Chapter 16

LEADING INDICATORS

MASSIMILIANO MARCELLINO*

IEP-Bocconi University, IGIER and CEPR
e-mail: *massimiliano.marcellino@uni-bocconi.it*

Contents

* I am grateful to two anonymous Referees, to participants at the Rady Conference on Economic Forecasting (UCSD) and to seminar participants at the Dutch Central Bank, University of Barcelona and Università Tor Vergata for useful comments on a previous draft; to Maximo Camacho, Jim Hamilton, Chang-Jin Kim, Charles Nelson, and Gabriel Perez-Quiros for sharing their code; to Ataman Ozyildirim at the Conference Board and Anirvan Banerji at ECRI for data and information; to Andrea Carriero and Alice Ghezzi for excellent research assistance; and to Nicola Scalzo for editorial assistance.

Handbook of Economic Forecasting, Volume 1
Edited by Graham Elliott, Clive W.J. Granger and Allan Timmermann

DOI: 10.1016/S1574-0706(05)01016-5

Abstract

In this chapter we provide a guide for the construction, use and evaluation of leading indicators, and an assessment of the most relevant recent developments in this field of economic forecasting. To begin with, we analyze the problem of indicator selection, choice of filtering methods, business cycle dating procedures to transform a continuous variable into a binary expansion/recession indicator, and methods for the construction of composite indexes. Next, we examine models and methods to transform the leading indicators into forecasts of the target variable. Finally, we consider the evaluation of the resulting leading indicator based forecasts, and review the recent literature on the forecasting performance of leading indicators.

Keywords

business cycles, leading indicators, coincident indicators, turning points, forecasting

JEL classification: E32, E37, C53

1. Introduction

Since the pioneering work of Mitchell and Burns (1938) and Burns and Mitchell (1946), leading indicators have attracted considerable attention, in particular by politicians and business people, who consider them as a useful tool for predicting future economic conditions. Economists and econometricians have developed more mixed feelings towards the leading indicators, starting with Koopmans' (1947) critique of the work of Burns and Mitchell, considered as an exercise in "measurement without theory". The resulting debate has stimulated the production of a vast literature that deals with the different aspects of the leading indicators, ranging from the choice and evaluation of the best indicators, possibly combined in composite indexes, to the development of more and more sophisticated methods to relate them to the target variable.

In this chapter we wish to provide a guide for the construction, use and evaluation of leading indicators and, more important, an assessment of the most relevant recent developments in this field of economic forecasting.

We start in Section 2 with a discussion of the choice of the target variable for the leading indicators, which can be a single variable, such as GDP or industrial production, or a composite coincident index, and the focus can be in anticipating either future values of the target or its turning points. We then evaluate the basic requirements for an economic variable to be a useful leading indicator, which can be summarized as:

(i) consistent timing (i.e., to systematically anticipate peaks and troughs in the target variable, possibly with a rather constant lead time);
(ii) conformity to the general business cycle (i.e., have good forecasting properties not only at peaks and troughs);
(iii) economic significance (i.e., being supported by economic theory either as possible causes of business cycles or, perhaps more importantly, as quickly reacting to negative or positive shocks);
(iv) statistical reliability of data collection (i.e., provide an accurate measure of the quantity of interest);
(v) prompt availability without major later revisions (i.e., being timely and regularly available for an early evaluation of the expected economic conditions, without requiring subsequent modifications of the initial statements);
(vi) smooth month to month changes (i.e., being free of major high frequency movements).

Once the choice of the target measure of aggregate activity and of the leading indicators is made, two issues emerge: first, the selection of the proper variable transformation, if any, and, second, the adoption of a dating rule that identifies the peaks and troughs in the series, and the associated expansionary and recessionary periods and their durations. The choice of the variable transformation is related to the two broad definitions of the cycle recognized in the literature, the so-called classical cycle and the growth or deviation cycle. In the case of the deviation cycle, the focus is on the deviations of the target variable from an appropriately defined trend rate of growth, while the classical cycle relies on the levels of the target variable. There is a large technical literature on variable transformation by filtering the data, and in Section 3 we review some of the key contributions in this area. We also compare alternative dating algorithms, highlighting their pros and cons.

In Section 4 we describe simple nonmodel based techniques for the construction of composite coincident or leading indexes. Basically, each component of the index should be carefully selected on the basis of the criteria mentioned above, properly filtered to enhance its business cycle characteristics, deal with seasonal adjustment and remove outliers, and standardized to make its amplitude similar or equal to that of the other index components. The components are then aggregated into the composite index using a certain weighting scheme, typically simple averaging.

From an econometric point of view, composite leading indexes constructed following the procedure sketched above are subject to several criticisms. For example, there is no explicit reference to the target variable in the construction of the composite leading index and the weighting scheme is fixed over time, with periodic revisions mostly due either to data issues, such as changes in the production process of an indicator, or to the past unsatisfactory performance of the index. The main counterpart of these problems is simplicity. Nonmodel based indexes are easy to build, easy to explain, and easy to interpret, which are very valuable assets, in particular for the general public and for policy-makers. Moreover, simplicity is often a plus also for forecasting.

Most of the issues raised for the nonmodel based approach to the construction of composite indexes are addressed by the model based procedures, which can be grouped into two main classes: dynamic factor models and Markov switching models.

Dynamic factor models were developed by Geweke (1977) and Sargent and Sims (1977), but their use became well known to most business cycle analysts after the publication of Stock and Watson's (1989) attempt to provide a formal probabilistic basis for Burns and Mitchell's coincident and leading indicators. The rationale of the approach is that a set of variables is driven by a limited number of common forces, and by idiosyncratic components that are uncorrelated across the variables under analysis. Stock and Watson (1989) estimated a coincident index of economic activity as the unobservable factor in a dynamic factor model for four coincident indicators: industrial production, real disposable income, hours of work and sales.

The main criticism Sims (1989) raised in his comment to Stock and Watson (1989) is the use of a constant parameter statistical model (estimated with classical rather than Bayesian methods). This comment relates to the old debate on the characterization of business cycles as extrinsic phenomena, i.e., generated by the arrival of external shocks propagated through a linear model, versus intrinsic phenomena, i.e., generated by the nonlinear development of the endogenous variables. The main problem with the latter view, at least implicitly supported also by Burns and Mitchell that treated expansions and recessions as two different periods, was the difficulty of casting it into a simple and testable statistical framework, an issue addressed by Hamilton (1989).

Hamilton's (1989) Markov switching model allows the growth rate of the variables (and possibly their dynamics) to depend on the status of the business cycle, which is modelled as a Markov chain. With respect to the factor model based analysis, there is again a single unobservable force underlying the evolution of the indicators but, first, it is discrete rather than continuous and, second, it does not directly affect or summarize the variables but rather indirectly determines their behavior that can change substantially over different phases of the cycle.

As in the case of Stock and Watson (1989), Hamilton (1989) has generated an impressive amount of subsequent research. Here it is worth mentioning the work by Diebold and Rudebusch (1996), which allows the parameters of the factor model in Stock and Watson (1989) to change over the business cycle according to a Markov process. Kim and Nelson (1998) estimated the same model but in a Bayesian framework using the Gibbs sampler, as detailed below, therefore addressing both of Sims' criticisms reported above. Unfortunately, both papers confine themselves to the construction of a coincident indicator and do not consider the issue of leading indicators.

In Sections 5 and 6 we review in detail the competing model based approaches to the construction of composite indexes and discuss their advantages and disadvantages.

In Section 7 we illustrate the practical implementation of the theoretical results by constructing and comparing a set of alternative indexes for the US. We find that all model based coincident indexes are in general very similar and close to the equal weighted ones. As a consequence, the estimation of the current economic condition is rather robust to the choice of method. The model based leading indexes are somewhat different from their nonmodel based counterparts. Their main advantage is that they are derived in a proper statistical framework that, for example, permits the computation of standard errors around the index, the unified treatment of data revisions and missing observations, and the possibility of using time-varying parameters.

In Section 8 we evaluate other approaches for forecasting using leading indicators. In particular, Section 8.1 deals with observed transition models, where the relationship between the target variable and the leading indicators can be made dependent on a set of observable variables, such as GDP growth or the interest rate. Section 8.2 considers neural network and nonparametric methods, where even less stringent hypotheses are imposed on the relationship between the leading indicators and their target. Section 8.3 focuses on the use of binary models for predicting business cycle phases, a topic that attracted considerable attention in the '90s, perhaps as a consequence of the influential study by Diebold and Rudebusch (1989). Finally, Section 8.4 analyzes forecast pooling procedures in the leading indicator context since, starting with the pioneering work of Bates and Granger (1969), it is well known that combining several forecasts can yield more accurate predictions than those of each of the individual forecasts.

In Section 9 we consider the methodological aspects of the evaluation of the forecasting performance of the leading indicators when used either in combination with simple rules to predict turning points [e.g., Vaccara and Zarnowitz (1978)], or as regressors in a model for (a continuous or discrete) target variable. We then discuss a set of empirical examples, to illustrate the theoretical concepts.

A review of the recent literature on the actual performance of leading indicators is contained in Section 10. Four main strands of research can be identified in this literature. First, the consequences of the use of real time information on the composite leading index and its components rather than the final releases. Second, the assessment of the relative performance of the new more sophisticated models for the coincident-leading indicators. Third, the evaluation of financial variables as leading indicators. Finally, the analysis of the behavior of the leading indicators during the two most recent US recessions as dated by the NBER, namely, July 1990–March 1991 and March 2001–November 2001.

To conclude, in Section 11 we summarize what we have learned about leading indicators in the recent past, and suggest directions for further research in this interesting and promising field of forecasting.

2. Selection of the target and leading variables

The starting point for the construction of leading indicators is the choice of the target variable, namely, the variable that the indicators are supposed to lead. Such a choice is discussed in the first subsection. Once the target variable is identified, the leading indicators have to be selected, and we discuss selection criteria in the second subsection.

2.1. Choice of target variable

Burns and Mitchell (1946, p. 3) proposed that:

> "…a cycle consists of expansions occurring at about the same time in many economic activities…"

Yet, later on in the same book (p. 72) they stated:

> "Aggregate activity can be given a definite meaning and made conceptually measurable by identifying it with gross national product."

These quotes underlie the two most common choices of target variable: either a single indicator that is closely related to GDP but available at the monthly level, or a composite index of coincident indicators.

GDP could provide a reliable summary of the current economic conditions if it were available on a monthly basis. Though both in the US and in Europe there is a growing interest in increasing the sampling frequency of GDP from quarterly to monthly, the current results are still too preliminary to rely on.

In the past, industrial production provided a good proxy for the fluctuations of GDP, and it is still currently monitored for example by the NBER business cycle dating committee and by the Conference Board in the US, in conjunction with other indicators. Yet, the ever rising share of services compared with the manufacturing, mining, gas and electric utility industries casts more and more doubts on the usefulness of IP as a single coincident indicator.

Another common indicator is the volume of sales of the manufacturing, wholesale and retail sectors, adjusted for price changes so as to proxy real total spending. Its main drawback, as in the case of IP, is its partial coverage of the economy.

A variable with a close to global coverage is real personal income less transfers, that underlies consumption decisions and aggregate spending. Yet, unusual productivity growth and favorable terms of trade can make income behave differently from payroll employment, the other most common indicator with economy wide coverage. More precisely, the monitored series is usually the number of employees on nonagricultural payrolls, whose changes reflect the net hiring (both permanent and transitory) and firing in the whole economy, with the exception of the smallest businesses and the agricultural sector.

Some authors focused on unemployment rather than employment, e.g., Boldin (1994) or Chin, Geweke and Miller (2000), on the grounds that the series is timely available and subject to minor revisions. Yet, typically unemployment is slightly lagging and not coincident.

Overall, it is difficult to identify a single variable that provides a good measure of current economic conditions, is available on a monthly basis, and is not subject to major later revisions. Therefore, it is preferable to consider combinations of several coincident indicators.

The monitoring of several coincident indicators can be done either informally, for example the NBER business cycle dating committee examines the joint evolution of IP, employment, sales and real disposable income [see, e.g., Hall et al. (2003)], or formally, by combining the indicators into a composite index. A composite coincident index can be constructed in a nonmodel based or in a model based framework, and we will review the main approaches within each category in Sections 4 and 5, respectively.

Once the target variable is defined, it may be necessary to emphasize its cyclical properties by applying proper filters, and/or to transform it into a binary expansion/recession indicator relying on a proper dating procedure. Both issues are discussed in Section 3.

2.2. *Choice of leading variables*

Since the pioneering work of Mitchell and Burns (1938), variable selection has rightly attracted considerable attention in the leading indicator literature; see, e.g., Zarnowitz and Boschan (1975a, 1975b) for a review of early procedures at the NBER and Department of Commerce. Moore and Shiskin (1967) formalized an often quoted scoring system [see, e.g., Boehm (2001), Phillips (1998–1999)], based mostly upon

(i) consistent timing as a leading indicator (i.e., to systematically anticipate peaks and troughs in the target variable, possibly with a rather constant lead time);
(ii) conformity to the general business cycle (i.e., have good forecasting properties not only at peaks and troughs);
(iii) economic significance (i.e., being supported by economic theory either as possible causes of business cycles or, perhaps more importantly, as quickly reacting to negative or positive shocks);
(iv) statistical reliability of data collection (i.e., provide an accurate measure of the quantity of interest);

(v) prompt availability without major later revisions (i.e., being timely and regularly available for an early evaluation of the expected economic conditions, without requiring subsequent modifications of the initial statements);
(vi) smooth month to month changes (i.e., being free of major high frequency movements).

Some of these properties can be formally evaluated at different levels of sophistication. In particular, the peak/trough dates of the target and candidate leading variables can be compared and used to evaluate whether the peak structure of the leading indicator systematically anticipated that of the coincident indicator, with a stable lead time (property (i)). An alternative procedure can be based on the statistical concordance of the binary expansion/recession indicators (resulting from the peak/trough dating) for the coincident and lagged leading variables, where the number of lags of the leading variable can be either fixed or chosen to maximize the concordance. A formal test for no concordance is defined below in Section 9.1. A third option is to run a logit/probit regression of the coincident expansion/recession binary indicator on the leading variable, evaluating the explanatory power of the latter. The major advantage of this procedure is that several leading indicators can be jointly considered to measure their partial contribution. Details on the implementation of this procedure are provided in Section 8.3.

To assess whether a leading indicator satisfies property (ii), conformity to the general business cycle, it is preferable to consider it and the target coincident index as continuous variables rather than transforming them into binary indicators. Then, the set of available techniques includes frequency domain procedures (such as the spectral coherence and the phase lead), and several time domain methods, ranging from Granger causality tests in multivariate linear models, to the evaluation of the marginal predictive content of the leading indicators in sophisticated nonlinear models, possibly with time varying parameters, see Sections 6 and 8 for details on these methods. Within the time domain framework it is also possible to consider a set of additional relevant issues such as the presence of cointegration between the coincident and leading indicators, the determination of the number lags of the leading variable, or the significance of duration dependence. We defer a discussion of these topics to Section 6.

Property (iii), economic significance, can be hardly formally measured, but it is quite important both to avoid the measurement without theory critique, e.g., Koopmans (1947), and to find indicators with stable leading characteristics. On the other hand, the lack of a commonly accepted theory of the origin of business cycles [see, e.g., Fuhrer and Schuh (1998)] makes it difficult to select a single indicator on the basis of its economic significance.

Properties (iv) and (v) have received considerable attention in recent years and, together with economic theory developments, underlie the more and more widespread use of financial variables as leading indicators (due to their exact measurability, prompt availability and absence of revisions), combined with the adoption of real-time datasets for the assessment of the performance of the indicators, see Section 10 for details on these issues. Time delays in the availability of leading indicators are particularly problematic for the construction of composite leading indexes, and have been treated differently in the literature and in practice. Either preliminary values of the composite indexes are constructed excluding the unavailable indicators and later revised, along the tradition of the NBER and later of the Department of Commerce and the Conference Board, or the unavailable observations are substituted with forecasts, as in the factor based approaches described in Section 6.2. The latter solution is receiving increasing favor also within the traditional methodology; see, e.g., McGuckin, Ozyildirim and Zarnowitz (2003). Within the factor based approaches the possibility of measurement error in the components of the leading index, due, e.g., to data revisions, can also be formally taken into account, as discussed in Section 5.1, but in practice the resulting composite indexes require later revisions as well. Yet, both for the traditional and for the more sophisticated methods, the revisions in the composite indexes due to the use of later releases of their components are minor.

The final property (vi), a smooth evolution in the leading indicator, can require a careful choice of variable transformations and/or filter. In particular, the filtering procedures discussed in Section 3 can be applied to enhance the business cycle characteristics of the leading indicators, and in general should be if the target variable is filtered. In general, they can provide improvements with respect to the standard choice of month to month differences of the leading indicator. Also, longer differences can be useful to capture sustained growth or lack of it [see, e.g., Birchenhall et al. (1999)] or differences with respect to the previous peak or trough to take into consideration the possible nonstationary variations of values at turning points [see, e.g., Chin, Geweke and Miller (2000)].

As in the case of the target variable, the use of a single leading indicator is dangerous because economic theory and experience teach that recessions can have different sources and characteristics. For example, the twin US recessions of the early '80s were mostly due to tight monetary policy, that of 1991 to a deterioration in the expectations climate because of the first Iraq war, and that of 2001 to the bursting of the stock market bubble and, more generally, to over-investment; see, e.g., Stock and Watson (2003b). In the Euro area, the three latest recessions according to the CEPR dating are also rather different, with the one in 1974 lasting only three quarters and characterized by synchronization across countries and coincident variables, as in 1992–1993 but contrary to the longer recession that started at the beginning of 1980 and lasted 11 quarters.

A combination of leading indicators into composite indexes can therefore be more useful in capturing the signals coming from different sectors of the economy. The construction of a composite index requires several steps and can be undertaken either in a nonmodel based framework or with reference to a specific econometric model of the evolution of the leading indicators, possibly jointly with the target variable. The two approaches are discussed in Sections 4 and 6, respectively.

3. Filtering and dating procedures

Once the choice of the target measure of aggregate activity (and possibly of the leading indicators) is made, two issues emerge: first the selection of the proper variable transformation, if any, and second the adoption of a dating rule that identifies the peaks and troughs in the series, and the associated expansionary and recessionary periods and their durations.

The choice of the variable transformation is related to the two broad definitions of the cycle recognized in the literature, the so-called classical cycle and the growth or deviation cycle. In the case of the deviation cycle, the focus is on the deviations of the rate of growth of the target variable from an appropriately defined trend rate of growth, while the classical cycle relies on the levels of the target variable.

Besides removing long term movements as in the deviation cycle, high frequency fluctuations can also be eliminated to obtain a filtered variable that satisfies the duration requirement in the original definition of Burns and Mitchell (1946, p. 3):

> "... in duration business cycles vary from more than one year to ten or twelve years; they are not divisible into shorter cycles of similar character with amplitudes approximating their own."

There is a large technical literature on methods of filtering the data. In line with the previous paragraph, Baxter and King (1999) argued that the ideal filter for cycle measurement must be customized to retain unaltered the amplitude of the business cycle periodic components, while removing high and low frequency components. This is known as a *band-pass* filter and, for example, when only cycles with frequency in the range 1.5–8 years are of interest, the theoretical frequency response function of the filter takes the rectangular form: $w(\omega) = \mathrm{I}(2\pi/(8s) \leqslant \omega \leqslant 2\pi/(1.5s))$, where $\mathrm{I}(\cdot)$ is the indicator function. Moreover, the phase displacement of the filter should always be zero, to preserve the timing of peaks and troughs; the latter requirement is satisfied by a symmetric filter.

