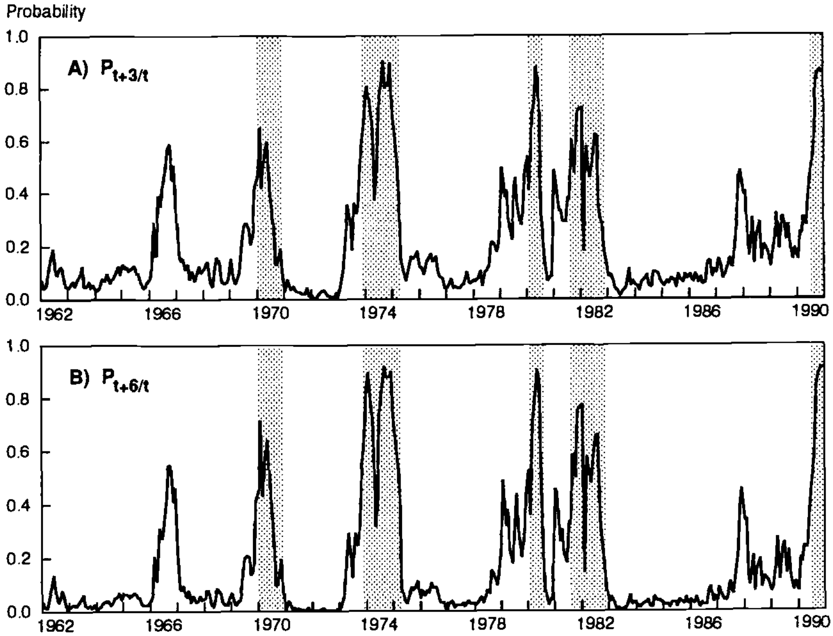

Fig. 2.6 Recession probabilities based on alternative nonfinancial indicators
Note: The dates on the horizontal axes denote t, the date through which the data are available for computing $P_{t+k|t}$. The recession probabilities were computed using the model of sec. 2.1, with the leading indicators: HSBP, LPHRM, LHEL, IVPAC, FSPCOMF, PMNO, HHSNTN. The series are based on data revised through 1991:2. The shaded areas represent NBER data recessions.

As a final exercise, the XLI model of section 2.1 was reestimated, replacing the leading indicators in panel B of table 2.1 with the seven indicators in the top-ranking index from the nonfinancial list (the first row in panel B, table 2.16). (In the notation of sec. 2.1, θ was reestimated, but μ was not.) The resulting recession probabilities $P_{t+k|t}$ are plotted for selected horizons in figure 2.6. The results are striking: the six-month-ahead recession probabilities $P_{t+6|t}$ computed with these indicators over July–October are, respectively, 84, 89, 92, and 91 percent. Evidently, had the short list of fifteen indicators examined in panel B been used over the summer of 1990 to produce recession forecasts, they would have predicted a recession starting in the fall of 1990. At the same time, it must be emphasized that, had they been used during the previous historical recessions, this set of indicators would have done substantially worse than those in the XRI: the index would have provided scant advance indication of recessions in 1969 and 1974 and only ambiguous signals in 1979 and 1981.

The composite indexes in this section put nonzero weight on only a few of the indicators and no weight on most. A natural alternative would be to construct a broadly based index that places weight on many or all of the indicators in table 2.13. One approach, analogous to the method used by the DOC to construct its index, would be to put equal weight on all the included indicators; another, advocated by Sims (1989), would be to impose strong prior restrictions on the weights so that, even though many coefficients would be estimated, these coefficients effectively would be a function of a much smaller number of coefficients that could be estimated more precisely.

Some initial calculations suggest that such broadly based indexes would also have performed poorly in the 1990 episode. For example, a six-month-ahead index of the form (21) and (22), constructed using the forty-one individual leading indicators in table 2.14 with equal weights on each indicator ($\alpha_i = 1/41$), results in an in-sample RMSE of 3.9 percent and an out-of-sample RMSE of 7.4 percent; neither represents a noticeable improvement over a constant forecast. When the weights $\{\sigma_i\}$ are estimated by OLS so that only forty-one parameters are estimated (in addition to the estimated lag coefficients in [22]), the in-sample RMSE drops to a low 2.2 percent, but the out-of-sample RMSE remains very large, 7.0 percent. These broadly based indicators do not exploit correlations across the individual leading indicators, so there might be room for improvement relative to these two crude indexes. Still, the poor performance of the preponderance of individual indicators in this episode, documented in tables 2.14 and 2.15, suggests that other more sophisticated broadly based indexes would also have performed poorly in this episode.