Given the two business cycle frequencies, $\omega_{c1} = 2\pi/(8s)$ and $\omega_{c2} = 2\pi/(1.5s)$, the band-pass filter is

$$w_{\mathrm{bp}}(L) = \frac{\omega_{c2} - \omega_{c1}}{\pi} + \sum_{j=1}^{\infty} \frac{\sin(\omega_{c2} j) - \sin(\omega_{c1} j)}{\pi j}\left(L^{j} + L^{-j}\right). \tag{1}$$

Thus, the ideal band-pass filter exists and is unique, but it entails an infinite number of leads and lags, so in practice an approximation is required. Baxter and King (1999) showed that the K-terms approximation to the ideal filter (1) that is optimal in the sense of minimizing the integrated squared approximation error is simply (1) truncated at lag K. They proposed using a three year window, i.e., $K = 3s$, as a valid rule of thumb for macroeconomic time series. They also constrained the weights to sum up to zero, so that the resulting approximation is a detrending filter; see, e.g., Stock and Watson (1999a) for an application.

As an alternative, Christiano and Fitzgerald (2003) proposed to project the ideal filter on the available sample. If $c_t = w_{\mathrm{bp}}(L)x_t$ denotes the ideal cyclical component, their proposal is to consider $\hat{c}_t = E(c_t \mid x_1, \ldots, x_T)$, where x_t is given a parametric linear representation, e.g., an ARIMA model. They also found that for a wide class of macroeconomic time series the filter derived under the random walk assumption for x_t is feasible and handy.

Baxter and King (1999) did not consider the problem of estimating the cycle at the extremes of the available sample (the first and last three years), which is inconvenient for a real-time assessment of current business conditions. Christiano and Fitzgerald (2003) suggested to replace the out of sample missing observations by their best linear prediction under the random walk hypothesis. Yet, this can upweight the last and the first available observations.

As a third alternative, Artis, Marcellino and Proietti (2004, AMP) designed a band-pass filter as the difference of two Hodrick and Prescott (1997) detrending filters with parameters $\lambda = 1$ and $\lambda = 677.13$, where these values are selected to ensure that $\omega_{c1} = 2\pi/(8s)$ and $\omega_{c2} = 2\pi/(1.5s)$. The resulting estimates of the cycle are comparable to the Baxter and King cycle, although slightly noisier, without suffering from unavailability of the end of sample estimates.

Working with growth rates of the coincident variables rather than levels, a convention typically adopted for the derivation of composite indexes, corresponds to the application of a filter whose theoretical frequency response function increases monotonically, starting at zero at the zero frequency. Therefore, growth cycles and deviation cycles need not be very similar.

In early post-war decades, especially in Western Europe, growth was relatively persistent and absolute declines in output were comparatively rare; the growth or deviation cycle then seemed to be of more analytical value, especially as inflexions in the rate of growth of output could reasonably be related to fluctuations in the levels of employment and unemployment. In more recent decades, however, there have been a number of instances of absolute decline in output, and popular description at any rate has focussed more on the classical cycle. The concern that de-trending methods can affect the information content of the series in unwanted ways [see, e.g., Canova (1999)] has reinforced the case for examining the classical cycle. The relationships among the three types of cycles are analyzed in more details below, after defining the dating algorithms to identify peaks and troughs in the series and, possibly, transform it into a binary indicator.

In the US, the National Bureau of Economic Research (http://www.nber.org) provides a chronology of the classical business cycle since the early '20s, based on the consensus of a set of coincident indicators concerning production, employment, real income and real sales, that is widely accepted among economists and policy-makers; see, e.g., Moore and Zarnowitz (1986). A similar chronology has been recently proposed for the Euro area by the Center for Economic Policy Research (http://www.cepr.org), see Artis et al. (2003).

Since the procedure underlying the NBER dating is informal and subject to substantial delays in the announcement of the peak and trough dates (which is rational to avoid later revisions), several alternative methods have been put forward and tested on the basis of their ability to closely reproduce the NBER classification.

The simplest approach, often followed by practitioners, is to identify a recession with at least two quarters of negative real GDP growth. Yet, the resulting chronology differs with respect to the NBER in a number of occasions; see, e.g., Watson (1991) or Boldin (1994).

A more sophisticated procedure was developed by Bry and Boschan (1971) and further refined by Harding and Pagan (2002). In particular, for quarterly data on the log-difference of GDP or GNP (Δx_t), Harding and Pagan defined an expansion terminating sequence, ETS_t, and a recession terminating sequence, RTS_t, as follows:

$$\begin{aligned} \mathrm{ETS}_t &= \left\{(\Delta x_{t+1} < 0) \cap (\Delta\Delta x_{t+2} < 0)\right\}, \\ \mathrm{RTS}_t &= \left\{(\Delta x_{t+1} > 0) \cap (\Delta\Delta x_{t+2} > 0)\right\}. \end{aligned} \tag{2}$$

The former defines a candidate point for a peak in the classical business cycle, which terminates the expansion, whereas the latter defines a candidate for a trough. When compared with the NBER dating, usually there are only minor discrepancies. Stock and Watson (1989) adopted an even more complicated rule for identifying peaks and troughs in their composite coincident index.

Within the Markov Switching (MS) framework, discussed in details in Sections 5 and 6, a classification of the observations into two regimes is automatically produced by comparing the probability of being in a recession with a certain threshold, e.g., 0.50. The turning points are then easily obtained as the dates of switching from expansion to recession, or vice versa. Among others, Boldin (1994) reported encouraging results using an MS model for unemployment, and Layton (1996) for the ECRI coincident index. Chauvet and Piger (2003) also confirmed the positive results with a real-time dataset and for a more up-to-date sample period.

Harding and Pagan (2003) compared their nonparametric rule with the MS approach, and further insight can be gained from Hamilton's (2003) comments on the paper and the authors' rejoinder. While the nonparametric rule produces simple, replicable and robust results, it lacks a sound economic justification and cannot be used for probabilistic statements on the current status of the economy. On the other hand, the MS model provides a general statistical framework to analyze business cycle phenomena, but the requirement of a parametric specification introduces a subjective element into the analysis and can necessitate careful tailoring. Moreover, if the underlying model is linear, the MS recession indicator is not identified while pattern recognition works in any case.

AMP developed a dating algorithm based on the theory of Markov chains that retains the attractive features of the nonparametric methods, but allows the computation of the probability of being in a certain regime or of a phase switch. Moreover, the algorithm can be easily modified to introduce depth or amplitude restrictions, and to construct diffusion indices. Basically, the transition probabilities are scored according to the pattern in the series x_t rather than within a parametric MS model. The resulting chronology for the Euro area is very similar to the one proposed by the CEPR, and a similar result emerges for the US with respect to the NBER dating, with the exception of the last recession, see Section 7 below for details.

An alternative parametric procedure to compute the probability of being in a certain cyclical phase is to adopt a probit or logit model where the dependent variable is the

NBER expansion/recession classification, and the regressors are the coincident indicators. For example, Birchenhall et al. (1999) showed that the fit of a logit model is very good in sample when the four NBER coincident indicators are used. They also found that the logit model outperformed an MS alternative, while Layton and Katsuura (2001) obtained the opposite ranking in a slightly different context.

The in-sample estimated parameters from the logit or probit models can also be used in combination with future available values of the coincident indicators to predict the future status of the economy, which is useful, for example, to conduct a real time dating exercise because of the mentioned delays in the NBER announcements.

So far, in agreement with most of the literature, we have classified observations into two phases, recessions and expansions, which are delimited by peaks and troughs in economic activity. However, multiphase characterizations of the business cycle are not lacking in the literature: the popular definition due to Burns and Mitchell (1946) postulated four states: expansion, recession, contraction, recovery; see also Sichel (1994) for an ex-ante three phases characterization of the business cycle, Artis, Krolzig and Toro (2004) for an ex-post three-phases classification based on a model with Markov switching, and Layton and Katsuura (2001) for the use of multinomial logit models.

To conclude, having defined several alternative dating procedures, it is useful to return to the different notions of business cycle and recall a few basic facts about their dating, summarizing results in AMP.

First, neglecting duration ties, classical recessions (i.e., peak-trough dynamics in x_t), correspond to periods of prevailing negative growth, $\Delta x_t < 0$. In effect, negative growth is a sufficient, but not necessary, condition for a classical recession under the Bry and Boschan dating rule and later extensions. Periods of positive growth can be observed during a recession, provided that they are so short lived that they do not determine an exit from the recessionary state.

Second, turning points in x_t correspond to Δx_t crossing the zero line (from above zero if the turning point is a peak, from below in the presence of a trough in x_t). This is strictly true under the calculus rule, according to which $\Delta x_t < 0$ terminates the expansion.

Third, if x_t admits the log-additive decomposition, $x_t = \psi_t + \mu_t$, where ψ_t denotes the deviation cycle, then growth is in turn decomposed into cyclical and residual changes:

$$\Delta x_t = \Delta\psi_t + \Delta\mu_t.$$

Hence, assuming that $\Delta\mu_t$ is mostly due to growth in trend output, deviation cycle recessions correspond to periods of growth below potential growth, that is $\Delta x_t < \Delta\mu_t$. Using the same arguments, turning points correspond to Δx_t crossing $\Delta\mu_t$. When the sum of potential growth and cyclical growth is below zero, that is $\Delta\mu_t + \Delta\psi_t < 0$, a classical recession also occurs.

Finally, as an implication of the previous facts, classical recessions are always a subset of deviation cycle recessions, and there can be multiple classical recessionary episodes within a period of deviation cycle recessions. This suggests that an analysis of the deviation cycle can be more informative and relevant also from the economic policy point of view, even though more complicated because of the filtering issues related to the extraction of the deviation cycle.

4. Construction of nonmodel based composite indexes

In the nonmodel based framework for the construction of composite indexes, the first element is the selection of the index components. Each component should satisfy the criteria mentioned in Section 2. In addition, in the case of leading indexes, a balanced representation of all the sectors of the economy should be achieved, or at least of those more closely related to the target variable.

The second element is the transformation of the index components to deal with seasonal adjustment, outlier removal, treatment of measurement error in first releases of indicators subject to subsequent revision, and possibly forecast of unavailable most recent observations for some indicators. These adjustments can be implemented either in a univariate framework, mostly by exploiting univariate time series models for each indicator, or in a multivariate context. In addition, the transformed indicators should be made comparable to be included in a single index. Therefore, they are typically detrended (using different procedures such as differencing, regression on deterministic trends, or the application of more general band-pass filters), possibly smoothed to eliminate high frequency movements (using moving averages or, again, band pass filters), and standardized to make their amplitudes similar or equal.

The final element for the construction of a composite index is the choice of a weighting scheme. The typical choice, once the components have been standardized, is to give them equal weights. This seems a sensible averaging scheme in this context, unless there are particular reasons to give larger weights to specific variables or sectors, depending on the target variable or on additional information on the economic situation; see, e.g., Niemira and Klein (1994, Chapter 3) for details.

A clear illustration of the nonmodel based approach is provided by (a slightly simplified version of) the step-wise procedure implemented by the Conference Board, CB (previously by the Department of Commerce, DOC) to construct their composite coincident index (CCI), see www.conference-board.org for details.

First, for each individual indicator, x_{it}, month-to-month symmetric percentage changes (spc) are computed as $x_{it\_\mathrm{spc}} = 200 * (x_{it} - x_{it-1})/(x_{it} + x_{it+1})$. Second, for each x_{it_spc} a volatility measure, v_i, is computed as the inverse of its standard deviation. Third, each x_{it_spc} is adjusted to equalize the volatility of the components, the standardization factor being $s_i = v_i / \sum_i v_i$. Fourth, the standardized components, $m_{it} = s_i x_{it\_\mathrm{spc}}$, are summed together with equal weights, yielding $m_t = \sum_i m_{it}$. Fifth, the index in levels is computed as

$$\mathrm{CCI}_t = \mathrm{CCI}_{t-1} * (200 + m_t)/(200 - m_t) \tag{3}$$

with the starting condition

$$\mathrm{CCI}_1 = (200 + m_1)/(200 - m_1).$$

Finally, rebasing CCI to average 100 in 1996 yields the $\mathrm{CCI}_{\mathrm{CB}}$.

From an econometric point of view, composite leading indexes (CLI) constructed following the procedure sketched above are subject to several criticisms, some of which are derived in a formal framework in Emerson and Hendry (1996). First, even though the single indicators are typically chosen according to some formal or informal bivariate analysis of their relationship with the target variable, there is no explicit reference to the target variable in the construction of the CLI, e.g., in the choice of the weighting scheme. Second, the weighting scheme is fixed over time, with periodic revisions mostly due either to data issues, such as changes in the production process of an indicator, or to the past unsatisfactory performance of the index. Endogenously changing weights that track the possibly varying relevance of the single indicators over the business cycle and in the presence of particular types of shocks could produce better results, even though their derivation is difficult. Third, lagged values of the target variable are typically not included in the leading index, while there can be economic and statistical reasons underlying the persistence of the target variable that would favor such an inclusion. Fourth, lagged values of the single indicators are typically not used in the index, while they could provide relevant information, e.g., because not only does the point value of an indicator matter but also its evolution over a period of time is important for anticipating the future behavior of the target variable. Fifth, if some indicators and the target variable are cointegrated, the presence of short run deviations from the long run equilibrium could provide useful information on future movements of the target variable. Finally, since the index is a forecast for the target variable, standard errors should also be provided, but their derivation is virtually impossible in the nonmodel based context because of the lack of a formal relationship between the index and the target.

The main counterpart of these problems is simplicity. Nonmodel based indexes are easy to build, easy to explain, and easy to interpret, which are very valuable assets, in particular for the general public and for policy-makers. Moreover, simplicity is often a plus also for forecasting. With this method there is no estimation uncertainty, no major problems of overfitting, and the literature on forecast pooling suggests that equal weights work pretty well in practice [see, e.g., Stock and Watson (2003a)] even though here variables rather than forecasts are pooled.

Most of the issues raised for the nonmodel based composite indexes are addressed by the model based procedures described in the next two sections, which in turn are in general much more complicated and harder to understand for the general public. Therefore, while from the point of view of academic research and scientific background of the methods there is little to choose, practitioners may well decide to base their preferences on the practical forecasting performance of the two approaches to composite index construction.

5. Construction of model based composite coincident indexes

Within the model based approaches for the construction of a composite coincident index (CCI), two main methodologies have emerged: dynamic factor models and Markov switching models. In both cases there is a single unobservable force underlying the current status of the economy, but in the former approach this is a continuous variable, while in the latter it is a discrete variable that evolves according to a Markov chain. We now review these two methodologies, highlighting their pros and cons.

5.1. Factor based CCI

Dynamic factor models were developed by Geweke (1977) and Sargent and Sims (1977), but their use became well known to most business cycle analysts after the publication of Stock and Watson's (1989, SW) attempt to provide a formal probabilistic basis for Burns and Mitchell's coincident and leading indicators, with subsequent refinements of the methodology in Stock and Watson (1991, 1992). The rationale of the approach is that a set of variables is driven by a limited number of common forces and by idiosyncratic components that are either uncorrelated across the variables under analysis or in any case common to only a limited subset of them. The particular model that SW adopted is the following:

$$\Delta x_t = \beta + \gamma(L)\,\Delta C_t + u_t, \tag{4}$$

$$D(L)u_t = e_t, \tag{5}$$

$$\phi(L)\,\Delta C_t = \delta + v_t, \tag{6}$$

where x_t includes the (logs of the) four coincident variables used by the CB for their $\mathrm{CCI}_{\mathrm{CB}}$, the only difference being the use of hours of work instead of employment since the former provides a more direct measure of fluctuations in labor input. C_t is the single factor driving all variables, while u_t is the idiosyncratic component; Δ indicates the first difference operator, L is the lag operator and $\gamma(L)$, $D(L)$, $\phi(L)$ are, respectively, vector, matrix and scalar lag polynomials. SW used first differenced variables since unit root tests indicated that the coincident indexes were integrated, but not cointegrated. The model is identified by assuming that $D(L)$ is diagonal and e_t and v_t are mutually and serially uncorrelated at all leads and lags, which ensures that the common and the idiosyncratic components are uncorrelated. Moreover, ΔC_t should affect contemporaneously at least one coincident variable. Notice that the hypothesis of one factor, ΔC_t, does not mean that there is a unique source of aggregate fluctuations, but rather that different shocks have proportional dynamic effects on the variables.

For estimation, the model in (4)–(6) is augmented by the identity

$$C_{t-1} = \Delta C_{t-1} + C_{t-2}, \tag{7}$$

and cast into state-space form. The Kalman filter can then be used to write down the likelihood function, which is in turn maximized to obtain parameter and factor estimates, all the details are presented in Stock and Watson (1991).

A few additional comments are in order. First, the composite coincident index, $\mathrm{CCI}_{\mathrm{SW}t}$, is obtained through the Kalman filter as the minimum mean squared error linear estimator of C_t using information on the coincident variables up to period t. Hence, the procedure can be implemented in real time, conditional on the availability of data on the coincident variables. By using the Kalman smoother rather than the filter, it is possible to obtain end of period estimates of the state of the economy, i.e., $C_{t|T}$. Second, it is possible to obtain a direct measure of the contribution of each coincident indicator in x_t to the index by computing the response of the latter to a unit impulse in the former. Third, since data on some coincident indicator are published with delay, they can be treated as missing observations and estimated within the state-space framework. Moreover, the possibility of measurement error in the first releases of the coincident indicators can also be taken into consideration by adding an error term to the measurement equation. This is an important feature since data revisions are frequent and can be substantial, for example as testified by the revised US GDP growth rate data for 2001. Fourth, a particular time varying pattern in the parameters of the lag polynomials $D(L)$ and $\phi(L)$ can be allowed by using a time-varying transition matrix. Fifth, standard errors around the coincident index can be computed, even though they were not reported by SW.

The cyclical structure of $\mathrm{CCI}_{\mathrm{SW}}$ closely follows the NBER expansions and recessions, and the correlation of two quarters growth rates in $\mathrm{CCI}_{\mathrm{SW}}$ and real GDP was about 0.86 over the period 1959–1987. Stock and Watson (1991) also compared their $\mathrm{CCI}_{\mathrm{SW}}$ with the DOC's one, finding that the overall relative importance of the single indicators is roughly similar (but the weights are different since the latter index is made up of contemporaneous indicators only), the correlation of the levels of the composite indexes was close to 0.94, again over the period 1959–1987, and the coherence of their growth rates at business cycle frequency was even higher.

These findings provide support for the simple averaging methodology originated at the NBER and then further developed at the DOC and the CB, but they also question the practical usefulness of the SW's approach, which is substantially more complicated. Overall, the SW methodology, and more generally model based index construction, are worth their cost since they provide a proper statistical framework that, for example, permits the computation of standard errors around the composite index, the unified treatment of data revisions and missing observations, the possibility of using time-varying parameters and, as we will see in more detail in the next section, a coherent framework for the development of composite leading indexes.

A possible drawback of SW's procedure is that it requires an ex-ante classification of variables into coincident and leading or lagging, even though this is common practice in this literature, and it cannot be directly extended to analyze large datasets because of computational problems, see Section 6.2 for details. Forni et al. (2000, 2001, FHLR henceforth) proposed an alternative factor based methodology that addresses both issues, and applied it to the derivation of a composite coincident indicator for the Euro area. They analyzed a large set of macroeconomic time series for each country of the Euro area using a dynamic factor model, and decomposed each time series into a common and an idiosyncratic component, where the former is the part of the variable explained by common Euro area shocks, the latter by variable specific shocks. The $\mathrm{CCI}_{\mathrm{FHLR}}$ is obtained as a weighted average of the common components of the interpolated monthly GDP series for each country, where the weights are proportional to GDP, and takes into account both within and across-countries cross correlations.

More specifically, the model FHLR adopted is

$$x_{it} = b_i'(L)v_t + \xi_{it}, \quad i = 1, \ldots, N, \; t = 1, \ldots, T, \tag{8}$$

where x_{it} is a stationary univariate random variable, v_t is a $q \times 1$ vector of common shocks, $\chi_{it} = x_{it} - \xi_{it}$ is the common component of x_{it}, and ξ_{it} is its idiosyncratic component. The shock v_t is an orthonormal white noise process, so that $\mathrm{var}(v_{jt}) = 1$, $\mathrm{cov}(v_t, v_{t-k}) = 0$, and $\mathrm{cov}(v_{jt}, v_{st-k}) = 0$ for any $j \neq s$, t and k. $\xi_N = \{\xi_{1t}, \ldots, \xi_{Nt}\}'$ is a wide sense stationary process, and $\mathrm{cov}(\xi_{jt}, v_{st-k}) = 0$ for any j, s, t and k. $b_i(L)$ is a $q \times 1$ vector of square summable, bilateral filters, for any i. Notice that SW's factor model (4) is obtained as a particular case of (8) when there is one common shock ($q = 1$), $b_i(L) = \gamma_i(L)/\phi(L)$, and the idiosyncratic components are assumed to be orthogonal.

Grouping the variables into $x_{Nt} = \{x_{1t}, \ldots, x_{Nt}\}'$, FHLR also required x_{Nt} (and χ_{Nt}, ξ_{Nt} that are similarly defined) to have rational spectral density matrices, Σ_N^x, Σ_N^χ, and Σ_N^ξ, respectively. To achieve identification, they assumed that the first (largest) idiosyncratic dynamic eigenvalue, λ_{N1}^ξ, is uniformly bounded, and that the first (largest) q common dynamic eigenvalues, $\lambda_{N1}^\chi, \ldots, \lambda_{Nq}^\chi$, diverge when N increases, where dynamic eigenvalues are the eigenvalues of the spectral density matrix; see, e.g., Brillinger (1981, Chapter 9). In words, the former condition limits the effects of ξ_{it} on other cross-sectional units. The latter, instead, requires v_t to affect infinitely many units.

Assuming that the number of common shocks is known, FHLR suggested to estimate the common component of χ_{it}, $\hat{\chi}_{it}$, as the projection of x_{it} on past, present, and future dynamic principal components of all variables, and proved that, under mild conditions, $\hat{\chi}_{it}$ is a consistent estimator of χ_{it} when N and T diverge. Once the common component is estimated, the idiosyncratic one is obtained simply as a residual, namely, $\hat{\xi}_{it} = x_{it} - \hat{\chi}_{it}$.

To determine the number of factors, q, FHLR suggested to exploit two features of the model: (a) the average over frequencies of the first q dynamic eigenvalues diverges, while the average of the $(q+1)$th does not; (b) there should be a big gap between the variance of x_{Nt} explained by the first q dynamic principal components and that explained by the $(q+1)$th principal component. As an alternative, an information criterion could be used, along the lines of Bai and Ng (2002).

The methodology was further refined by Altissimo et al. (2001) and Forni et al. (2003a) for real time implementation, and it is currently adopted to produce the CEPR's composite coincident indicator for the Euro area, Eurocoin (see www.cepr.org). In particular, they exploited the large cross-sectional dimension for forecasting indicators available with delay and for filtering out high frequency dynamics. Alternative coincident indexes for the Euro area following the SW methodology were proposed by

Proietti and Moauro (2004), while Carriero and Marcellino (2005) compared several methodologies, finding that they yield very similar results.

5.2. Markov switching based CCI

The main criticism Sims (1989) raised in his comment to Stock and Watson (1989) is the use of a constant parameter model (even though, as remarked above, their framework is flexible enough to allow for parameter variation), and a similar critique can be addressed to FHLR's method. Hamilton's (1989) Markov switching model is a powerful response to this criticism, since it allows the growth rate of the variables (and possibly their dynamics) to depend on the status of the business cycle. A basic version of the model can be written as

$$\Delta x_t = c_{s_t} + A_{s_t} \Delta x_{t-1} + u_t, \tag{9}$$

$$u_t \sim \text{i.i.d. } N(0, \Sigma), \tag{10}$$

where, as in (4), x_t includes the coincident variables under analysis (or a single composite index), while s_t measures the status of the business cycle, with $s_t = 1$ in recessions and $s_t = 0$ in expansions, and both the deterministic component and the dynamics can change over different business cycle phases. The binary state variable s_t is not observable, but the values of the coincident indicators provide information on it.

With respect to the factor model based analysis, there is again a single unobservable force underlying the evolution of the indicators but, first, it is discrete rather than continuous and, second, it does not directly affect or summarize the variables but rather indirectly determines their behavior that can change substantially over different phases of the cycle.

To close the model and estimate its parameters, an equation describing the behavior of s_t is required, and it cannot be of autoregressive form as (6) since s_t is a binary variable. Hamilton (1989) proposed to adopt the Markov switching (MS) model, where

$$\Pr(s_t = j \mid s_{t-1} = i) = p_{ij}, \tag{11}$$

as previously considered by Lindgren (1978) and Neftci (1982) in simpler contexts. For expositional purposes we stick to the two states hypothesis, though there is some empirical evidence that three states can further improve the specification, representing recession, high growth and normal growth; see, e.g., Kim and Murray (2002) for the US and Artis, Krolzig and Toro (2004) for the Euro area.

In our business cycle context, the quantity of special interest is an estimate of the unobservable current status of the economy and, assuming a mean square error loss function, the best estimator coincides with the conditional expectation of s_t given current and past information on x_t, which in turn is equivalent to the conditional probability

$$\zeta_{t|t} = \begin{pmatrix} \Pr(s_t = 0 \mid x_t, x_{t-1}, \ldots, x_1) \\ \Pr(s_t = 1 \mid x_t, x_{t-1}, \ldots, x_1) \end{pmatrix}. \tag{12}$$

Using simple probability rules, it follows that

$$\zeta_{t|t} = \begin{pmatrix} \frac{f(x_t \mid s_t=0,\ x_{t-1},\ldots,x_1)\Pr(s_t=0 \mid x_{t-1},\ldots,x_1)}{f(x_t \mid x_{t-1},\ldots,x_1)} \\ \frac{f(x_t \mid s_t=1,\ x_{t-1},\ldots,x_1)\Pr(s_t=1 \mid x_{t-1},\ldots,x_1)}{f(x_t \mid x_{t-1},\ldots,x_1)} \end{pmatrix}, \tag{13}$$

where

$$\begin{aligned}
&\Pr(s_t = i \mid x_{t-1},\ldots,x_1) = \sum_{j=0}^{1} p_{ji} \Pr(s_{t-1} = j \mid x_{t-1},\ldots,x_1),\\
&f(x_t \mid s_t = i,\ x_{t-1},\ldots,x_1)\\
&\quad = \frac{1}{(2\pi)^{T/2}} |\Sigma|^{-1/2}\\
&\qquad \times \exp\bigl[-(\Delta x_t - c_i - A_i\,\Delta x_{t-1})'\Sigma^{-1}(\Delta x_t - c_i - A_i\,\Delta x_{t-1})/2\bigr],\\
&f(x_t, s_t = i \mid x_{t-1},\ldots,x_1)\\
&\quad = f(x_t \mid s_t = i,\ x_{t-1},\ldots,x_1)\Pr(s_t = i \mid x_{t-1},\ldots,x_1),\\
&f(x_t \mid x_{t-1},\ldots,x_1) = \sum_{j=0}^{1} f(x_t, s_t = j \mid x_{t-1},\ldots,x_1), \quad i = 0, 1.
\end{aligned} \tag{14}$$

Hamilton (1994) or Krolzig (1997) provide additional details on these computations, and formulae to calculate $\zeta_{t|T}$, i.e., the smoothed estimator of the probability of being in a given status in period t. Notice also that the first and last rows of (14) provide, respectively, the probability of the state and the density of the variables conditional on past information only, that will be used in Section 6.3 in a related context for forecasting.

For comparison and since it is rather common in empirical applications [see, e.g., Niemira and Klein (1994) for the US and Artis et al. (1995) for the UK], it is useful to report Neftci's (1982) formula to compute the (posterior) probability of a turning point given the available data, as refined by Diebold and Rudebusch (1989). Defining

$$\Pi_t = \Pr(s_t = 1 \mid x_t,\ldots,x_1), \tag{15}$$

the formula is

$$\begin{aligned}
&\Pi_t = \frac{A_1}{B_1 + C_1},\\
&A_1 = \bigl(\Pi_{t-1} + p_{01}(1 - \Pi_{t-1})\bigr) f(x_t \mid s_t = 1,\ x_{t-1},\ldots,x_1),\\
&B_1 = \bigl(\Pi_{t-1} + p_{01}(1 - \Pi_{t-1})\bigr) f(x_t \mid s_t = 1,\ x_{t-1},\ldots,x_1),\\
&C_1 = (1 - \Pi_{t-1})(1 - p_{01}) f(x_t \mid s_t = 0,\ x_{t-1},\ldots,x_1).
\end{aligned} \tag{16}$$

The corresponding second element of $\zeta_{t|t}$ in (13) can be written as

$$\begin{aligned}
\Pi_t &= \frac{A_2}{B_2 + C_2}, \\
A_2 &= \big(\Pi_{t-1} - \Pi_{t-1} p_{01} + p_{01}(1 - \Pi_{t-1})\big) f(x_t \mid s_t = 1,\ x_{t-1}, \ldots, x_1), \\
B_2 &= \big(\Pi_{t-1} - \Pi_{t-1} p_{01} + p_{01}(1 - \Pi_{t-1})\big) f(x_t \mid s_t = 1,\ x_{t-1}, \ldots, x_1), \\
C_2 &= \big((1 - \Pi_{t-1})(1 - p_{01}) + \Pi_{t-1} p_{01}\big) f(x_t \mid s_t = 0,\ x_{t-1}, \ldots, x_1).
\end{aligned} \tag{17}$$

Since in practice the probability of transition from expansion to recession, p_{01}, is very small [e.g., Diebold and Rudebusch (1989) set it at 0.02], the term $\Pi_{t-1} p_{01}$ is also very small and the two probabilities in (16) and (17) are very close. Yet, in general it is preferable to use the expression in (17) which is based on a more general model. Notice also that when $\Pi_t = 1$ the formula in (16) gives a constant value of 1 [e.g., Diebold and Rudebusch (1989) put an ad hoc upper bound of 0.95 for the value that enters the recursive formula], while this does not happen with (17).

The model in (9)–(11) can be extended in several dimensions, for example, to allow for more states and cointegration among the variables [see, e.g., Krolzig, Marcellino and Mizon (2002)] or time-varying probabilities as, e.g., in Diebold, Lee and Weinbach (1994) or Filardo (1994). The latter case is of special interest in our context when past values of the leading indicators, y, are the driving forces of the probabilities, as in Filardo (1994), who substituted (11) with

$$\Pr(s_t = i \mid s_{t-1} = j,\ x_{t-1}, \ldots, x_1,\ y_{t-1}, \ldots, y_1) = \frac{\exp(\theta y_{t-1})}{1 + \exp(\theta y_{t-1})}, \tag{18}$$

so that the first row of (14) should be modified into

$$\begin{aligned}
\Pr(s_t = i \mid x_{t-1}, \ldots, x_1) = {} & \frac{\exp(\theta y_{t-1})}{1 + \exp(\theta y_{t-1})} \Pr(s_{t-1} = j \mid x_{t-1}, \ldots, x_1) \\
& + \frac{1}{1 + \exp(\theta y_{t-1})} \Pr(s_{t-1} = i \mid x_{t-1}, \ldots, x_1).
\end{aligned} \tag{19}$$

Another example is provided by Filardo and Gordon (1998), who used a probit model rather than a logistic specification for $\Pr(s_t = i \mid s_{t-1} = j,\ x_{t-1}, \ldots,\ x_1,\ y_{t-1}, \ldots, y_1)$, while Ravn and Sola (1999) warned against possible parameter instability of relationships such as (18). Raj (2002) provides a more detailed review of these and other extensions of the MS model.

Factor models and Markov switching specifications capture two complementary and fundamental features of business cycles, namely, the diffusion of slow-down and recovery across many series and the different behavior of several indicators in expansions and recessions. They are not only flexible and powerful statistical tools but can also be given sound justifications from an economic theory point of view; see, e.g., the overview in Diebold and Rudebusch (1996). The latter article represents also one of the earliest attempts to combine the two approaches, by allowing the factor underlying SW's model to evolve according to a Markov switching model. To provide support for their ideas, they fitted univariate and multivariate MS models to, respectively, the DOC's composite coincident indicator and its components, finding substantial evidence in favor of the MS specifications. Yet, they did not jointly estimate the factor MS model. Such a task was tackled by Chauvet (1998) and Kim and Yoo (1995), using an approximated maximum likelihood procedure developed by Kim (1994), and by Kim and Nelson (1998) and Filardo and Gordon (1999) using Gibbs sampler techniques introduced by Albert and Chib (1993a), Carter and Kohn (1994), and Shepard (1994).

In particular, Kim and Nelson (1998) substituted Equation (6) in SW's model with

$$\begin{aligned} &\phi(L)(\Delta C_t - \mu_{s_t} - \delta) = v_t, \\ &\mu_{s_t} = \mu_0 + \mu_1 s_t, \end{aligned} \tag{20}$$

where the transition probabilities are either constant or follow a probit specification. They compared the (posterior) regime probabilities from the factor MS model estimated with the four SW's components with those from a univariate MS model for IP, concluding that the former are much more closely related with the NBER expansion/recession classification. Yet, such a result is not surprising since Filardo (1994) showed that time-varying probabilities are needed for the univariate MS model to provide a close match with the NBER classification. When the original SW's model is estimated using the Gibbs sampling approach, the posterior distributions of the parameters are very close to those obtained using (20) instead of (6), the main difference being a slightly larger persistence of the estimated factor. Filardo and Gordon (1999), focusing on the 1990 recession, also found a similar performance of the standard and MS factor model, while a multivariate MS model with time-varying probabilities performed best during the recessionary part of 1990 (but not significantly better in the remaining months). Finally, Kim and Nelson (1998) also found a close similarity of their composite coincident indicator and the equal weighted DOC's one, with correlation in the growth rates above 0.98.

Finally, notice that if the probability of the states is time varying, e.g., as in (18), and the indicators in y_t include a measure of the length of the current recession (or expansion), it is possible to allow and test for duration dependence, namely, for whether the current or past length of a business cycle phase influences its future duration. The test is based on the statistical significance of the parameter associated with the duration indicator in an equation such as (18). Earlier studies using nonparametric techniques, such as Diebold and Rudebusch (1990) or Diebold, Rudebusch and Sichel (1993), detected positive duration dependence for recessions but not for expansions. Such a finding was basically confirmed by Durland and McCurdy (1994) using a semi-Markov model with duration depending only on calendar time, by Filardo and Gordon (1998) in a univariate Markov switching framework that also relates duration to macroeconomic variables, and by Kim and Nelson (1998) in their multivariate factor MS model. Therefore, another interesting question to be addressed in Sections 6 and 8 is whether leading indicators can be used to predict the duration of a business cycle phase.

In summary, no clear cut ranking of the multivariate model based approaches to CCI construction emerges, but the resulting indexes are in general very similar and close to the equal weighted ones, as we will see in the examples of Section 7. The positive aspect of this result is that estimation of the current economic condition is rather robust to the choice of method. Another implication is that pooling methods can be expected to yield no major improvements because of high correlation of all the indicators; see, e.g., Carriero and Marcellino (2005), but this is an issue that certainly deserves further investigation.

6. Construction of model based composite leading indexes

Leading indicators are hardly of any use without a rule to transform them into a forecast for the target variable. These rules range from simple nonparametric procedures that monitor the evolution of the leading indicator and transform it into a recession signal, e.g., the three-consecutive-declines in the $\mathrm{CLI}_{\mathrm{CB}}$ rule [e.g., Vaccara and Zarnowitz (1978)], to sophisticated nonlinear models for the joint evolution of the leading indicators and the target variable, which can be used to predict growth rates, turning points, and expected duration of a certain business cycle phase. In this section we discuss the methods that are directly related to those reviewed in the previous section in the context of CCIs. In particular, Section 6.1 deals with linear models, Section 6.2 with factor based models, and Section 6.3 with Markov switching models. Examples are provided in the next section, while other approaches are considered in Section 8 below.

6.1. VAR based CLI

A linear VAR provides the simplest model based framework to understand the relationship between coincident and leading indicators, the construction of regression based composite leading indexes, the role of the latter in forecasting, and the consequences of invalid restrictions or unaccounted cointegration.

Let us group the m coincident indicators in the vector x_t, and the n leading indicators in y_t. For the moment, we assume that (x_t, y_t) is weakly stationary and its evolution is described by the VAR(1):

$$\begin{aligned}\begin{pmatrix} x_t \\ y_t \end{pmatrix} &= \begin{pmatrix} c_x \\ c_y \end{pmatrix} + \begin{pmatrix} A & B \\ C & D \end{pmatrix}\begin{pmatrix} x_{t-1} \\ y_{t-1} \end{pmatrix} + \begin{pmatrix} e_{xt} \\ e_{yt} \end{pmatrix},\\ \begin{pmatrix} e_{xt} \\ e_{yt} \end{pmatrix} &\sim \text{i.i.d.}\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \Sigma_{xx} & \Sigma_{xy} \\ \Sigma_{yx} & \Sigma_{yy} \end{pmatrix}\right).\end{aligned} \tag{21}$$

It immediately follows that the expected value of x_{t+1} conditional on the past is

$$E(x_{t+1} \mid x_t, x_{t-1}, \ldots,\ y_t, y_{t-1}, \ldots) = c_x + Ax_t + By_t, \tag{22}$$

so that for y to be a useful set of leading indicators it must be $B \neq 0$. When $A \neq 0$, lagged values of the coincident variables also contain useful information for forecasting. Both hypotheses are easily testable and, in case both $A = 0$ and $B = 0$ are rejected, a composite regression based leading indicator for x_{t+1} (considered as a vector) can be constructed as

$$\mathrm{CLI1}_t = \hat{c}_x + \hat{A} x_t + \widehat{B} y_t, \tag{23}$$

where the ˆ indicates the OLS estimator. Standard errors around this CLI can be constructed using standard methods for VAR forecasts; see, e.g., Lütkepohl (2006). Moreover, recursive estimation of the model provides a convenient tool for continuous updating of the weights.