2.5 Discussion and Conclusions

The foregoing analysis has focused on the empirical performance of one particular forecasting tool—the XLI/XRI model—but also has implications that apply more generally to forecasting exercises based on a broad set of leading indicators. Focusing initially on the XRI and the XLI, we draw six conclusions. First, there is only weak evidence for nonlinearities in the data not captured in the linear model sketched in section 2.1. The main exception is the usefulness of the estimated duration of a downturn in predicting when that downturn will end. This duration dependence, however, appears to be restricted to downturns and is significant only at very short forecasting horizons.

Second, forecasts through September 1990 based on the XLI/XRI model performed quite well. For example, the absolute three-month-ahead forecast error in June 1990 was only 0.4 percent at an annual rate. However, the model failed to forecast the precipitous declines that started in October.

Third, there is no compelling evidence that the recession definition or the algorithm for computing recession probabilities was misspecified. Rather, the failure of the model in the fall of 1990 can be attributed to large forecast errors in the conditional means.

Fourth, the forecast errors in the conditional means do not appear to be the result of overparameterization or the imprecise estimation of too many parameters, given the list of included indicators. A key piece of evidence for this conclusion is that simple composite indexes with the same variables as the XLI but fewer parameters and a different, simpler structure exhibit performance comparable to that of the XLI, both in and out of sample. Moreover, these simple alternative indexes that use the same set of indicators make forecast errors as large as or larger than those made by the XLI model during the fall of 1990.

Fifth, the short-horizon recession probabilities produced by the model performed relatively well during this episode. In October, the first month in which there were large declines in the coincident indicators, $P_{t|t}$ was 28 percent; by November, it was 80 percent. Thus, the index can claim the modest success of "forecasting" that the economy was already in a recession, once the downturn had begun in earnest.

Sixth, the key source of difficulty was with the choice of indicators included in the model. The financial variables in the XRI and the XLI behaved quite differently over the summer of 1990 than they had during the preceding recessions in 1973, 1980, and 1981. Prior to those earlier recessions, the Treasury-bill yield curve was sharply inverted, while in June–September 1990 it sloped upward. The corporate paper–Treasury-bill spread in June 1990 was one-third its average value in the month before the previous three recessions. Although these indicators failed to predict the 1990 recession, a few (but not many) alternative indicators would have provided advanced warning had they been incorporated into the XRI. The strongest evidence of this is the performance of the alternative seven-indicator index, constructed from the list of fifteen indicators that excluded the poorly performing interest rate and exchange rate indicators and included selected indicators that, as it turned out, performed well, such as consumer expectations. Had this set of indicators been used as the basis of the XRI, the index would have registered much larger recession probabilities and would have reduced the out-of-sample RMSE by almost half.

These results also suggest some more general conclusions and areas for future research. While the results of sections 2.4.4 and 2.4.5 indicate that the downturn could have been forecast—in the sense that there were composite indexes that, had they been constructed, would have performed well during the summer and fall of 1990—the decisions on which variables to include and which to omit were made with the benefit of hindsight. The challenging, unsolved—and hardly new—problem that this underscores is developing an appropriate methodology for the identification and selection of leading indicators.

Appendix
Variable Definitions

All data were obtained from Citibase, with the exception of those denoted "(AC)," which were calculated by the authors. A single asterisk (*) indicates that log first-differences of the variable were used, a double asterisk (**) that first-differences of the variable were used.

Coincident Indicators

IP*

Industrial production: total index (1977 = 100; seasonally adjusted).

GMYXP8*

Personal income: total less transfer payments, 1982$ ($billion, seasonally adjusted at an annual rate).

MT82*

Manufacturing and trade sales: total, 1982$ ($million, seasonally adjusted).

LPMHUADJ (AC)*

Citibase series LPMHU (employee-hours in nonagricultural establishments [billion hours, seasonally adjusted at an annual rate]), adjusted for short sampling weeks in 1970:9, 1974:4, 1979:4, 1981:9, and 1982:1. If the sampling week was short in month t, the adjusted series was computed as ½ (LPMHU_{t+1} + LPMHU_{t-1}).

Leading Indicators in the XLI

HSBP

Housing authorized: index of new private housing units (1967 = 100; seasonally adjusted).