A similar procedure can be followed when the target variable is dated $t + h$ rather than t. For example, when $h = 2$,

$$\begin{aligned} \mathrm{CLI1}_t^{h=2} &= \hat{c}_x + \hat{A}\hat{x}_{t+1|t} + \widehat{B}\hat{y}_{t+1|t} \\ &= \hat{c}_x + \hat{A}\big(\hat{c}_x + \hat{A}x_t + \widehat{B}y_t\big) + \widehat{B}\big(\hat{c}_y + \widehat{C}x_t + \widehat{D}y_t\big). \end{aligned} \tag{24}$$

As an alternative, the model in (21) can be rewritten as

$$\begin{pmatrix} x_t \\ y_t \end{pmatrix} = \begin{pmatrix} \tilde{c}_x \\ \tilde{c}_y \end{pmatrix} + \begin{pmatrix} \tilde{A} & \widetilde{B} \\ \widetilde{C} & \widetilde{D} \end{pmatrix} \begin{pmatrix} x_{t-h} \\ y_{t-h} \end{pmatrix} + \begin{pmatrix} \tilde{e}_{xt} \\ \tilde{e}_{yt} \end{pmatrix}, \tag{25}$$

where a ˜ indicates that the new parameters are a combination of those in (21), and $\tilde{e}_{xt}$ and $\tilde{e}_{yt}$ are correlated of order $h - 1$. Specifically,

$$\begin{aligned} &\begin{pmatrix} \tilde{c}_x \\ \tilde{c}_y \end{pmatrix} = \left(I + \begin{pmatrix} A & B \\ C & D \end{pmatrix} + \cdots + \begin{pmatrix} A & B \\ C & D \end{pmatrix}^{h-1} \right) \begin{pmatrix} c_x \\ c_y \end{pmatrix}, \\ &\begin{pmatrix} \tilde{A} & \widetilde{B} \\ \widetilde{C} & \widetilde{D} \end{pmatrix} = \begin{pmatrix} A & B \\ C & D \end{pmatrix}^{h}, \\ &\begin{pmatrix} \tilde{e}_{xt} \\ \tilde{e}_{yt} \end{pmatrix} = \left(I + \begin{pmatrix} A & B \\ C & D \end{pmatrix} + \cdots + \begin{pmatrix} A & B \\ C & D \end{pmatrix}^{h-1} \right) \begin{pmatrix} e_{xt} \\ e_{yt} \end{pmatrix}. \end{aligned} \tag{26}$$

The specification in (25) can be estimated by OLS, and the resulting CLI written as

$$\widetilde{\mathrm{CLI1}}_t^h = \hat{\tilde{c}}_x + \hat{\tilde{A}} x_t + \widehat{\widetilde{B}} y_t. \tag{27}$$

The main disadvantage of this latter method, often called dynamic estimation, is that a different model has to be specified for each forecast horizon h. On the other hand, no model is required for the leading indicators, and the estimators of the parameters in (25) can be more robust than those in (21) in the presence of mis-specification; see, e.g., Clements and Hendry (1996) for a theoretical discussion and Marcellino, Stock and Watson (2005) for an extensive empirical analysis of the two competing methods (showing that dynamic estimation is on average slightly worse than the iterated method for forecasting US macroeconomic time series). For the sake of simplicity, in the rest of the paper we will focus on $h = 1$ whenever possible.

Consider now the case where the target variable is a composite coincident indicator,

$$\mathrm{CCI}_t = w x_t, \tag{28}$$

where w is a $1 \times m$ vector of weights as in Section 4. To construct a model based CLI for the CCI in (28) two routes are available. First, and more common, we could model CCI_t and y_t with a finite order VAR, say

$$\begin{pmatrix} \mathrm{CCI}_t \\ y_t \end{pmatrix} = \begin{pmatrix} d_{\mathrm{CCI}} \\ d_y \end{pmatrix} + \begin{pmatrix} e(L) & F(L) \\ g(L) & H(L) \end{pmatrix} \begin{pmatrix} \mathrm{CCI}_{t-1} \\ y_{t-1} \end{pmatrix} + \begin{pmatrix} u_{\mathrm{CCI}t} \\ u_{yt} \end{pmatrix}, \tag{29}$$

where L is the lag operator and the error process is white noise. Repeating the previous procedure, the composite leading index for $h = 1$ is

$$\mathrm{CLI2}_t = \hat{d}_{\mathrm{CCI}} + \hat{e}(L)\mathrm{CCI}_t + \widehat{F}(L)y_t. \tag{30}$$

Yet, in this case the VAR is only an approximation for the generating mechanism of (wx_t, y_t), since in general the latter should have an infinite number of lags or an MA component.

The alternative route is to stick to the model in (21), and construct the CLI as

$$\mathrm{CLI3}_t = w\mathrm{CLI1}_t, \tag{31}$$

namely, aggregate the composite leading indicators for each of the components of the CCI, using the same weights as in the CCI. Lütkepohl (1987) showed in a related context that in general aggregating the forecasts (CLI3) is preferable than forecasting the aggregate (CLI2) when the variables are generated by the model in (21), while this is not necessarily the case if the model in (21) is also an approximation and/or the x variables are subject to measurement error; see also Lütkepohl (2006). Stock and Watson (1992) overall found little difference in the performance of CLI2 and CLI3.

Both CLI2 and CLI3 are directly linked to the target variable, incorporate distributed lags of both the coincident and the leading variables (depending on the lag length of the VAR), the weights can be easily periodically updated using recursive estimation of the model, and standard errors around the point forecasts (or the whole distribution under a distributional assumption for the error process in the VAR) are readily available. Therefore, this simple linear model based procedure already addresses several of the main criticisms to the nonmodel based composite index construction, see Section 4.

In this context the dangers of using a simple average of the y variables as a composite leading index are also immediately evident, since the resulting index can provide an inefficient forecast of the CCI unless specific restrictions on the VAR coefficients in (21) are satisfied. In particular, indicating by i_n a $1 \times n$ vector with elements equal to $1/n$, the equal weight composite leading index

$$\mathrm{CLI}_{\mathrm{EW}t} = i_n y_t \tag{32}$$

is optimal and coincides with CLI3 if and only if

$$wc_x = 0, \quad wA = 0, \quad wB = i_n, \tag{33}$$

which imposes $1 + m + n$ restrictions on the parameters of the x equations in (21). In higher order VARs, the product of the weights w and the coefficients of longer lags of x and y in the x equations should also be equal to zero. Notice that these are all testable assumptions as long as $m+n$ is small enough with respect to the sample size to leave sufficient degrees of freedom for the VAR parameter estimation. For example, in the case of the Conference Board, $m+n=14$ and monthly data are available for about 45 years for a total of more than 500 observations. Auerbach (1982) found that a regression based CLI in sample performed better than the equal weighted $\mathrm{CLI}_{\mathrm{CB}}$ for industrial production and the unemployment rate, but not out of sample.

If the restrictions in (33) are not satisfied but it is desired to use in any case $\mathrm{CLI}_{\mathrm{EW}}$ (or more generally a given CLI) to forecast the CCI, it can be possible to improve upon its performance by constructing a VAR for the two composite indexes CCI and $\mathrm{CLI}_{\mathrm{EW}}$ $(wx_t, i_n y_t)$, say

$$\begin{pmatrix} \mathrm{CCI}_t \\ \mathrm{CLI}_{\mathrm{EW}t} \end{pmatrix} = \begin{pmatrix} f_{\mathrm{CCI}} \\ f_{\mathrm{CLI}_{\mathrm{EW}}} \end{pmatrix} + \begin{pmatrix} e(L) & f(L) \\ g(L) & h(L) \end{pmatrix} \begin{pmatrix} \mathrm{CCI}_{t-1} \\ \mathrm{CLI}_{\mathrm{EW}t-1} \end{pmatrix} + \begin{pmatrix} v_{\mathrm{CCI}t} \\ v_{\mathrm{CLI}_{\mathrm{EW}}t} \end{pmatrix} \tag{34}$$

and construct the new composite index as

$$\mathrm{CLI4}_t = \hat{f}_{\mathrm{CCI}} + \hat{e}(L)\mathrm{CCI}_t + \hat{f}(L)\mathrm{CLI}_{\mathrm{EW}t}. \tag{35}$$

This is for example the methodology adopted by Koch and Rasche (1988), who analyzed a VAR for IP, as a coincident indicator, and the equal weighted DOC leading index. Since CLI4 has a dynamic structure and also exploits past information in the CCI, it can be expected to improve upon $\mathrm{CLI}_{\mathrm{EW}}$. Moreover, since the VAR in (34) is much more parsimonious than both (21) and (29), CLI4 could perform in practice even better than the other composite indexes, in particular in small samples.

A point that has not gained attention in the literature but can be of importance is the specification of the equations for the (single or composite) leading indicators. Actually, in all the models we have considered so far, the leading variables depend on lags of the coincident ones, which can be an unreliable assumption from an economic point of view. For example, the interest rate spread depends on future expected short term-interest rates and the stock market index on future expected profits and dividends, and these expectations are positively and highly correlated with the future expected overall economic conditions. Therefore, the leading variables could depend on future expected coincident variables rather than on their lags. For example, the equations for y_t in the model for (x_t, y_t) in (21) could be better specified as

$$y_t = c_y + Cx^e_{t+1|t-1} + Dy_{t-1} + e_{yt}, \tag{36}$$

where $x^e_{t+1|t-1}$ indicates the expectation of x_{t+1} conditional on information available in period $t-1$. Combining these equations with those for x_t in (21), it is possible to obtain a closed form expression for $x^e_{t+1|t-1}$, which is

$$x^e_{t+1|t-1} = (I - BC)^{-1}\big(c_x + Ac_x + Bc_y + A^2 x_{t-1} + (AB + BD)y_{t-1}\big). \tag{37}$$

Therefore, a VAR specification such as that in (21) can also be considered as a reduced form of a more general model where the leading variables depend on expected future coincident variables. A related issue is whether the coincident variables, x_t, could also depend on their future expected values, as it often results in new-Keynesian models; see, e.g., Walsh (2003). Yet, the empirical evidence in Fuhrer and Rudebusch (2004) provides little support for this hypothesis.

Another assumption we have maintained so far is that both the coincident and the leading variables are weakly stationary, while in practice it is likely that the behavior of most of these variables is closer to that of integrated process. Following Sims, Stock and Watson (1990), this is not problematic for consistent estimation of the parameters of VARs in levels such as (21), and therefore for the construction of the related CLIs, even though inference is complicated and, for example, hypotheses on the parameters such as those in (33) could not be tested using standard asymptotic distributions. An additional complication is that in this literature, when the indicators are I(1), the VAR models are typically specified in first differences rather than in levels, without prior testing for cointegration. Continuing the VAR(1) example, the adopted model would be

$$\begin{pmatrix} \Delta x_t \\ \Delta y_t \end{pmatrix} = \begin{pmatrix} c_x \\ c_y \end{pmatrix} + \begin{pmatrix} e_{xt} \\ e_{yt} \end{pmatrix}, \tag{38}$$

rather than possibly

$$\begin{aligned} \begin{pmatrix} \Delta x_t \\ \Delta y_t \end{pmatrix} &= \begin{pmatrix} c_x \\ c_y \end{pmatrix} - \left(\begin{pmatrix} I_m & 0 \\ 0 & I_n \end{pmatrix} - \begin{pmatrix} A & B \\ C & D \end{pmatrix} \right) \begin{pmatrix} x_{t-1} \\ y_{t-1} \end{pmatrix} + \begin{pmatrix} e_{xt} \\ e_{yt} \end{pmatrix} \\ &= \begin{pmatrix} c_x \\ c_y \end{pmatrix} - \alpha\beta' \begin{pmatrix} x_{t-1} \\ y_{t-1} \end{pmatrix} + \begin{pmatrix} e_{xt} \\ e_{yt} \end{pmatrix}, \end{aligned} \tag{39}$$

where β is the matrix of cointegrating coefficients and α contains the loadings of the error correction terms. As usual, omission of relevant variables yields biased estimators of the parameters of the included regressors, which can translate into biased and inefficient composite leading indicators. See Emerson and Hendry (1996) for additional details and generalizations and, e.g., Clements and Hendry (1999) for the consequences of omitting cointegrating relations when forecasting. As long as $m + n$ is small enough with respect to the sample size, the number and composition of the cointegrating vectors can be readily tested [see, e.g., Johansen (1988) for tests within the VAR framework], and the specification in (39) used as a basis to construct model based CLIs that also take cointegration into proper account. Hamilton and Perez-Quiros (1996) found cointegration to be important for improving the forecasting performance of the $\mathrm{CLI}_{\mathrm{DOC}}$.

Up to now we have implicitly assumed, as it is common in most of the literature that analyzes CCIs and CLIs within linear models, that the goal of the composite leading index is forecasting a continuous variable, the CCI. Yet, leading indicators were originally developed for forecasting business cycle turning points. Simulation based methods can be used to derive forecasts of a binary recession/expansion indicator, and these in turn can be exploited to forecast the probability that a recession will take place within, or at, a certain horizon.

Let us consider the model in (29) and assume that the parameters are known and the errors are normally distributed. Then, drawing random numbers from the joint distribution of the errors for period $t+1, \ldots, t+n$ and solving the model forward, it is possible to get a set of simulated values for $(\mathrm{CCI}_{t+1}, \Delta y_{t+1}), \ldots, (\mathrm{CCI}_{t+n}, \Delta y_{t+n})$. Repeating the exercise many times, a histogram of the realizations provides an approximation for the conditional distribution of $(\mathrm{CCI}_{t+1}, \Delta y_{t+1}), \ldots, (\mathrm{CCI}_{t+n}, \Delta y_{t+n})$ given the past. Given this distribution and a rule to transform the continuous variable CCI into a binary recession indicator, e.g., the three months negative growth rule, the probability that a given future observation can be classified as a recession is computed as the fraction of the relevant simulated future values of the CCI that satisfy the rule.

A related problem that could be addressed within this framework is forecasting the beginning of the next recession, which is given by the time index of the first observation that falls into a recessionary pattern. Assuming that in period t the economy is in expansion, the probability of a recession after q periods, i.e., in $t+q$, is equal to the probability that $\mathrm{CCI}_{t+1}, \ldots, \mathrm{CCI}_{t+q-1}$ belong to an expansionary pattern while CCI_{t+q} to a recessionary one.

The procedure can be easily extended to allow for parameter uncertainty by drawing parameter values from the distribution of the estimators rather than treating them as fixed. Normality of the errors is also not strictly required since resampling can be used; see, e.g., Wecker (1979), Kling (1987), and Fair (1993) for additional details and examples.

Bayesian techniques are also available for forecasting turning points in linear models, see, e.g., Geweke and Whiteman (2006). In particular, Zellner and Hong (1991) and Zellner, Hong and Gulati (1990) addressed the problem in a decision-theoretic framework, using fixed parameter AR models with leading indicators as exogenous regressors. In our notation, the model can be written as

$$x_t = z_t' \beta + u_t, \quad u_t \sim \text{i.i.d. } N\big(0, \sigma^2\big), \tag{40}$$

where $z_t' = (x_{t-1}, y_{t-1})$, x_t is a univariate coincident variable or index, y_t is the $1 \times n$ vector of leading indicators, and β is a $k \times 1$ parameter vector, with $k = n + 1$.

Zellner, Hong and Gulati (1990) and Zellner, Hong and Min (1991) used annual data and declared a downturn (DT) in year $T+1$ if the annual growth rate observations satisfy

$$x_{T-2}, x_{T-1} < x_T > x_{T+1}, \tag{41}$$

while no downturn (NDT) happens if

$$x_{T-2}, x_{T-1} < x_T \leqslant x_{T+1}. \tag{42}$$

Similar definitions were proposed for upturns and no upturns.

The probability of a DT in $T+1$, p_{DT}, can be calculated as

$$p_{\mathrm{DT}} = \int_{-\infty}^{x_T} p(x_{T+1} \mid A_1, D_T)\, \mathrm{d}x_{T+1}, \tag{43}$$

where A_1 indicates the condition $(x_{T-2}, x_{T-1} < x_T)$, D_T denotes the past sample and prior information as of period T, and p is the predictive probability density function (pdf) defined as

$$p(x_{T+1} \mid D_T) = \int_\theta f(x_{T+1} \mid \theta, D_T)\pi(\theta \mid D_T)\,\mathrm{d}\theta, \tag{44}$$

where $f(x_{T+1} \mid \theta, D_T)$ is the pdf for x_{T+1} given the parameter vector $\theta = (\beta, \sigma^2)$ and D_T, while $\pi(\theta \mid D_T)$ is the posterior pdf for θ obtained by Bayes' Theorem.

The predictive pdf is constructed as follows. First, natural conjugate prior distributions are assumed for β and σ, namely, $p(\beta \mid \sigma) \sim N(0, \sigma^2 I \times 10^6)$ and $p(\sigma) \sim \mathrm{IG}(v_0 s_0)$, where IG stands for inverted gamma and v_0 and s_0 are very small numbers; see, e.g., Canova (2004, Chapter 9) for details. Second, at $t = 0$, the predictive pdf $p(x_1 \mid D_0)$ is a Student-t, namely, $t_{v_0} = (x_1 - z_1'\hat{\beta}_0)/s_0 a_0$ has a univariate Student-t density with v_0 degrees of freedom, where $a_0^2 = 1 + z_1' z_1 10^6$ and $\hat{\beta}_0 = 0$. Third, the posterior pdfs obtained period by period using the Bayes' Theorem are used to compute the period by period predictive pdfs. In particular, the predictive pdf for x_{T+1} is again Student-t and

$$t_{v_T} = \bigl(x_{T+1} - z_{T+1}'\hat{\beta}_T\bigr)/s_T a_T \tag{45}$$

has a univariate Student-t pdf with v_T degrees of freedom, where

$$\begin{aligned}
\hat{\beta}_T &= \hat{\beta}_{T-1} + \bigl(Z_{T-1}' Z_{T-1}\bigr)^{-1} z_T \bigl(x_T - z_T'\hat{\beta}_{T-1}\bigr)\big/\bigl[1 + z_T'(Z_T' Z_T)^{-1} z_T\bigr],\\
a_T^2 &= 1 + z_{T+1}'\bigl(Z_T' Z_T\bigr)^{-1} z_{T+1},\\
v_T &= v_{T-1} + 1,\\
v_T s_T^2 &= v_{T-1} s_{T-1} + \bigl(x_T - z_T'\hat{\beta}_T\bigr)^2 + \bigl(\hat{\beta}_T - \hat{\beta}_{T-1}\bigr)' Z_{T-1}' Z_{T-1}\bigl(\hat{\beta}_T - \hat{\beta}_{T-1}\bigr),
\end{aligned}$$

and $Z_T' = (z_T, z_{T-1}, \ldots, z_1)$. Therefore,

$$\Pr(x_{T+1} < x_T \mid D_T) = \Pr\bigl(t_{v_T} < \bigl(x_T - z_{T+1}'\hat{\beta}_T\bigr)\big/s_T a_T \bigm| D_T\bigr),$$

which can be analytically evaluated using the Student-t distribution with v_T degrees of freedom.

Finally, if the loss function is symmetric (i.e., the loss from wrongly predicting NDT in the case of DT is the same as predicting DT in the case of NDT), then a DT is predicted in period $T + 1$ if $p_{\mathrm{DT}} > 0.5$. Otherwise, the cut-off value depends on the loss structure, see also Section 8.3.