MDU82S (AC)

Manufacturers' unfilled orders: durable goods industries, total ($million, seasonally adjusted) (MDU), deflated by the producer price index: durable manufacturing goods (not seasonally adjusted) (PWDMD): log first-difference, smoothed. PWDMD was seasonally adjusted prior to deflating by removing average monthly growth rates.

EXNWT2FS (AC)

EXNWT2 is the nominal weighted exchange rate between the United States and France, Italy, Japan, the United Kingdom, and West Germany, constructed using shares of total real imports as weights. EXNWT2FS is the log first-difference, smoothed, led by one month.

LHNAPSS (AC)

LHNAPS is persons at work: part time for economic reasons—slack work, nonagricultural industries (thousands, seasonally adjusted). LHNAPSS is the log first-difference, smoothed.

FYGT10FS (AC)

FYGT10 is the interest rate: U.S. Treasury constant maturities, 10 year (% per annum, not seasonally adjusted). FYGT10FS is the first-difference, smoothed, led by one month.

CP6_GM6F (AC)

FYCP − FYGM6, led by one month, where FYCP is the interest rate: commercial paper, 6 month (% per annum, not seasonally adjusted), and FYGM6 is the interest rate: U.S. Treasury bills, secondary market, 6 month (% per annum, not seasonally adjusted).

G10_G1F (AC)

FYGT10 − FYGT1, led by one month, where FYGT10 is the interest rate: U.S. Treasury constant maturities, 10 year (% per annum, not seasonally adjusted), and FYGT1 is the interest rate: U.S. Treasury constant maturities, 1 year (% per annum, not seasonally adjusted).

Leading Indicators in the XLI 2

LPHRM

Average weekly hours of production workers: manufacturing (seasonally adjusted).

IPXMCA**

Capacity utilization rate: manufacturing total (% of capacity, seasonally adjusted).

LHEL*

Index of help wanted advertising in newspapers (1967 = 100; seasonally adjusted).

IVPAC

Vendor performance: % of companies reporting slower deliveries (% not seasonally adjusted).

Financial Variables

FSPCOMF (AC)*

S&P'S common stock price index: composite (1941–43 = 10), led by one month.

FM1D82*

Money stock: M1 in 1982$ ($billion, seasonally adjusted).

FM2D82*

Money stock: M2 in 1982$ ($billion, seasonally adjusted).

FMBASE*

Monetary base, adjusted for reserve requirement changes (Federal Reserve Bank of St. Louis) ($billion, seasonally adjusted).

CCI30M*

Consumer installment loans: delinquency rate, 30 days and over (% seasonally adjusted).

FCBCUCY (AC)

Change in business and consumer credit outstanding (%, seasonally adjusted at an annual rate) (FCBCUC) minus the annual percentage growth in total nominal person income (GMPY).

FYFFF

Interest rate: Federal funds (effective) (% per annum, not seasonally adjusted), led by one month.

BAA_G10F (AC)

FYBAAC − FYGT10, led by one month, where FYBAAC is the bond yield: Moody's Baa Corporate (% per annum), and FYGT10 is the interest rate: U.S. Treasury constant maturities, 10 year (% per annum, not seasonally adjusted).

YLD_DUMF (AC)

Inverted yield curve dummy, led by one month. Dummy variable that takes on a value of 1 when G10_G1 is negative.

Employment Variables

LUINC*

Average weekly initial claims, state unemployment insurance, excluding Puerto Rico (thousands, seasonally adjusted).

LHU5*

Unempoyment by duration: persons unemployed less than 5 weeks (thousands, seasonally adjusted).

LHELX

Employment: ratio; help wanted ads: no. unemployed current labor force.

Sales and Consumption Variables

IPCD*

Industrial production: durable consumer goods (1977 = 100; seasonally adjusted).

GMCD82*

Personal consumption expenditures: durable goods, 1982$.

RTR82*

Retail sales: total, 1982$ ($million, seasonally adjusted).

Inventories and Orders

MPCON8*

Contracts and orders for plant and equipment in 1982$ ($billion, seasonally adjusted).

MOCM82*

Manufacturing new orders: consumer goods and material, 1982$ ($billion, seasonally adjusted).

MDO82*

Manufacturing new orders: durable goods industries, 1982$ ($billion, seasonally adjusted).

IVMT82*

Manufacturing and trade inventories: total, 1982$ ($billion, seasonally adjusted).