While the analysis in Zellner, Hong and Gulati (1990) is univariate, the theory for Bayesian VARs is also well developed, starting with Doan, Litterman and Sims (1984). A recent model in this class was developed by Zha (1998) for the Atlanta FED, and its performance in turning point forecasting is evaluated by Del Negro (2001). In this case the turning point probabilities are computed by simulations from the predictive pdf rather than analytically, in line with the procedure illustrated above in the classical context.

To conclude, a common problem of VAR models is their extensive parameterization, which prevents the analysis of large data sets. Canova and Ciccarelli (2001, 2003) proposed Bayesian techniques that partly overcome this problem, extending previous analysis by, e.g., Zellner, Hong and Min (1991), and providing applications to turning point forecasting; see Canova (2004, Chapter 10) for an overview. As an alternative, factor models can be employed, as we discuss in the next subsection.

6.2. Factor based CLI

The idea underlying Stock and Watson's (1989, SW) methodology for the construction of a CCI, namely that a single common force drives the evolution of several variables, can also be exploited to construct a CLI. In particular, if the single leading indicators are also driven by the (leads of the) same common force, then a linear combination of their present and past values can contain useful information for predicting the CCI.

To formalize the intuition above, following SW, Equation (6) in Section 5.1 is substituted with

$$\Delta C_t = \delta_C + \lambda_{CC}(L)\, \Delta C_{t-1} + \Lambda_{Cy}(L)\, \Delta y_{t-1} + v_{ct}. \tag{46}$$

and, to close the model, equations for the leading indicators are also added

$$\Delta y_t = \delta_y + \lambda_{yC}(L)\, \Delta C_{t-1} + \Lambda_{yy}(L)\, \Delta y_{t-1} + v_{yt}, \tag{47}$$

where v_{ct} and v_{yt} are i.i.d. and uncorrelated with the errors in (5).

The model in (4), (5), (46), (47) can be cast into state space form and estimated by maximum likelihood through the Kalman filter. SW adopted a simpler two-step procedure, where in the first step the model (4)–(6) is estimated, and in the second step the parameters of (46), (47) are obtained conditional on those in the first step. This procedure is robust to mis-specification of Equations (46), (47), in particular the estimated CCI coincides with that in Section 5.1, but it can be inefficient when either the whole model is correctly specified or, at least, the lags of the leading variables contain helpful information for estimating the current status of the economy. Notice also that the "forecasting" system (46), (47) is very similar to that in (29), the main difference being that here C_t is unobservable and therefore substituted with the estimate obtained in the first step of the procedure, which is $\mathrm{CCI}_{\mathrm{SW}}$. Another minor difference is that SW constrained the polynomials $\lambda_{yC}(L)$ and $\Lambda_{yy}(L)$ to eliminate higher order lags, while $\lambda_{CC}(L)$ and $\Lambda_{Cy}(L)$ are left unrestricted; see SW for the details on the lag length determination.

The SW composite leading index is constructed as

$$\mathrm{CLI}_{\mathrm{SW}} = \widehat{C}_{t+6|t} - C_{t|t}, \tag{48}$$

namely, it is a forecast of the 6-month growth rate in the $\mathrm{CCI}_{\mathrm{SW}}$, where the value in $t+6$ is forecasted and that in t is estimated. This is rather different from the NBER tradition, represented nowadays by the $\mathrm{CLI}_{\mathrm{CB}}$ that, as mentioned, aims at leading turning points in the level of the CCI. Following the discussion in Section 3, focusing on growth rather than on levels can be more interesting in periods of prolonged expansions.

A few additional comments are in order about SW's procedure. First, the leading indicators should depend on expected future values of the coincident index rather than on its lags, so that a better specification for (47) is along the lines of (36). Yet, we have seen that in the reduced form of (36) the leading indicators depend on their own lags and on those of the coincident variables, and a similar comment holds in this case. Second, the issue of parameter constancy is perhaps even more relevant in this enlarged model, and in particular for forecasting. Actually, in a subsequent (1997) revision of the procedure, SW made the deterministic component of (46), δ_C, time varying; in particular, it evolves according to a random walk. Third, dynamic estimation of Equation (46) would avoid the need of (47). This would be particularly convenient in this framework where the dimension of y_t is rather large, and a single forecast horizon is considered, $h = 6$. Fourth, rather than directly forecasting the $\mathrm{CCI}_{\mathrm{SW}}$, the components of x_t could be forecasted and then aggregated into the composite index using the in sample weights, along the lines of (31). Fifth, while SW formally tested for lack of cointegration among the components of x_t, they did not do it among the elements of y_t, and of (x_t, y_t), namely, there could be omitted cointegrating relationships either among the leading indicators, or among them and the coincident indicators. Finally, the hypothesis of a single factor driving both the coincident and the leading indicators should be formally tested.

Otrok and Whiteman (1998) derived a Bayesian version of SW's CCI and CLI. As in the classical context, the main complication is the nonobservability of the latent factor. To address this issue, a step-wise procedure is adopted where the posterior distribution of all unknown parameters of the model is determined conditional on the latent factor, then the conditional distribution of the latent factor conditional on the data and the other parameters is derived, the joint posterior distribution for the parameters and the factor is sampled using a Markov Chain Monte Carlo procedure using the conditional distributions in the first two steps, and a similar route is followed to obtain the marginal predictive pdf of the factor, which is used in the construction of the leading indicator; see Otrok and Whiteman (1998), Kim and Nelson (1998), Filardo and Gordon (1999) for details and Canova (2004, Chapter 11) for an overview.

The SW's methodology could also be extended to exploit recent developments in the dynamic factor model literature. In particular, a factor model for all the potential leading indicators could be considered, and the estimated factors used to forecast the coincident index or its components. Let us sketch the steps of this approach, more details can be found in Stock and Watson (2006).

The model for the leading indicators in (47) can be replaced by

$$\Delta y_t = \Lambda f_t + \xi_t, \tag{49}$$

where the dimension of Δy_t can be very large, possibly larger than the number of observations (so that no sequential indicator selection procedure is needed), f_t is an $r \times 1$ vector of common factors (so that more than one factor can drive the indicators), and ξ_t is a vector containing the idiosyncratic component of each leading indicator. Precise moment conditions on f_t and ξ_t, and requirements on the loadings matrix Λ, are given in Stock and Watson (2002a, 2002b). Notice that f_t could contain contemporaneous and lagged values of factors, so that the model is truly dynamic even though the representation in (49) is static.

Though the model in (49) is a simple extension of that for the construction of SW's composite coincident index in (4), its estimation is complicated by the possibly very large number of parameters, that makes maximum likelihood computationally not feasible. Therefore, Stock and Watson (2002a, 2002b) defined the factor estimators, $\hat{f}_t$, as the minimizers of the objective function

$$V_{nT}(f, \Lambda) = \frac{1}{nT} \sum_{i=1}^{n} \sum_{t=1}^{T} (y_{it} - \Lambda_i f_t)^2. \tag{50}$$

It turns out that the optimal estimators of the factors are the r eigenvectors corresponding to the r largest eigenvalues of the $T \times T$ matrix $n^{-1} \sum_{i=1}^{n} \underline{y}_i \underline{y}_i'$, where $\underline{y}_i = (y_{i1}, \ldots, y_{iT})$, and these estimators converge in probability to the space spanned by the true factors f_t. See Bai (2003) for additional inferential results, Bai and Ng (2002) for results related to the choice of the number of factors, r, Boivin and Ng (2003) for issues related to the choice of the size of the dataset (i.e., the number of leading indicators in our case), and Kapetanios and Marcellino (2003) for an alternative (parametric) estimation procedure.

The factors driving the leading indicators, possibly coinciding with (leads of) those driving the coincident indicators, can be related to the coincident composite index by replacing Equation (46) with

$$\Delta C_t = \delta_C + \lambda_{CC}(L)\,\Delta C_{t-1} + \lambda_{Cy}(L) f_{t-1} + v_{ct}. \tag{51}$$

Another important result proved by Stock and Watson (2002a, 2002b) is that the factors in the equation above can be substituted by their estimated counterparts, $\hat{f}_t$, without (asymptotically) modifying the mean square forecast error; see also Bai and Ng (2003) for additional results.

A forecasting procedure based on the use of (49) and (51), produced good results for the components of the CCI_{SW}, Stock and Watson (2002a, 2002b), but also for predicting macroeconomic variables for the Euro area, the UK, and the Accession countries, see, respectively, Marcellino, Stock and Watson (2003), Artis, Banerjee and Marcellino (2005), and Banerjee, Marcellino and Masten (2005). Yet, in these studies the set of indicators for factor extraction was not restricted to those with leading properties, and the target variable was not the composite coincident index. Camba-Mendez et al. (2001) used only leading indicators on the largest European countries for factor extraction (estimating iteratively the factor model cast in state-space form), and confirmed the good forecasting performance of the estimated factors when inserted in a VAR for predicting GDP growth.

The alternative factor based approach by FHLR described in Section 5.1 can also be used to construct a CLI. The leading variables are endogenously determined using the phase delay of their common components with respect to CCI_{FHLR} (the weighted average of the common components of interpolated monthly GDP for Euro area countries). An equal weight average of the resulting leading variables is the CLI_{FHLR}. Future values of the CCI_{FHLR} are predicted with a VAR for CCI_{FHLR}, CLI_{FHLR}. Further refinements of the methodology are presented in Forni et al. (2003a), with applications in Forni et al. (2003b).

All the factor based methods we have considered up to now focus on predicting continuous variables. Therefore, as in the case of linear models, we now discuss how to forecast discrete variables related to business cycle dynamics. In particular, we review the final important contribution of SW, further refined in Stock and Watson (1992), namely, the construction of a pattern recognition algorithm for the identification of recessions, and the related approach for computing recession probabilities.

As mentioned in Section 3, a recession is broadly defined by the three Ds: duration, a recession should be long enough; depth, there should be a substantial slowdown in economic activity; and diffusion, such a slowdown should be common to most sectors of the economy. Diffusion requires several series or a composite index to be monitored, and SW were in favor of the latter option, using their CCI (which, we recall, in the cumulated estimate of ΔC_t in Equation (4)). Moreover, SW required a recession to be characterized by ΔC_t falling below a certain boundary value, b_{rt} (depth), for either (a) six consecutive months or (b) nine months with no more than one increase during the middle seven months (duration), where (b) is the same as requiring ΔC_t to follow for seven of nine consecutive months including the first and the last month. Expansions were treated symmetrically, with b_{et} being the counterpart of b_{rt}, and both b_{rt} and b_{et} were treated as i.i.d. normal random variables.

A particular month is classified as a recession if it falls in a recessionary pattern as defined above. In particular, suppose that it has to be decided whether month t belongs to a recessionary pattern. Because of the definition of a recessionary pattern, the longest span of time to be considered is given by $\Delta C_{t-8}, \ldots, \Delta C_{t-1}$ and $\Delta C_{t+1}, \ldots, \Delta C_{t+8}$. For example, it could be that ΔC_t is below the threshold b_{rt} and also $\Delta C_{t-i} < b_{rt-i}$ for $i = 1, \ldots, 5$; in this case the sequence $\Delta C_{t-5}, \ldots, \Delta C_t$ is sufficient to classify period t as a recession. But it could be that $\Delta C_{t-i} > b_{rt-i}$ for $i = 1, \ldots, 8$, $\Delta C_t < b_{rt}$, $\Delta C_{t+1} > b_{rt+1}$, and $\Delta C_{t+i} < b_{rt+i}$ for $i = 2, \ldots, 8$, which requires to consider the whole sequence of 17 periods $\Delta C_{t-8}, \ldots, \Delta C_t, \ldots, \Delta C_{t+8}$ to correctly classify period t as a recession. Notice also that the sequence for ΔC_t has to be compared with the corresponding sequence of thresholds, $b_{rt-8}, \ldots, b_{rt}, \ldots, b_{rt+8}$.

The binary recession indicator, R_t, takes the value 1 if ΔC_t belongs to a recessionary pattern, and 0 otherwise. The expansion indicator is defined symmetrically, but is also worth noting that the definition of recession is such that there can be observations that are classified neither as recessions nor as expansions. Also, there is no role for duration dependence or correlation, in the sense that the probability of recession is independent of the length of the current expansion or recession, and of past values of R_t.

The evaluation of the probability of recession in period $t + h$ conditional on information on the present and past of the CCI and of the leading indicators (and on the fact that $t + h$ belongs either to an expansionary or to a recessionary pattern), requires the integration of a 34-dimensional distribution, where 17 dimensions are due to the evaluation of an (estimated and forecasted) sequence for ΔC_t that spans 17 periods, and the remaining ones from integration with respect to the distribution of the threshold parameters. Stock and Watson (1992) described in details a simulation based procedure to perform numerically the integration, and reported results for their composite recession indicator, $\mathrm{CRI}_{\mathrm{SW}}$, that evaluates in real time the probability that the economy will be in a recession 6-months ahead.

Though a rule that transforms the $\mathrm{CRI}_{\mathrm{SW}}$ into a binary variable is not defined, high values of the $\mathrm{CRI}_{\mathrm{SW}}$ should be associated with realizations of recessions. Using the NBER dating as a benchmark, SW found the in-sample performance of the CRI quite satisfactory, as well as that of the CLI. Yet, out of sample, in the recessions of 1990 and 2001, both indicators failed to provide strong early warnings, an issue that is considered in more detail in Section 10.3.

To conclude, it is worth pointing out that the procedure underlying SW's CRI is not specific to their model. Given the definition of a recessionary pattern, any model that relates a CCI to a set of leading indicators or to a CLI can be used to compute the probability of recession in a given future period using the same simulation procedure as SW but drawing the random variables from the different model under analysis. The simplest case is when the model for the coincident indicator and the leading indexes is linear, which is the situation described at the end of the previous subsection.

6.3. *Markov switching based CLI*

The MS model introduced in Section 5.2 to define an intrinsic coincident index, and in Section 3 to date the business cycle, can also be exploited to evaluate the forecasting properties of a single or composite leading indicator. In particular, a simplified version of the model proposed by Hamilton and Perez-Quiros (1996) can be written as

$$\begin{aligned}
\Delta x_t - c_{s_t} &= a(\Delta x_{t-1} - c_{s_{t-1}}) + b(\Delta y_{t-1} - d_{s_{t+r-1}}) + u_{xt},\\
\Delta y_t - d_{s_{t+r}} &= c(\Delta x_{t-1} - c_{s_{t-1}}) + d(\Delta y_{t-1} - d_{s_{t+r-1}}) + u_{yt},\\
u_t &= (u_{xt}, u_{yt})' \sim \text{i.i.d. } N(0, \Sigma),
\end{aligned} \tag{52}$$

where x and y are univariate, s_t evolves according to the constant transition probability Markov chain defined in (11), and the leading characteristics of y are represented not only by its influence on future values of x but also by its being driven by future values of the state variable, s_{t+r}.

The main difference between (52) and the MS model used in Section 5.2, Equation (9), is the presence of lags and leads of the state variable. This requires to define a new state variable, s_t^*, such that

$$s_t^* = \begin{cases}
1 & \text{if } s_{t+r} = 1,\ s_{t+r-1} = 1,\ \ldots,\ s_{t-1} = 1,\\
2 & \text{if } s_{t+r} = 0,\ s_{t+r-1} = 1,\ \ldots,\ s_{t-1} = 1,\\
3 & \text{if } s_{t+r} = 1,\ s_{t+r-1} = 0,\ \ldots,\ s_{t-1} = 1,\\
\vdots & \qquad\qquad\vdots\\
2^{r+2} & \text{if } s_{t+r} = 0,\ s_{t+r-1} = 0,\ \ldots,\ s_{t-1} = 0.
\end{cases} \tag{53}$$

The transition probabilities of the Markov chain driving s_t^* can be derived from (11), and in the simplest case where $r = 1$ they are summarized by the matrix

$$P = \begin{pmatrix} p_{11} & 0 & 0 & 0 & p_{11} & 0 & 0 & 0 \\ p_{10} & 0 & 0 & 0 & p_{10} & 0 & 0 & 0 \\ 0 & p_{01} & 0 & 0 & 0 & p_{01} & 0 & 0 \\ 0 & p_{00} & 0 & 0 & 0 & p_{00} & 0 & 0 \\ 0 & 0 & p_{11} & 0 & 0 & 0 & p_{11} & 0 \\ 0 & 0 & p_{10} & 0 & 0 & 0 & p_{10} & 0 \\ 0 & 0 & 0 & p_{01} & 0 & 0 & 0 & p_{01} \\ 0 & 0 & 0 & p_{00} & 0 & 0 & 0 & p_{00} \end{pmatrix}, \tag{54}$$

whose ith, jth element corresponds to the probability that $s_t^* = i$ given that $s_{t-1}^* = j$.

The quantity of major interest is the probability that s_t^* assumes a certain value given the available information, namely,

$$\zeta_{t|t} = \begin{pmatrix} \Pr(s_t^* = 1 \mid x_t, x_{t-1}, \ldots, x_1,\ y_t, y_{t-1}, \ldots, y_1) \\ \Pr(s_t^* = 2 \mid x_t, x_{t-1}, \ldots, x_1,\ y_t, y_{t-1}, \ldots, y_1) \\ \vdots \\ \Pr(s_t^* = 2^{r+2} \mid x_t, x_{t-1}, \ldots, x_1,\ y_t, y_{t-1}, \ldots, y_1) \end{pmatrix}, \tag{55}$$

which is the counterpart of Equation (12) in this more general context. The vector $\zeta_{t|t}$ and the conditional density of future values of the variables given the past, $f(x_{t+1}, y_{t+1} \mid s_{t+1}^*, x_t, \ldots, x_1,\ y_t, \ldots, y_1)$, can be computed using the sequential procedure outlined in Section 5.2; see Hamilton and Perez-Quiros (1996) and Krolzig (2004) for details. The latter can be used for forecasting future values of the coincident variable, the former to evaluate the current status of the economy or to forecast its future status up to period $t + r$. For example, the probability of being in a recession today is given by the sum of the rows of $\zeta_{t|t}$ corresponding to those values of s_t^* characterized by $s_t = 1$, while the probability of being in a recession in period $t + r$ is given by the sum of the rows of $\zeta_{t|t}$ corresponding to those values of s_t^* characterized by $s_{t+r} = 1$. To make inference on states beyond period $t + r$, it is possible to use the formula

$$\zeta_{t+m|t} = P^m \zeta_{t|t}, \tag{56}$$

which is a direct extension of the first row of (14).

Hamilton and Perez-Quiros (1996) found that their model provides only a weak signal of recession in 1960, 1970 and 1990. Moreover, the evidence in favor of the nonlinear cyclical factor is weak and the forecasting gains for predicting GNP growth or its turning point are minor with respect to a linear VAR specification. Even weaker evidence in favor of the MS specification was found when a cointegrating relationship between GNP and lagged CLI is included in the model. The unsatisfactory performance of the MS model could be due to the hypothesis of constant probability of recessions, as in the univariate context; see, e.g., Filardo (1994). Evidence supporting this claim, based on the recession of 1990, is provided by Filardo and Gordon (1999).

Chauvet (1998) found a good performance also for the factor MS model in tracking the recession of 1990 using the proper version of $\zeta_{t|t}$ in that context. This is basically the only forecasting application of the factor MS models described in Section 2.1, so that further research is needed to close the gap. For example, SW's procedure for the CLI construction could be implemented using Kim and Nelson's (1998) MS version of the factor model, or a switching element could be introduced in the SW's VAR Equations (46) and (47).