IVM1D8 (AC)

Log first-difference of IVM1: real manufacturing inventories, materials and supplies: all manufacturing industries (materials and supplies inventories), deflated by the total inventories price deflator, IVMT/IVMT82, where IVMT is total nominal manufacturing inventories. Growth rate in 1982:1 is average of growth rates for 1981:12 and 1982:2 to adjust for accounting change in 1982:1.

IVM2D8 (AC)

Log first-difference of IVM2: real manufacturing and trade inventories: work in process, all manufacturing industries (seasonally adjusted), (work in progress inventories), deflated by the total inventories price deflator, IVMT/IVMT82, where IVMT is total nominal manufacturing inventories. Growth rate in 1982:1 is average of growth rates for 1981:12 and 1982:2 to adjust for accounting change in 1982:1.

IVM3D8 (AC)

Log first-difference of IVM3: real manufacturing inventories: finished goods, all manufacturing industries (finished goods inventories), deflated by the total inventories price deflator, IVMT/IVMT82, where IVMT is total nominal manufacturing inventories. Growth rate in 1982:1 is average of growth rates of 1981:12 and 1982:2 to adjust for accounting change in 1982:1.

Additional Indicators

DLBLNPAP (AC)

Log first-difference of the ratio of the volume of bank loans to the volume of commercial paper, where bank loans are commercial bank loans to the nonfarm corporate business sector and the nonfarm corporate sector, excluding mortgages and bankers' acceptances. The original series is quarterly and was distributed to a monthly basis as follows: the growth rate from QI to QII (say) was used as the data for June, July, and August (because growth from QII to QIII includes lending through September). *Source:* Federal Reserve Board, quarterly flow-of-funds data bank (kindly provided by A. Kashyap, J. Stein, and D. Wilcox).

PMI

Purchasing managers' index (seasonally adjusted).

PMNO

National Association of Purchasing Managers new orders index (%).

HHSNTN

University of Michigan index of consumer expectations.

HHST (AC)

HHSENT interpolated with HHSNTR: 1953–1977:4, HHSENT, University of Michigan index of consumer sentiment, (1966:I = 100; not seasonally adjusted); 1978:1–1990:12, HHSNTR, University of Michigan index of consumer sentiment (February 1966 = 100; not seasonally adjusted).

PW561*

Producer price index: crude petroleum (1982 = 100; not seasonally adjusted).

PW561R (AC)*

PW561/PW, where PW = producer price index: all commodities (1982 = 100; not seasonally adjusted).

FTM333

U.S. merchandise imports: petroleum and petroleum products ($million, seasonally adjusted).

FTM333R

FTM333/PW, where PW = producer price index: all commodities (1982 = 100; not seasonally adjusted).

Aggregate Indexes

XLI (AC)

NBER experimental index of leading indicators.

MTREC (AC)

Expected length of current recession (construction is described in the text).

MTEXP (AC)

Expected length of current expansion (construction is described in the text).

MTTOT (AC)

$(P_{t|t}) \times$ (MTREC) + $(1 - P_{t|t}) \times$ (MTEXP) (construction is described in the text).

DLEAD*

U.S. Department of Commerce composite index of 11 leading indicators (1982 = 100; seasonally adjusted).

DL3D (AC)

Dummy variable taking on the value of 1 at time t if 3 consecutive downturns of DLEAD have occurred. That is, DL3D(t) = 1 if ΔDLEAD(t) < 0, ΔDLEAD($t - 1$) < 0, and ΔDLEAD($t - 2$) < 0, and DL3D (t) = 0 otherwise.

DL3U (AC)

Dummy variable taking on the value of 1 at time t if 3 consecutive upturns of DLEAD have occurred. That is, DL3U(t) = 1 if ΔDLEAD(t) > 0, ΔDLEAD($t - 1$) > 0, and ΔDLEAD($t - 2$) > 0, and DL3U(t) = 0 otherwise.

IP3D (AC)

Dummy variable taking on the value of 1 at time t if 3 consecutive downturns of IP have occurred. That is, IP3D(t) = 1 if ΔIP(t) < 0, ΔIP($t - 1$) $<$ 0, and ΔIP($t - 2$) < 0, and IP3D(t) = 0 otherwise.

IP3U (AC)

Dummy variable taking on the value of 1 at time t if 3 consecutive upturns of IP have occurred. That is, $\text{IP3U}(t) = 1$ if $\Delta\text{IP}(t) < 0$, $\Delta\text{IP}(t-1) < 0$, and $\Delta\text{IP}(t-2) < 0$, and $\text{IP3D}(t) = 0$ otherwise.

References

Auerbach, A. J. 1982. The index of leading indicators: "Measurement without theory," thirty-five years later. *Review of Economics and Statistics* 64 (4): 589–95.

Burns, A. F., and W. C. Mitchell. 1946. *Measuring business cycles.* New York: NBER.

Friedman, B. M., and K. N. Kuttner. In press. Another look at the evidence on money-income causality. *Journal of Econometrics.*

Geweke, J. 1977. The dynamic factor analysis of economic time series. In *Latent variables in socio-economic models,* ed. D. J. Aigner and A. S. Goldberger. Amsterdam: North-Holland.

Hamilton, J. D. 1989. A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica* 57:357–84.

Harvey, A. C., R. Blake, M. Desai, and C. McKenzie. 1981. Data revisions. In *Applied time series analysis of economic data,* ed. A. Zellner. Economic Research Report no. ER-5. Washington, D.C.: U.S. Department of Commerce.

Hymans, S. 1973. On the use of leading indicators to predict cyclical turning points. *Brookings Papers on Economic Activity* (2), 339–84.

Kling, J. L. 1987. Predicting the turning points of business and economic time series. *Journal of Business* 60, (2): 201–38.

Mankiw, N. G., D. Runkle, and M. D. Shapiro. 1984. Are preliminary announcements of the money supply rational forecasts? *Journal of Monetary economics* 14:15–27.

Mankiw, N. G., and M. D. Shapiro. 1986. News of noise analysis of GNP revisions. *Survey of Current Business* 66 (May): 20–25.

Mitchell, W. C., and A. F. Burns. [1938] 1961. Statistical indicators of cyclical revivals. Reprinted in *Business cycle indicators,* ed. G. H. Moore. Princeton, N.J.: Princeton University Press.

Neftci, S. N. 1982. Optimal prediction of cyclical downturns. *Journal of Economic Dynamics and Control* 4:225–41.

Sargent, T. J., and C. A. Sims. 1977. Business cycle modeling without pretending to have too much *a priori* economic theory. In *New methods in business cycle research,* ed. C. Sims. Minneapolis: Federal Reserve Bank of Minneapolis.

Sims, C. A. 1989. Comment on "New indexes of leading and coincident economic indicators." *NBER Macroeconomics Annual,* 394–97.

Singleton, K. 1980. A latent time series model of the cyclical behavior of interest rates. *International Economic Review* 21 (3): 559–75.

Stekler, H. O., and M. Schepsman. 1973. Forecasting with an index of leading series. *Journal of the American Statistical Association* 68 (342): 291–96.

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

———. 1989. New indexes of leading and coincident economic indicators. *NBER Macroeconomics Annual,* 351–94.

———. 1991. A probability model of the coincident economic indicators. In *Leading economic indicators: New approaches and forecasting records,* ed. K. Lahiri and G. H. Moore. New York: Cambridge University Press.

Vaccara, B. N., and V. Zarnowitz. 1978. Forecasting with the index of leading indicators. Working Paper no. 244. Cambridge, Mass.: NBER.
Wecker, W. E. 1979. Predicting the turning points of a time series. *Journal of Business* 52 (1): 35–50.
Zarnowitz, V., and C. Boschan. 1975a. Cyclical indicators: An evaluation and new leading indexes. *Business Conditions Digest* (May), v–xxiv.
———. 1975b. New composite indexes of coincident and lagging indicators. *Business Conditions Digest* (November), v–xxii.
Zarnowitz, V., and G. H. Moore. 1982. Sequential signals of recession and recovery. *Journal of Business* 55 (1): 57–85.
Zellner, A., C. Hong, and G. M. Gulati. 1987. Turning points in economic time series, loss structures and Bayesian forecasting. University of Chicago, Graduate School of Business. Typescript.