The MS model can also be used to derive analytic forecasts of recession (or expansion) duration. Suppose that x_t follows the simpler MS model in (9)–(11) and that it is known that in period t the economy is in a recession, i.e., $s_t = 1$. Then,

$$\begin{aligned}
&\Pr(s_{t+1} = 1 \mid x_t, \ldots, x_1) = p_{11}, \\
&\Pr(s_{t+2} = 1,\ s_{t+1} = 1 \mid x_t, \ldots, x_1) \\
&\quad = \Pr(s_{t+2} = 1 \mid s_{t+1} = 1,\ x_t, \ldots, x_1) \Pr(s_{t+1} = 1 \mid x_t, \ldots, x_1) = p_{11}^2, \\
&\quad \vdots
\end{aligned} \tag{57}$$

and the probability that the recession ends in period $t + n$ is

$$\Pr(s_{t+n} = 0,\ s_{t+n-1} = 1,\ \ldots,\ s_{t+1} = 1 \mid x_t, \ldots, x_1) = (1 - p_{11}) p_{11}^{n-1}. \tag{58}$$

Instead, if (11) is substituted with (18), i.e., the state probabilities are time-varying, then

$$\begin{aligned}
&\Pr(s_{t+n} = 0,\ s_{t+n-1} = 1,\ \ldots,\ s_{t+1} = 1 \mid x_t, \ldots, x_1) \\
&\quad = (1 - \hat{p}_{11,t+n}) \prod_{j=1}^{n-1} \hat{p}_{11,t+j}
\end{aligned} \tag{59}$$

with

$$\hat{p}_{11,t+j} = E\left(\frac{\exp(\theta y_{t+j-1})}{1 + \exp(\theta y_{t+j-1})} \;\middle|\; x_t, \ldots, x_1,\ y_t, \ldots, y_1 \right). \tag{60}$$

It follows that an estimator of the expected remaining duration of the recession, τ, in period t is given by

$$\hat{\tau} = E(\tau \mid s_t = 1) = \sum_{i=1}^{\infty} i(1 - \hat{p}_{11,t+i}) \prod_{j=1}^{i-1} \hat{p}_{11,t+j}, \tag{61}$$

which simplifies to

$$\hat{\tau} = E(\tau \mid s_t = 1) = \sum_{i=1}^{\infty} i(1 - p_{11}) p_{11}^{i-1}, \tag{62}$$

for constant probabilities. An interesting issue is therefore whether the leading indicators are useful to predict τ or not.

To conclude, Bayesian methods for the estimation of Markov switching models were developed by Albert and Chib (1993a), McCulloch and Tsay (1994), Filardo and Gordon (1994) and several other authors; see, e.g., Filardo and Gordon (1999) for a comparison of Bayesian linear, MS and factor models for coincident indicators, and Canova (2004, Chapter 11) for an overview. Yet, to the best of our knowledge, there are no applications to forecasting turning points with Bayesian MS models while, for example, a Bayesian replication of the Hamilton and Perez-Quiros (1996) exercise would be feasible and interesting.

7. Examples of composite coincident and leading indexes

In this section we provide empirical examples to illustrate some of the theoretical methods introduced so far. In particular, in the first subsection we compare several composite coincident indexes obtained with different methodologies, while in the second subsection we focus on leading indexes.

7.1. Alternative CCIs for the US

In Figure 1 we graph four composite coincident indexes for the US over the period 1959:1–2003:12: the Conference Board's equal weighted nonmodel based CCI, the

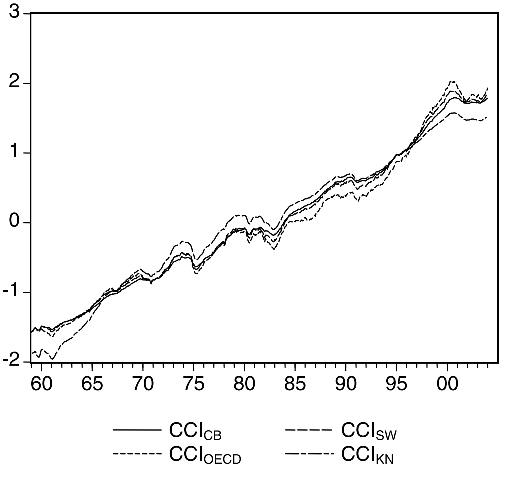

Figure 1. Composite coincident indexes. The figure reports the Conference Board's composite coincident indicator (CCI_{CB}), the OECD reference coincident series (CCI_{OECD}), Stock and Watson's coincident index (CCI_{SW}), and the coincident index derived from the four components in CCI_{CB} modeled with a dynamic factor model as in Kim and Nelson (1998) (CCI_{KN}). All indexes have been normalized to have zero mean and unit standard deviation.

Table 1
Correlation of composite coincident indexes (6-month percentage change)

	CCI_{CB}	CCI_{OECD}	CCI_{SW}	CCI_{KN}
CCI_{CB}	1			
CCI_{OECD}	0.941	1		
CCI_{SW}	0.979	0.969	1	
CCI_{KN}	0.943	0.916	0.947	1

Note: Common sample is 1970:01–2003:11.

OECD coincident reference series which is a transformation of IP, the Stock and Watson's (1989) factor model based CCI, and the Kim and Nelson's (1998) Bayesian MS factor model based CCI computed using the four coincident series combined in the CCI_{CB}. For the sake of comparability, all indexes are normalized to have zero mean and unit standard deviation.

The figure highlights the very similar behavior of all the CCIs, which in particular share the same pattern of peaks and troughs. The visual impression is confirmed by the correlations for the levels, and by those for the 6-month percentage changes reported in Table 1, the lowest value being 0.916 for CCI_{KN} and CCI_{OECD}. These values are in line with previous studies, see Section 5, and indicate that it is possible to achieve a close to complete agreement on the status of the economy.

In Figure 2 we consider dating the US classical and deviation cycles. In the upper panel we graph the CCI_{CB} and the NBER expansion/recession classification. The figure highlights that the NBER recessions virtually coincide with the peak-trough periods in the CCI_{CB}. In the middle panel we graph the CCI_{CB} and the expansion/recession classification resulting from the AMP dating. The results are virtually identical with respect to the NBER (see also the first two columns of Table 3), with the noticeable difference that AMP identifies a double dip at the beginning of the new century with recessions in 2000:10–2001:12 and 2002:7–2003:4 versus 2001:3–2001:11 for the NBER. In the lower panel of Figure 2 we graph the HP band pass filtered CCI_{CB}, described in Section 3, and the AMP dating for the resulting deviation cycle. As discussed in Section 3, the classical cycle recessions are a subset of those for the deviation cycle, since the latter capture periods of lower growth even if not associated with declines in the level of the CCI.

Finally, in Figure 3 we report the (filtered) probability of recessions computed with two methods. In the upper panel we graph the probabilities resulting from the Kim and Nelson's (1998) Bayesian MS factor model applied to the four coincident series combined in the CCI_{CB}. In the lower panel those from the AMP nonparametric MS approach applied to the CCI_{CB}. The results in the two panels are very similar, and the matching of peaks in these probabilities and NBER dated recessions is striking. The latter result supports the use of these methods for real-time dating of the business cycle. It is also worth noting that both methods attribute a probability close to 60% for a second

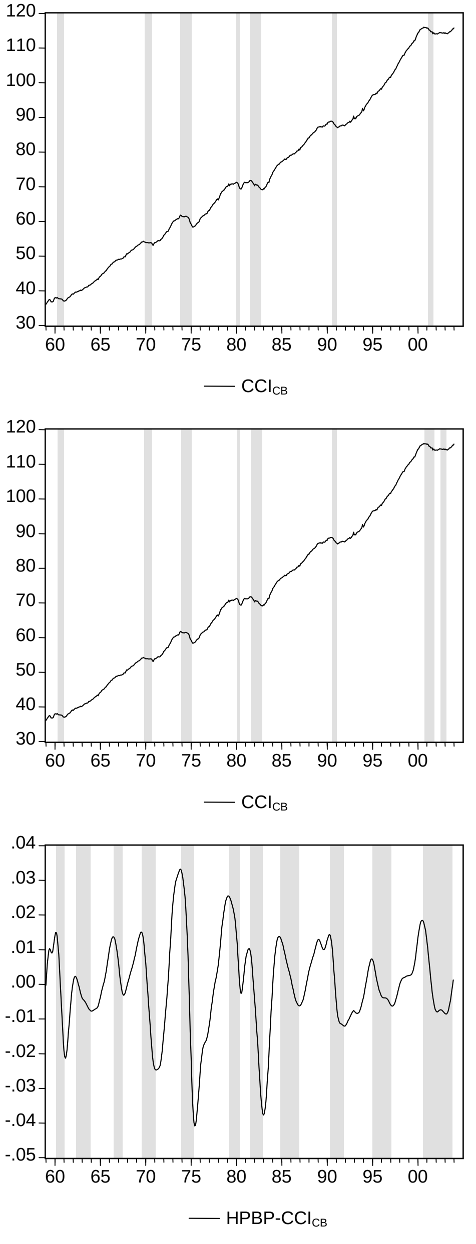

Figure 2. Classical and deviation cycles. Upper panel: $\mathrm{CCI}_{\mathrm{CB}}$ and NBER dated recessions (shaded areas). Middle panel: $\mathrm{CCI}_{\mathrm{CB}}$ and recessions dated with Artis, Marcellino and Proietti (2004) algorithm (shaded areas). Lower panel: HP-band pass filtered $\mathrm{CCI}_{\mathrm{CB}}$ and recessions dated with Artis, Marcellino and Proietti (2004) algorithm (shaded areas).

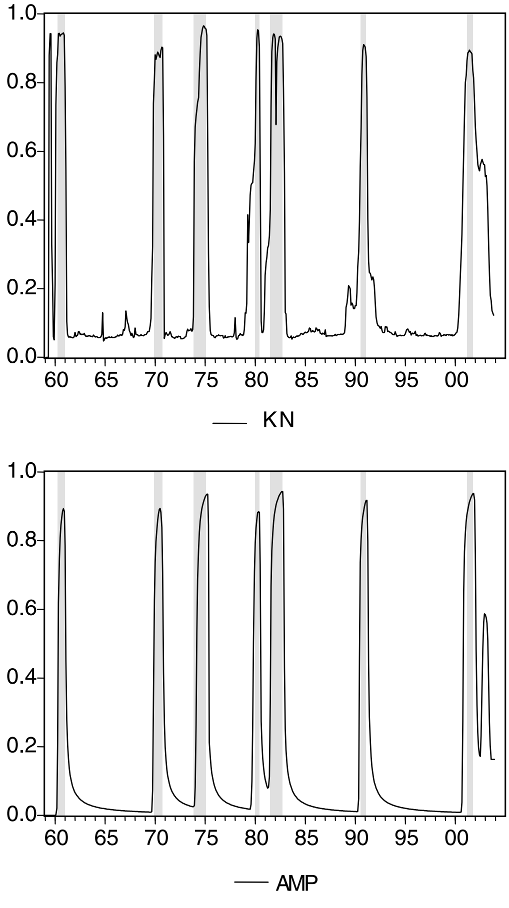

Figure 3. Probability of recession and NBER dated recessions. The upper panel reports the (filtered) probability of recession computed from a dynamic factor model for the four components in the CCI_{CB} using the Kim and Nelson's (1998) methodology. The lower panel reports the (filtered) probability of recession computed using the algorithm in Artis, Marcellino and Proietti (2004) applied to the CCI_{CB}. The shaded areas are the NBER dated recessions.

short recession at the beginning of the century, in line with the AMP dating reported in the middle panel of Figure 2 but in contrast with the NBER dating.

7.2. *Alternative CLIs for the US*

We start this subsection with an analysis of the indicator selection process for Stock and Watson's (1989, SW) model based composite leading index, described in detail in

Section 6.2, and of the construction of two nonmodel based indexes for the US produced by official agencies, the Conference Board, $\mathrm{CLI_{CB}}$, and the OECD, $\mathrm{CLI_{OECD}}$.

SW started with a rather large dataset of about 280 series, yet smaller than Mitchell and Burns' original selection of 487 candidate indicators. The series can be divided into ten groups: "measures of output and capacity utilization; consumption and sales; inventories and orders; money and credit quantity variables; interest rates and asset prices; exchange rates and foreign trade; employment, earnings and measures of the labor force; wages and prices; measures of government fiscal activity; and other variables", SW (p. 365).

The bivariate relationships between each indicator, properly transformed, and the growth of the $\mathrm{CCI_{DOC}}$ were evaluated using frequency domain techniques (the coherence and the phase lead), and time domain techniques (Granger causality tests and marginal predictive content for $\mathrm{CCI_{DOC}}$ beyond that of $\mathrm{CLI_{DOC}}$). The choice of $\mathrm{CCI_{DOC}}$ rather than $\mathrm{CCI_{SW}}$ as the target variable can raise some doubts, but the latter was likely not developed yet at the time, and in addition the two composite coincident indexes are highly correlated. Some series were retained even if they performed poorly on the basis of the three criteria listed above, because either economic theory strongly supported their inclusion or they were part of the $\mathrm{CLI_{DOC}}$. After this first screening, 55 variables remained in the list of candidate components of the composite leading index.

It is interesting that SW mentioned the possibility of using all the 55 series for the construction of an index, but abandoned the project for technical reasons (at the time construction of a time series model for all these variables was quite complicated) and because it would be difficult to evaluate the contribution of each component to the index. About ten years later, the methodology to address the former issue became available, see Stock and Watson (2002a, 2002b) and the discussion in Section 6.2 above, but the latter issue remains, the trade-off between parsimony and broad coverage of the index is still unresolved.

The second indicator selection phase is based on a step-wise regression procedure. The dependent variable is $\mathrm{CCI}_{\mathrm{SW}t+6} - \mathrm{CCI}_{\mathrm{SW}t}$, i.e., the six months growth rate in the SW composite coincident index, that is also the target variable for SW composite leading index, see Section 6.2. Different sets of variables (including their lags as selected by the AIC) are used as regressors, variables in each set are retained on the basis of their marginal explanatory power, the best variables in each original set are grouped into other sets of regressors, and the procedure is repeated until a small number of indicators remains in the list.

At the end, seven variables (and their lags) were included in the composite index, as listed in Table 1 in SW. They are:

(i) an index of new private housing authorized,
(ii) the growth rate of manufacturers' unfilled orders for durable goods industries,
(iii) the growth rate in a trade weighted nominal exchange rate,
(iv) the growth rate of part-time work in nonagricultural industries,
(v) the difference of the yield on constant-maturity portfolio of 10-years US treasury bonds,

(vi) the spread between interest rates on 6-months corporate paper and 6-months US treasury bills,
(vii) the spread between the yield on 10-years and 1-year US Treasury bonds.

The only change in the list so far took place in 1997, when the maturity in (vi) became 3 months. SW also discussed theoretical explanations for the inclusion of these variables (and exclusion of others). The most innovative variables in SW's CLI_{SW} are the financial spreads, whose forecasting ability became the focus of theoretical and empirical research in subsequent years. Yet, following an analysis of the performance of their CLI_{SW} during the 1990 recession, see Section 10.3, Stock and Watson (1992) also introduced a nonfinancial based index (CLI2_{SW}).

A potential problem of the extensive variable search underlying the final selection of index components, combined with parameter estimation, is overfitting. Yet, when SW checked the overall performance of their selection procedure using Monte Carlo simulations, the results were satisfactory. Even better results were obtained by Hendry and Krolzig (1999, 2001) for their automated model selection procedure, PcGets; see Banerjee and Marcellino (2005) for an application to leading indicator selection for the US.

A final point worth noting about SW's indicator selection procedure is the use of variable transformations. First, seasonally adjusted series are used. Second, a stationarity transformation is applied for the indicator to have similar properties as the target. Third, some series are smoothed because of high frequency noise, in particular, (ii)–(v) in the list above. The adopted filter is $f(L) = 1+2L+2L^2+L^3$. Such a filter is chosen with reference to the target variable, the 6-month growth of CCI, and to the use of first differenced indicators, since $f(L)(1-L)$ is a band-pass filter with gains concentrated at periods of four months to one year. Finally, if the most recent values of some of the seven indicators are not available, they are substituted with forecasts in order to be able to use as timely information as possible. Zarnowitz and Braun (1990), in their comment to SW, pointed out that smoothing the indicators contributes substantially to the good forecasting performance of SW's CLI, combined with the use of the most up-to-date information.

The practice of using forecasts when timely data are not available is now supported also for the CLI_{CB} [see McGuckin, Ozyildirim and Zarnowitz (2003)], but not yet implemented in the published version of the index. The latter is computed following the same steps as for the coincident index, the CCI_{CB} described in Section 4, but with a different choice of components. In particular, the single indicators combined in the index include average weekly hours, manufacturing; average weekly initial claims for unemployment insurance; manufacturers' new orders, consumer good and materials (in 1996$); vendor performance, slower deliveries diffusion index; manufacturers' new orders, nondefense capital goods; building permits, new private housing units; stock prices, 500 common stocks; money supply (in 1996$); interest rate spread, 10-year Treasury bond less federal funds; and the University of Michigan's index of consumer expectations.

This list originates from the original selection of Mitchell and Burns (1938), but only two variables passed the test of time: average weekly hours in the manufacturing sector and the Standard and Poor's stock index (that replaces the Dow Jones index of industrial common stock prices); see Moore (1983) for an historical perspective. Both variables are not included in the CLI_{SW}, since their marginal contribution in forecasting the 6-month growth of the CCI_{SW} is not statistically significant. Other major differences in the components of the two composite leading indexes are the inclusion in CLI_{CB} of M2 and of the index of consumer expectations (the relationship of M2 with the CCI_{SW} is found to be unstable, while consumer expectations were added to CLI_{CB} in the '90s so that the sample is too short for a significant evaluation of their role); and the exclusion from CLI_{CB} of an exchange rate measure and of the growth in part time work (yet, the former has a small weight in the CLI_{SW}, while the latter is well proxied by the average weekly hours in manufacturing and the new claims for unemployment insurance).

The third CLI for the US we consider is the OECD composite short leading index, CLI_{OECD} (see www.oecd.org). Several points are worth making. First, the target is represented by the turning points in the growth cycle of industrial production, where the trend component is estimated using a modified version of the phase average trend (PAT) method developed at the NBER [see OECD (**?**), Niemira and Klein (1994) for details], and the Bry and Boschan (1971) methodology is adopted for dating peaks and troughs. All of these choices are rather questionable, since industrial production is a lower and lower share of GDP (though still one of the most volatile components), theoretically sounder filters such as those discussed in Section 3 are available for detrending, and more sophisticated procedures are available for dating, see again Section 3. On the other hand, since the OECD computes the leading index for a wide variety of countries, simplicity and robustness are also relevant for them.

Second, the criteria for the selection of the components of the index are broadly in line with those listed in Section 2. The seven chosen indicators as listed in the OECD web site include dwellings started; net new orders for durable goods, share price index; consumer sentiment indicator; weekly hours of work, manufacturing; purchasing managers index; and the spread of interest rates. Overall, there is a strong similarity with the elements of the CLI_{CB}.