Comment Kenneth F. Wallis

Notable among recent work on cyclical indicators is the attempt by James H. Stock and Mark W. Watson (henceforth S–W) to recast the traditional approach in a suitable form for the application of the techniques of modern time-series econometrics. Two previous papers (Stock and Watson 1989, 1991) describe new indexes of coincident and leading indicators, which are based on the dynamic single-index model of Sargent and Sims (1977), and a new recession index, which is an estimate of the probability that the economy will be in recession in six months' time, where a recession is defined as a particular pattern of movements in the unobserved single index or "state of the economy." The publication of probability forecasts in economics is itself a considerable innovation. In the present paper, S–W provide a detailed account of the construction of the recession index and of its forecasting performance both in and out of sample; in the light of its forecast failure in late 1990, they then return to the question of the selection of variables used as indicators in the model.

The evaluation of economic forecasts has a large literature, to which important early contributions came from an NBER project directed by Victor Zarnowitz in the late 1960s. This literature scarcely considers probability forecasts, however, since these have scarcely featured in economics. Rather, the literature on the theory and practice of probability forecasting is largely to be found in the meteorological journals, stimulated by the inclusion of a statement of the probability of precipitation in U.S. weather forecasts over a considerable period of years (for a review, see Dawid 1986). A simple summary indicator of forecast performance is a *reliability diagram* or *calibration curve*,

Kenneth F. Wallis is professor of econometrics at the University of Warwick and director of the ESRC Macroeconomic Modelling Bureau.

in which the observed relative frequency of the event is calculated over subsets of occasions for which the forecast probability was at, or close to, preassigned values and then plotted against those forecast probability values: in the absence of sampling fluctuations, a diagonal line indicates perfect reliability or that the forecaster was "well calibrated" (Dawid 1982). Of course, large samples are always a help, and daily rainfall forecasts are thirty times more frequent than monthly economic forecasts; nevertheless, we can make a start by plotting $\bar{R}$ against $\bar{P}$ within the columns of panel B of S–W's table 2.3. This dramatically illustrates the lack of complete-sample reliability of the recession probability forecasts, which were in general too high. While the criterion of perfect reliability or complete calibration is not by itself a sufficient condition for forecasts to be good—it would be achieved in the present case by a forecaster whose one-month-ahead recession probability forecast was always equal to 0.146—it has often been taken to be a minimum desirable property.

Another useful diagrammatic presentation of probability forecasts is found in an exception to the general comment above about the lack of an economics literature, namely, the contribution to the early NBER project by Fels and Hinshaw (1968). For each turning point during 1948–61, they are concerned with how well it was first forecast and then, ex post, recognized, so for a number of different forecasters they plot $P_{\tau|t}$ for fixed τ, the turning-point date, against t, the date of the forecast, in S–W's notation. (Mention of Fels and Hinshaw is not intended to detract from the originality of S–W's probability forecasts: Fels and Hinshaw's analysis was based on their own subjective assessment of the odds on a turning point implicit in the forecasters' language.) In the present case, this plot amounts to reading the entries in table 2.6 along a diagonal moving upward to the right, setting τ at October 1990 if that is eventually declared to be the turning point. Corresponding plots for previous turning points can be constructed from data shown in the various panels of figure 2.1: although these are within sample, the resulting comparison dramatically illustrates the relative forecast failure of the recession index in the most recent episode.

Attention then turns to the choice of indicator variables to include in the model and the appropriateness of the model itself. The first of these has already been the subject of comment by Sims (1989) and Zarnowitz and Braun (1989), who are now to some extent entitled to say we told you so. Sims questioned the heavy dependence of the new leading index on variables, three out of seven, that are functions of interest rates, while, in comparing the S–W and Department of Commerce (DOC) indexes, Zarnowitz and Braun questioned the inclusion of the nominal exchange rate and the exclusion of the DOC index's vendor performance variable. This criticism is by and large validated by S–W's evidence from the recent past: whereas the financial variables did not perform well, forecasts of the recession could have been improved by a selection of the more traditional indicators, together with a consumer sentiment variable that was added to the DOC set in January 1989. Why, then, were these variables not selected for inclusion in the original model?

The dynamic single-index model is a linear time-invariant model, and the selection and weighting of leading indicator variables is likewise based on linear regression methods. Whereas Sims (1989) expresses disappointment at the use of a model without time-varying coefficients, my reservation concerns its linearity. These are not unrelated, of course, since a linear approximation to a nonlinear model in general has time-varying or, more precisely, state-dependent coefficients. In a forecasting context, however, some model of the time variation is necessary, which requires either a return to the nonlinear model or, if this is unknown, the use of one of the statistical models of time-varying coefficients, which typically rest on underlying constant parameters. In the context of the statistical modeling of the business cycle, features such as its asymmetry are well established, as is the fact that these cannot be adequately accommodated by linear constant-parameter models. See, for example, Hamilton (1989), Pfann (1991), and earlier references given by these authors, whose own work illustrates the distinction drawn by S–W (Stock and Watson 1989, 356–57) between the "intrinsic" and the "extrinsic" views of cycles. In the former view, expansions and recessions are regarded as periods of distinctly different economic behavior, defined by intrinsic shifts in the data-generating process, whereas, in the latter view, expansions and recessions are extrinsic patterns that result from the adaptation of a stable structure to random shocks. Thus, Hamilton (1989) uses a Markov switching regression to characterize changes in the parameters of a linear autoregressive process, whereas Pfann (1991) uses a nonlinear autoregression with constant coefficients to capture asymmetries. Here, as in other areas of economics, competing views need to be tested against one another and, no doubt, against other models yet to be developed. Extension is also required from the univariate to the multivariate setting in which S–W are located, and rightly so, since the business cycle is about comovements in a broad range of macroeconomic aggregates. But the evidence from these different models suggests that S–W should not underestimate the evidence of nonlinearity found in their own rather limited testing and indeed supports their call for further research. If the regime-shift view is upheld, then, at a simple level, the selection of indicators for predicting recessions might be based on weighted rather than ordinary least squares.

Judgment plays an essential part in the selection of indicators, the dating of turning points, and various other aspects of traditional business-cycle analysis, and its elimination is the ultimate target of the modern model-based methods. Judgment, of course, accommodates nonlinearity of unspecified form, but it is not just seat of the pants, Zarnowitz and Braun (1989) noting the reliance of the traditional approach on business-cycle theory and selection criteria that are formally stated, albeit informally weighted; formalization of the former surely requires elements of the structural modeling approach. On the other hand, Sims (1989) notes that S–W have only partially formalized the selection process and that therefore the forecast uncertainty that results from uncertainty about which variables belong in the model is not completely known: "The criteria for [respecifying or adjusting the model] should be more explicit, if we are to have much improvement over the current judgmental DOC procedures" (p. 397). The seasonal adjustment problem provides an interesting parallel. Here, too, the traditional methods were developed in the absence of a formal probability model and formal statistical criteria; subsequently, model-based methods were developed as an alternative and as an aid to understanding the behavior and characteristics of the traditional methods. They have not replaced the traditional methods in the official statistical agencies, however, one objection being the need for skilled judgment in the model-based methods (specifically in the choice of the ARIMA representations used at various points). This might seem a somewhat contrary objection since more generally the traditionalists feel that the modern methods overlook much that is of value in their own use of judgment. Perhaps the same tension will persist in business-cycle analysis, with the result that a probability model that provides a complete rationalization of the traditional methods cannot be attained. Many avenues remain to be explored, however, and, in the meantime, S–W deserve congratulations for having clearly advanced the debate, in particular by establishing a more scientific foundation.

References

Dawid, A. P. 1982. The well-calibrated Bayesian [with discussion]. *Journal of the American Statistical Association* 77:605–13.

———. 1986. Probability forecasting. In *Encyclopedia of statistical sciences,* ed. S. Kotz and N. L. Johnson, vol. 7. New York: Wiley.

Fels, R., and C. E. Hinshaw. 1968. *Forecasting and recognizing business cycle turning points.* Studies in Business Cycles, no. 17. New York: Columbia University Press (for the NBER).

Hamilton, J. D. 1989. A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica* 57:357–84.

Pfann, G. A. 1991. Employment and business cycle asymmetries: A data based study. Discussion Paper no. 39. Minneapolis: Federal Reserve Bank of Minneapolis, Institute for Empirical Macroeconomics.

Sargent, T. J., and C. A. Sims. 1977. Business cycle modeling without pretending to have too much *a priori* economic theory. In *New methods in business cycle research,* ed. C. A. Sims. Minneapolis: Federal Reserve Bank of Minneapolis.

Sims, C. A. 1989. Comment on Stock and Watson (1989). *NBER Macroeconomics Annual,* 394–97.

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

———. 1991. A probability model of the coincident economic indicators. In *Leading economic indicators: New approaches and forecasting records,* ed. K. Lahiri and G. H. Moore. New York: Cambridge University Press.

Zarnowitz, V., and P. Braun. 1989. Comment on Stock and Watson (1989). *NBER Macroeconomics Annual,* 397–408.