Third, as for CLI_{CB}, the components are first standardized and then aggregated with equal weights. More precisely, each indicator is detrended with the PAT method; smoothed according to its months for cyclical dominance (MCD) values to reduce irregularity [see OECD (**?**) for details]; transformed to homogenize the cyclical amplitudes; standardized by subtracting the mean from the observed values and then dividing the resulting difference by the mean of the absolute values of the differences from the mean; and finally aggregated. When timely data for an indicator are not available, the indicator is not included in the preliminary release of the composite leading index.

Finally, the composite index is adjusted to ensure that its cyclical amplitude on average agrees with that of the detrended reference series. The trend restored version of the index is also computed and published, to get comparability with the IP series.

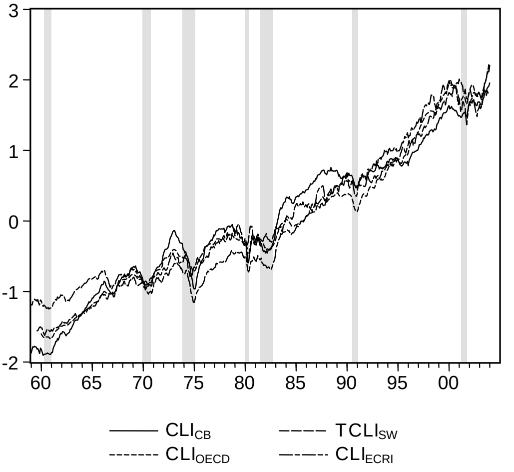

Figure 4. Composite leading indexes. The figure reports the Conference Board composite leading index (CLI_{CB}), the OECD leading index (CLI_{OECD}), a transformation of Stock and Watson's leading index ($TCLI_{SW}$, see text), the ECRI leading index (CLI_{ECRI}), and the NBER dated recessions (shaded areas). All indexes have been normalized to have zero mean and unit standard deviation.

A fourth CLI commonly monitored for the US is the Economic Cycle Research Institute's weekly leading index (see www.businesscycle.com). The precise parameters and procedural details underlying the construction of the CLI_{ECRI} are proprietary, the methodology is broadly described in Boschan and Banerji (1990).

In Figure 4 we graph the four composite leading indexes for the US we have described: the Conference Board's leading index (CLI_{CB}), the OECD leading index (CLI_{OECD}), the ECRI's weekly leading index (CLI_{ECRI}), and a transformation of Stock and Watson's (1989) composite leading index ($TCLI_{SW}$), their leading index plus their coincident index that yields a 6-month ahead forecast for the level of the coincident index, see Section 6.2. For comparability, all indexes are normalized to have zero mean and unit standard deviation. In the same figure we graph the NBER dated recessions (shaded areas).

Visual inspection suggests that the four indices move closely together, and their peaks anticipate NBER recessions. These issues are more formally evaluated in Tables 2 and 3. In Table 2 we report the correlations of the 6-month percentage changes of the four indices, which are indeed high, in particular when the '60s are excluded from the sample, the lowest value being 0.595 for CLI_{SW} and CLI_{ECRI}.

In Table 3 we present a descriptive analysis of the peak and trough structure of the four leading indexes (obtained with the AMP algorithm), compared either with the NBER dating or with the dating of the CCI_{CB} resulting from the AMP algorithm. The $TCLI_{SW}$ has the worst performance in terms of missed peaks and troughs, but it is worth recalling that the goal of the CLI_{SW} is not predicting turning points but the 6-month growth rate of the CCI_{SW}. The other three leading indexes missed no peaks or troughs,

Table 2
Correlation of composite leading indexes (6-month percentage change)

	CLI_{CB}	CLI_{OECD}	CLI_{SW}	CLI_{ECRI}
CLI_{CB}	1			
CLI_{OECD}	0.891	1		
CLI_{SW}	0.719	0.601	1	
CLI_{ECRI}	0.817	0.791	0.595	1

Note: Common sample is 1970:01–2003:11.

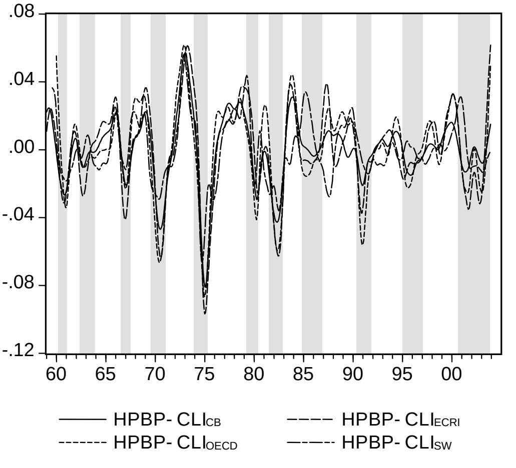

Figure 5. Filtered composite leading indexes with AMP dated recessions for deviation cycle of CCI_{CB}. The figure reports the HP-band pass filtered versions of the four CLIs in Figure 4, and the Artis, Marcellino and Proietti (2004) dating of the HP band pass filtered versions of the CCI_{CB} (shaded areas).

with the exception of the 2002 peak identified only by the AMP dating algorithm. Yet, they gave three false alarms, in 1966, 1984–1985, and 1994–1995. The average lead for recessions is about 9–10 months for all indexes (slightly shorter for $TCLI_{SW}$), but for expansions it drops to only 3–4 months for CLI_{OECD} and CLI_{ECRI}. Based on this descriptive analysis, the CLI_{CB} appears to yield the best overall leading performance. Yet, these results should be interpreted with care since they are obtained with the final release of the leading indicators rather than with real time data, see Section 10.1.

In Figure 5 we graph the HP band pass filtered versions of the four composite leading indexes, with the AMP deviation cycle dating (shaded areas). Again the series move closely together, slightly less so for the HPBP-$TCLI_{SW}$, and their peaks anticipate dated recessions.

Table 3
Classical cycles, dating of coincident and leading indexes

Peak						Trough					
Coincident		Leading (AMP)				Coincident		Leading (AMP)			
NBER	AMP	CB	OECD	ECRI	SW	NBER	AMP	CB	OECD	ECRI	SW
Apr 1960	May 1960	Jan 1959*	Jan 1960*	Jan 1959	Aug 1959*	Feb 1961	Feb 1961	Mar 1960	Dec 1960	Oct 1960	May 1960
				Jan 1962						Jun 1962	
		Apr 1966	Apr 1966	Apr 1966	Feb 1966			Dec 1966	Nov 1966	Dec 1966	Jul 1966
Dec 1969	Nov 1969	May 1969	Jan 1969	Jan 1969	**MISSING**	Nov 1970	Nov 1970	Apr 1970	Apr 1970	Jul 1970	**MISSING**
Nov 1973	Dec 1973	Feb 1973	Feb 1973	Jun 1973	Jan 1973	Mar 1975	Mar 1975	Jan 1975	Dec 1974	Jan 1975	Aug 1974
Jan 1980	Feb 1980	Nov 1978	Aug 1978	Nov 1978	Jun 1979	Jul 1980	Jul 1980	Apr 1980	Apr 1980	May 1980	Aug 1981
Jul 1981	Aug 1981	Nov 1980	Nov 1980	May 1981	**MISSING**	Nov 1982	Dec 1982	Jan 1982	Feb 1982	Aug 1982	**MISSING**
			Feb 1984		Oct 1985				Sep 1984		Jun 1986
		Jul 1988						Jun 1989			
Jul 1990	Jul 1990	Feb 1990	Mar 1990	Oct 1989	Feb 1990	Mar 1991	Mar 1991	Jan 1991	Dec 1990	Dec 1990	Jan 1991
		Nov 1994	Dec 1994					May 1995	Apr 1995		
				May 1998						Oct 1998	
Mar 2001	Oct 2000	Feb 2000	Feb 2000	Feb 2000	**MISSING**	Nov 2001	Dec 2001	Mar 2001	Oct 2001	Oct 2001	**MISSING**
	Jul 2002	**MISSING**	May 2002	**MISSING**	Feb 2002		Apr 2003	**MISSING**	**MISSING**	Apr 2003	**MISSING**

	NBER	AMP	NBER	AMP	NBER	AMP	NBER	AMP	NBER	AMP	NBER	AMP	NBER	AMP	NBER	AMP
Average lead	10	11	9	9	9	10	7	8	9	9	4	4	3	3	8	9
St. dev.	4.23	4.28	4.30	5.31	5.13	4.75	3.78	2.50	4.30	5.31	2.89	3.04	1.11	1	5.38	5.80
False alarms	3	3	3	3	3	3	2	2	3	3	3	3	3	3	2	2
Missing	0	1	0	0	0	1	2	4	0	1	0	0	0	1	3	4

Note: Shaded values are false alarms, 'MISSING' indicates a missed turning point. Leads longer than 18 months are considered false alarms. Negative leads are considered missed turning points. AMP: dating based on algorithm in Artis, Marcellino and Proietti (2004).
* indicates no previous available observation. Based on final release of data.

Table 4
Correlations of HP band pass filtered composite leading indexes

	$\text{HPBP-CLI}_{\text{CB}}$	$\text{HPBP-CLI}_{\text{OECD}}$	$\text{HPBP-CLI}_{\text{ECRI}}$	$\text{HPBP-CLI}_{\text{SW}}$
$\text{HPBP-CLI}_{\text{CB}}$	1			
$\text{HPBP-CLI}_{\text{OECD}}$	0.919	1		
$\text{HPBP-CLI}_{\text{ECRI}}$	0.906	0.882	1	
$\text{HPBP-CLI}_{\text{SW}}$	0.703	0.595	0.645	1

Note: Common sample is 1970:01–2003:11.

From Table 4, the $\text{HPBP-TCLI}_{\text{SW}}$ is the least correlated with the other indexes, correlation coefficients are in the range 0.60–0.70, while for the other three indexes the lowest correlation is 0.882.

From Table 5, the ranking of the indexes in terms of lead-time for peaks and troughs is similar to that in Table 3. In this case there is no official dating of the deviation cycle, so that we use the AMP algorithm applied to the $\text{HPBP-CCI}_{\text{CB}}$ as a reference. The $\text{HPBP-CLI}_{\text{CB}}$ confirms its good performance, with an average lead time of 7 months for recessions, 10 months for expansions, and just one missed signal and two false alarms. The $\text{HPBP-CLI}_{\text{ECRI}}$ is a close second, while the $\text{HPBP-TCLI}_{\text{SW}}$ remains the worst, with 3–4 missed signals.

Finally, the overall good performance of the simple nonmodel based CLI_{CB} deserves further attention. We mentioned that it is obtained by cumulating, using the formula in (3), an equal weighted average of the one month symmetric percent changes of ten indicators. The weighted average happens to have a correlation of 0.960 with the first principal component of the ten members of the CLI_{CB}. The latter provides a nonparametric estimator for the factor in a dynamic factor model, see Section 6.2 and Stock and Watson (2002a, 2002b) for details. Therefore, the CLI_{CB} can also be considered as a good proxy for a factor model based composite leading indicator.

8. Other approaches for prediction with leading indicators

In this section we discuss other methods to transform leading indicators into a forecast for the target variable. In particular, Section 8.1 deals with observed transition models, Section 8.2 with neural network and nonparametric methods, Section 8.3 with binary models, and Section 8.4 with forecast pooling procedures. Examples are provided in the next section, after having defined formal evaluation criteria for leading indicator based forecasts.

8.1. Observed transition models

In the class of MS models described in Sections 5.2 and 6.3, the transition across states is abrupt and driven by an unobservable variable. As an alternative, in smooth transition (ST) models the parameters evolve over time at a certain speed, depending on the

Table 5
Deviations cycles, dating of coincident and leading indexes

Peak					Trough				
Coincident	Leading				Coincident	Leading			
CB	CB	OECD	ECRI	SW	CB	CB	OECD	ECRI	SW
Mar 1960	May 1959	Feb 1960	Jul 1959	Sep 1959	Mar 1961	Nov 1960	Jan 1961	Oct 1960	Jan 1961
May 1962	Jan 1962	Jan 1962	Dec 1961	**MISSING**	Jan 1964	Sep 1962	Nov 1962	Sep 1962	**MISSING**
				Apr 1963					May 1964
Jul 1967	Feb 1966	Mar 1966	Feb 1966	Jan 1967	Aug 1967	Feb 1967	Jan 1967	Dec 1966	Jan 1967
Aug 1969	Feb 1969	Dec 1968	Feb 1969	Dec 1967	Mar 1971	Jul 1970	Jun 1970	Aug 1970	Jun 1970
Dec 1973	Feb 1973	Jan 1973	May 1973	Jan 1973	Jun 1975	Feb 1975	Jan 1975	Jan 1975	Oct 1974
Mar 1979	Sep 1978	Sep 1978	Dec 1978	May 1979	Jul 1980	May 1982	Apr 1980	Jun 1982	Feb 1980
Jul 1981	**MISSING**	Mar 1981	**MISSING**	Sep 1980	Jan 1983	**MISSING**	May 1982	**MISSING**	Jun 1982
Nov 1984	Jan 1984	Dec 1983	Oct 1983	Apr 1985	Jan 1987	Jan 1986	May 1985	Oct 1985	Aug 1987
			Jun 1987					Apr 1988	
May 1990	Sep 1987	Aug 1987	Nov 1989	Jan 1990	Dec 1991	Dec 1990	Jan 1991	Nov 1990	Jul 1991
		Feb 1993					Jul 1993		
Jan 1995	Jun 1994	Jun 1994	Oct 1993	Jan 1994	Mar 1997	Nov 1995	Aug 1995	Feb 1995	Oct 1994
				Aug 1995					May 1997
		Nov 1997					Oct 1998		
Aug 2000	Jan 2000	Mar 2000	Mar 2000	Jan 2001	Dec 2003*	May 2001	Dec 2003*	Dec 2003*	Nov 2003*
	May 2002					Dec 2003*			
Aver. lead	7	6	7	8		10	7	10	6
St. dev.	2.28	3.21	3.80	3.25		4.67	4.03	4.47	2.31
False alarms	2	2	1	2		1	4	2	1
Missing	1	0	1	4		1	0	1	3

Note: Shaded values are false alarms, 'MISSING' indicates a missed turning point. Leads longer than 18 months are considered false alarms. Negative leads are considered missed turning points. AMP: dating based on algorithm in Artis, Marcellino and Proietti (2004).

* indicates last available observation. Based on final release of data.

behavior of observable variables. In particular, the ST-VAR, that generalizes the linear model in (21) can be written as

$$\begin{aligned}
\Delta x_t &= c_x + A\,\Delta x_{t-1} + B\,\Delta y_{t-1} + (c_x + A\,\Delta x_{t-1} + B\Delta y_{t-1})F_x + u_{xt},\\
\Delta y_t &= c_y + C\Delta x_{t-1} + D\Delta y_{t-1} + (c_y + C\Delta x_{t-1} + D\Delta y_{t-1})F_y + u_{yt},\\
u_t &= (u_{xt}, u_{yt})' \sim \text{i.i.d. } N(0, \Sigma),
\end{aligned} \tag{63}$$

where

$$F_x = \frac{\exp(\theta_0 + \theta_1 z_{t-1})}{1 + \exp(\theta_0 + \theta_1 z_{t-1})}, \qquad F_y = \frac{\exp(\phi_0 + \phi_1 z_{t-1})}{1 + \exp(\phi_0 + \phi_1 z_{t-1})}, \tag{64}$$

and z_{t-1} contains lags of x_t and y_t.

The smoothing parameters θ_1 and ϕ_1 regulate the shape of parameter change over time. When they are equal to zero, the model becomes linear, while for large values the model tends to a self-exciting threshold model [see, e.g., Potter (1995), Artis, Galvao and Marcellino (2003)], whose parameters change abruptly as in the MS case. In this sense the ST-VAR provides a flexible tool for modelling parameter change.

The transition function F_x is related to the probability of recession. In particular, when the values of z_{t-1} are much smaller than the threshold value, θ_0, the value of F_x gets close to zero, while large values lead to values of F_x close to one. This is a convenient feature in particular when F_x only depends on lags of y_t, since it provides direct evidence on the usefulness of the leading indicators to predict recessions. As an alternative, simulation methods as in Section 6.1 can be used to compute the probabilities of recession.

Details on the estimation and testing procedures for ST models, and extensions to deal with more than two regimes or time-varying parameters, are reviewed, e.g., by van Dijk, Teräsvirta and Franses (2002), while Teräsvirta (2006) focuses on the use of ST models in forecasting. In particular, as it is common with nonlinear models, forecasting more than one-step ahead requires the use of simulation techniques, unless dynamic estimation is used as, e.g., in Stock and Watson (1999b) or Marcellino (2003).

Univariate versions of the ST model using leading indicators as transition variables were analyzed by Granger, Teräsvirta and Anderson (1993), while Camacho (2004), Anderson and Vahid (2001), and Camacho and Perez-Quiros (2002) considered the VAR case. The latter authors found a significant change in the parameters only for the constant, in line with the MS specifications described in the previous subsection and with the time-varying constant introduced by SW to compute their CLI.

Finally, Bayesian techniques for the analysis of smooth transition models were developed by Lubrano (1995), and by Geweke and Terui (1993) and Chen and Lee (1995) for threshold models; see Canova (2004, Chapter 11) for an overview. Yet, there are no applications to forecasting using leading indicators.

8.2. Neural networks and nonparametric methods

The evidence reported so far, and that summarized in Section 10 below, is not sufficient to pin down the best parametric model to relate the leading to the coincident indicator, different sample periods or indicators can produce substantially different results. A possible remedy is to use artificial neural networks, which can provide a valid approximation to the generating mechanism of a vast class of nonlinear processes; see, e.g., Hornick, Stinchcombe and White (1989), and Swanson and White (1997), Stock and Watson (1999b), Marcellino (2003) for their use as forecasting devices.

In particular, Stock and Watson (1999b) considered two types of univariate neural network specifications. The single layer model with n_1 hidden units (and a linear component) is

$$x_t = \beta_0' z_t + \sum_{i=1}^{n_1} \gamma_{1i} g\left(\beta_{1i}' z_t\right) + e_t, \tag{65}$$

where $g(z)$ is the logistic function, i.e., $g(z) = 1/(1 + e^{-z})$, and z_t includes lags of the dependent variable. Notice that when $n_1 = 1$ the model reduces to a linear specification with a logistic smooth transition in the constant. A more complex model is the double layer feedforward neural network with n_1 and n_2 hidden units:

$$x_t = \beta_0' z_t + \sum_{j=1}^{n_2} \gamma_{2j} g\left(\sum_{i=1}^{n_1} \beta_{2ji} g\left(\beta_{1i}' z_t\right)\right) + e_t. \tag{66}$$

The parameters of (65) and (66) can be estimated by nonlinear least squares, and forecasts obtained by dynamic estimation.

While the studies using NN mentioned so far considered point forecasts, Qi (2001) focused on turning point prediction. The model she adopted is a simplified version of (66), namely,

$$r_t = g\left(\sum_{i=1}^{n_1} \beta_{2i} g\left(\beta_{1i}' z_t\right)\right) + e_t, \tag{67}$$

where z_t includes lagged leading indicators in order to evaluate their forecasting role, and r_t is a binary recession indicator. Actually, since $g(\cdot)$ is the logistic function, the predicted values from (67) are constrained to lie in the [0, 1] interval. As for (65) and (66), the model is estimated by nonlinear least squares, and dynamic estimation is adopted when forecasting.

An alternative way to tackle the uncertainty about the functional form of the relationship between leading and coincident indicators is to adopt a nonparametric specification, with the cost for the additional flexibility being the required simplicity of the model. Based on the results from the parametric models they evaluated, Camacho and Perez-Quiros (2002) suggested the specification,

$$x_t = m(y_{t-1}) + e_t, \tag{68}$$

estimated by means of the Nadaraya–Watson estimator; see also Härdle and Vieu (1992). Therefore,

$$\hat{x}_t = \left(\sum_{j=1}^{T} K\left(\frac{y_{t-1} - y_j}{h}\right) x_j\right) \Bigg/ \left(\sum_{j=1}^{T} K\left(\frac{y_{t-1} - y_j}{h}\right)\right), \tag{69}$$

where $K(\cdot)$ is the Gaussian kernel and the bandwidth h is selected by leave-one-out cross validation.

The model is used to predict recessions according to the two negative quarters rule. For example,

$$\Pr(x_{t+2} < 0,\ x_{t+1} < 0 \mid y_t) = \int_{y_{t+2}<0} \int_{y_{t+1}<0} f(x_{t+2}, x_{t+1} \mid y_t)\, \mathrm{d}x_{t+2}\, \mathrm{d}x_{t+1}, \tag{70}$$

and the densities are estimated using an adaptive kernel estimator; see Camacho and Perez-Quiros (2002) for details.

Another approach that imposes minimal structure on the leading-coincident indicator connection is the pattern recognition algorithm proposed by Keilis-Borok et al. (2000). The underlying idea is to monitor a set of leading indicators, comparing their values to a set of thresholds, and when a large fraction of the indicators rise above the threshold a recession alarm, A_t, is sent. Formally, the model is

$$A_t = \begin{cases} 1 & \text{if } \sum_{k=1}^{N} \Psi_{kt} \geqslant N - b, \\ 0 & \text{otherwise,} \end{cases} \tag{71}$$

where $\Psi_{kt} = 1$ if $y_{kt} \geqslant c_k$, and $\Psi_{kt} = 0$ otherwise. The salient features of this approach are the tight parameterization (only $N+1$ parameters, $b, c_1, \ldots, c_N$), which is in general a plus in forecasting, the transformation of the indicators into binary variables prior to their combination (from y_{kt} to Ψ_{kt} and then summed with equal weights), and the focus on the direct prediction of recessions, A_t is a $0/1$ variable.

Keilis-Borok et al. (2000) used 6 indicators: SW's CCI defined in Section 5.1 and five leading indicators, the interest rate spread, a short term interest rate, manufacturing and trade inventories, weekly initial claims for unemployment, and the index of help wanted advertising. They analyzed three different versions of the model in (71) where the parameters are either judgementally assigned or estimated by nonlinear least squares, with or without linear filtering of the indicators, finding that all versions perform comparably and satisfactory, producing (in a pseudo-out-of-sample context) an early warning of the five recessions over the period 1961 to 1990. Yet, the result should be interpreted with care because of the use of the finally released data and of the selection of the indicators using full sample information, consider, e.g., the use of the spread which was not common until the end of the '80s.

8.3. Binary models

In the models we have analyzed so far to relate coincident and leading indicators, the dependent variable is continuous, even though forecasts of business cycle turning points are feasible either directly (MS or ST models) or by means of simulation methods (linear or factor models). A simpler and more direct approach treats the business cycle phases as a binary variable, and models it using a logit or probit specification.

In particular, let us assume that the economy is in recession in period t, $R_t = 1$, if the unobservable variable s_t is larger than zero, where the evolution of s_t is governed by

$$s_t = \beta' y_{t-1} + e_t. \tag{72}$$

Therefore,

$$\Pr(R_t = 1) = \Pr(s_t > 0) = F\big(\beta' y_{t-1}\big), \tag{73}$$

where $F(\cdot)$ is either the cumulative normal distribution function (probit model), or the logistic function (logit model). The model can be estimated by maximum likelihood, and the estimated parameters combined with current values of the leading indicators to provide an estimate of the recession probability in period $t+1$, i.e.,

$$\widehat{R}_{t+1} = \Pr(R_{t+1} = 1) = F\big(\hat{\beta}' y_t\big). \tag{74}$$

The logit model was adopted, e.g., by Stock and Watson (1991) and the probit model by Estrella and Mishkin (1998), while Birchenhall et al. (1999) provided a statistical justification for the former in a Bayesian context [on the latter, see also Zellner and Rossi (1984) and Albert and Chib (1993b)]. Binary models for European countries were investigated by Estrella and Mishkin (1997), Bernard and Gerlach (1998), Estrella, Rodrigues and Schich (2003), Birchenhall, Osborn and Sensier (2001), Osborn, Sensier and Simpson (2001), Moneta (2003).

Several points are worth discussing about the practical use of the probit or logit models for turning point prediction. First, often in practice the dating of R_t follows the NBER expansion/recession classification. Since there are substantial delays in the NBER's announcements, it is not known in period t whether the economy is in recession or not. Several solutions are available to overcome this problem. Either the model is estimated with data up to period $t-k$ and it is assumed that β remains constant in the remaining part of the sample; or R_t is substituted with an estimated value from an auxiliary binary model for the current status of the economy, e.g., using the coincident indicators as regressors [see, e.g., Birchenhall et al. (1999)] or one of the alternative methods for real-time dating of the cycle described in Section 2.2 is adopted.

Second, as in the case of dynamic estimation, a different model specification is required for each forecast horizon. For example, if an h-step ahead prediction is of interest, the model in (72) should be substituted with

$$s_t = \gamma_h' y_{t-h} + u_{t,h}. \tag{75}$$

This approach typically introduces serial correlation and heteroskedasticity into the error term $u_{t,h}$, so that the logit specification combined with nonlinear least-squares estimation and robust estimation of the standard errors of the parameters can be preferred over standard maximum likelihood estimation, compare, for example, (67) in the previous subsection which can be considered as a generalization of (75). Notice also that $\hat{\gamma}_h' y_{t-h}$ can be interpreted as an h-step ahead composite leading indicator. As an alternative, the model in (72) could be complemented with an auxiliary specification for y_t, say,

$$y_t = A y_{t-1} + v_t \tag{76}$$

so that

$$\begin{aligned}\Pr(R_{t+h} = 1) = \Pr(s_{t+h} > 0) &= \Pr\bigl(\beta' A^{h-1} y_t + \eta_{t+h-1} + e_{t+h} > 0\bigr)\\ &= F_{\eta+e}\bigl(\beta' A^{h-1} y_t\bigr)\end{aligned} \tag{77}$$

with $\eta_{t+h-1} = \beta' v_{t+h-1} + \beta' A v_{t+h-2} + \cdots + \beta' A^{h-1} v_t$. In general, the derivation of $F_{\eta+e}(\cdot)$ is quite complicated, and the specification of the auxiliary model for y_t can introduce additional noise. Dueker (2003) extended and combined Equations (72) and (76) into

$$\begin{pmatrix} s_t \\ y_t \end{pmatrix} = \begin{pmatrix} a & B \\ c & D \end{pmatrix} \begin{pmatrix} s_{t-1} \\ y_{t-1} \end{pmatrix} + \begin{pmatrix} e_{st} \\ e_{yt} \end{pmatrix}, \tag{78}$$

which is referred to as Qual-VAR because of its similarity with the models considered in Section 6.1. The model composed of the equation for s_t alone is the dynamic ordered probit studied by Eichengreen, Watson and Grossman (1985), who derived its likelihood and the related maximum likelihood estimators. Adding the set of equations for y_t has the main advantage of closing the model for forecasting purposes. Moreover, Dueker (2003) showed that the model can be rather easily estimated using Gibbs sampling techniques, and Dueker and Wesche (2001) found sizeable forecasting gains with respect to the standard probit model, in particular during recessionary periods.

Third, the construction of the probability of a recession within a certain period, say $t+2$, is complicated within the binary model framework. The required probability is given by $\Pr(R_{t+1} = 0,\ R_{t+2} = 1) + \Pr(R_{t+1} = 1,\ R_{t+2} = 0) + \Pr(R_{t+1} = 1,\ R_{t+2} = 1)$. Then, either from (75),

$$\begin{aligned}\Pr(R_{t+1} = 1,\ R_{t+2} = 1) &= \Pr(s_{t+1} > 0,\ s_{t+2} > 0)\\ &= \Pr\bigl(u_{t+1,1} > -\gamma_1' y_t,\ u_{t+2,2} > -\gamma_2' y_t\bigr),\end{aligned} \tag{79}$$

or from (77),

$$\begin{aligned}&\Pr(R_{t+1} = 1,\ R_{t+2} = 1) = \Pr(s_{t+1} > 0,\ s_{t+2} > 0)\\ &\quad = \Pr\bigl(\beta' y_t + e_{t+1} > 0,\ \beta' A y_t + \beta' v_{t+1} + e_{t+2} > 0\bigr),\end{aligned} \tag{80}$$

and similar formulae apply for $\Pr(R_{t+1}=0,\ R_{t+2}=1)$ and $\Pr(R_{t+1}=1,\ R_{t+2}=0)$. As long as the joint distributions in (79) and (80) are equivalent to the product of the marginal ones, as in this case assuming that v_t are uncorrelated with e_t, and the error terms are i.i.d., an analytic solution can be found. For higher values of h simulation methods are required. For example, a system made up of the models resulting using Equation (75) for different values of h can be jointly estimated and used to simulate the probability values in (79). A similar approach can be used to compute the probability that an expansion (or a recession) will have a certain duration. A third, simpler alternative, is to define another binary variable directly linked to the event of interest, in this case,

$$R2_t = \begin{cases} 0 & \text{if no recession in period } t+1,\ t+2, \\ 1 & \text{if at least one recession in } t+1,\ t+2, \end{cases} \tag{81}$$

and then model $R2_t$ with a probit or logit specification as a function of indicators dated up to period $t-1$. The problem of this approach is that it is not consistent with the model for R_t in Equations (72), (73). The extent of the mis-specification should be evaluated in practice and weighted with the substantial simplification in the computations. A final, more promising, approach is simulation of the Qual-VAR model in (78), along the lines of the linear model in Section 6.1.

Fourth, an additional issue that deserves investigation is the stability of the parameters over time, and in particular across business cycle phases. Chin, Geweke and Miller (2000) proposed to estimate different parameters in expansions and recessions, using an exogenous classification of the states based on their definition of turning points. Dueker (1997, 2002) suggested to make the switching endogenous by making the parameters of (72) evolve according to a Markov chain. Both authors provided substantial evidence in favor of parameters instability.

Fifth, an alternative procedure to compute the probability of recession in period t consists of estimating logit or probit models for a set of coincident indicators, and then aggregating the resulting forecasts. The weights can be either those used to aggregate the indicators into a composite index, or they can be determined within a pooling context, as described in the next subsection.

Sixth, Pagan (2005) points out that the construction of the binary R_t indicator matters, since it can imply that the indicator is not i.i.d. as required by the standard probit or logit analysis.

Finally, as in the case of MS or ST models, the estimated probability of recession, $\hat{r}_{t+1}$, should be transformed into a 0/1 variable using a proper rule. The common choices are of the type $\hat{r}_t \geqslant c$ where c is either 0.5, a kind of uninformative Bayesian prior, or equal to the sample unconditional recession probability. Dueker (2002) suggested to make the cutoff values also regime dependent, say c_0 and c_1, and to compare the estimated probability with a weighted combination of c_0 and c_1 using the related regime probabilities. In general, as suggested, e.g., by Zellner, Hong and Gulati (1990) and analyzed in details by Lieli (2004), the cutoff should be a function of the preferences of the forecasters.

8.4. Pooling

Since the pioneering work of Bates and Granger (1969), it is well known that pooling several forecasts can yield a mean square forecast error (msfe) lower than that of each of the individual forecasts; see Timmermann (2006) for a comprehensive overview. Hence, rather than selecting a preferred forecasting model, it can be convenient to combine all the available forecasts, or at least some subsets.

Several pooling procedures are available. The three most common methods in practice are linear combination, with weights related to the msfe of each forecast [see, e.g., Granger and Ramanathan (1984)], median forecast selection, and predictive least squares, where a single model is chosen, but the selection is recursively updated at each forecasting round on the basis of the past forecasting performance.

Stock and Watson (1999b) and Marcellino (2004) presented a detailed study of the relative performance of these pooling methods, using a large dataset of, respectively, US and Euro area macroeconomic variables, and taking as basic forecasts those produced by a range of linear and nonlinear models. In general simple averaging with equal weights produces good results, more so for the US than for the Euro area. Stock and Watson (2003a) focused on the role of pooling for GDP growth forecasts in the G7 countries, using a larger variety of pooling methods, and dozens of models. They concluded that median and trimmed mean pooled forecasts produce a more stable forecasting performance than each of their component forecasts. Incidentally, they also found pooled forecasts to perform better than the factor based forecasts discussed in Section 6.2.

Camacho and Perez-Quiros (2002) focused on pooling leading indicator models, in particular they considered linear models, MS and ST models, probit specifications, and the nonparametric model described in Section 8.2, using regression based weights as suggested by Granger and Ramanathan (1984). Hence, the pooled forecast is obtained as

$$\hat{x}_{t+1|t} = w_1\hat{x}_{t+1|t,1} + w_2\hat{x}_{t+1|t,2} + \cdots + w_p\hat{x}_{t+1|t,p}, \tag{82}$$

and the weights, w_i, are obtained as the estimated coefficients from the linear regression

$$x_t = \omega_1\hat{x}_{t|t-1,1} + \omega_2\hat{x}_{t|t-1,2} + \cdots + \omega_p\hat{x}_{t|t-1,p} + u_t \tag{83}$$

which is estimated over a training sample using the forecasts from the single models to be pooled, $\hat{x}_{t|t-1,i}$, and the actual values of the target variable.

Camacho and Perez-Quiros (2002) evaluated the role of pooling not only for GDP growth forecasts but also for turning point prediction. The pooled recession probability is obtained as

$$\hat{r}_{t+1|t} = F(a_1\hat{r}_{t+1|t,1} + a_2\hat{r}_{t+1|t,2} + \cdots + a_p\hat{r}_{t+1|t,p}), \tag{84}$$

where $F(\cdot)$ is the cumulative distribution function of a normal variable, and the weights, a_i, are obtained as the estimated parameters in the probit regression

$$r_t = F(\alpha_1\hat{r}_{t|t-1,1} + \alpha_2\hat{r}_{t|t-1,2} + \cdots + \alpha_p\hat{r}_{t|t-1,p}) + e_t, \tag{85}$$

which is again estimated over a training sample using the recession probabilities from the single models to be pooled, $\hat{r}_{t|t-1,i}$, and the actual values of the recession indicator, r_t.

The pooling method described above was studied from a theoretical point of view by Li and Dorfman (1996) in a Bayesian context. A more standard Bayesian approach to forecast combination is the use of the posterior odds of each model as weights; see, e.g., Zellner and Min (1993). When all models have equal prior odds, this is equivalent to the use of the likelihood function value of each model as its weight in the pooled forecast.

9. Evaluation of leading indicators

In this section we deal with the evaluation of the forecasting performance of the leading indicators when used either in combination with simple rules to predict turning points, or as regressors in one of the models described in the previous sections to forecast either the growth rate of the target variable or its turning points. In the first subsection we consider methodological aspects while in the second subsection we discuss empirical examples.

9.1. Methodology

A first assessment of the goodness of leading indicators can be based on standard in-sample specification and mis-specification tests of the models that relate the indicators to the target variable.

The linear model in (21) provides the simplest framework to illustrate the issues. A first concern is whether it is a proper statistical model of the relationships among the coincident and the leading variables. This requires the estimated residuals to mimic the assumed i.i.d. characteristics of the errors, the parameters to be stable over time, and the absence of nonlinearity. Provided these hypotheses are not rejected, the model can be used to assess additional properties, such as Granger causality of the leading for the coincident indicators, or to evaluate the overall goodness of fit of the equations for the coincident variables (or for the composite coincident index). The model also offers a simple nesting framework to evaluate the relative merits of competing leading indicators, whose significance can be assessed by means of standard testing procedures. For a comprehensive analysis of the linear model see, e.g., Hendry (1995).

The three steps considered for the linear model, namely, evaluation of the goodness of the model from a statistical point of view, testing of hypotheses of interest on the parameters, and comparison with alternative specifications should be performed for each of the approaches listed in Sections 6 and 8. In particular, Hamilton and Raj (2002) and Raj (2002) provide up-to-date results for Markov-switching models, van Dijk, Teräsvirta and Franses (2002) for smooth transition models, while, e.g., Marcellino and Mizon (2004) present a general framework for model comparison.

Yet, in-sample analyses are more useful to highlight problems of a certain indicator or methodology than to provide empirical support in their favor, since they can be biased by over-fitting and related problems due to the use of the same data for model specification, estimation, and evaluation. A more sound appraisal of the leading indicators can be based on their out of sample performance, an additional reason for this being that forecasting is their main goal.

When the target is a continuous variable, such as the growth of a CCI over a certain period, standard forecast evaluation techniques can be used. In particular, the out-of-sample mean square forecast error (MSFE) or mean absolute error (MAE) provide standard summary measures of forecasting performance. Tests for equal forecast accuracy can be computed along the lines of Diebold and Mariano (1995), Clark and McCracken (2001), the standard errors around the MSFE of a model relative to a benchmark can be computed following West (1996), and tests for forecast encompassing can be constructed as in Clark and McCracken (2001). West (2006) provides an up-to-date survey of forecast evaluation techniques.

Moreover, as discussed in Section 6, simulation methods are often employed to compute the joint distribution of future values of the CCI to produce recession forecasts. Such a joint distribution can be evaluated using techniques developed in the density forecast literature; see, e.g., Corradi and Swanson (2006).

When the target variable, R_t, is a binary indicator while the (out of sample) forecast is a probability of recession, P_t, similar techniques can be used since the forecast error is a continuous time variable. For example, Diebold and Rudebusch (1989) defined the accuracy of the forecast as

$$\mathrm{QPS} = \frac{1}{T}\sum_{t=1}^{T} 2(P_t - R_t)^2, \tag{86}$$

where QPS stands for quadratic probability score, which is the counterpart of the MSFE. The range of QPS is $[0, 2]$, with 0 for perfect accuracy. A similar loss function that assigns more weight to larger forecast errors is the log probability score,

$$\mathrm{LPS} = -\frac{1}{T}\sum_{t=1}^{T}\bigl((1 - R_t)\log(1 - P_t) + R_t \log P_t\bigr). \tag{87}$$

The range of LPS is $[0, \infty]$, with 0 for perfect accuracy.

Furthermore, Stock and Watson (1992) regressed $R_{t+k} - \mathrm{CRI}_{t+k|t}$, i.e., the difference of their indicator of recession and the composite recession index, on available information in period t, namely

$$R_{t+k} - \mathrm{CRI}_{t+k|t} = z_t\beta + e_t, \tag{88}$$

where the regressors in z_t are indicators included or excluded in SW's CLI. The error term in the above regression is heteroskedastic, because of the discrete nature of R_t, and serially correlated, because of the k-period ahead forecast horizon. Yet, robust t- and

F-statistics can be used to test the hypothesis of interest, $\beta = 0$, that is associated with correct model specification when z_t contains indicators included in the CLI, or with an efficient use of the information in the construction of the recession forecast when z_t contains indicators excluded from the CLI. Of course, the model in (88) can also be adopted when the dependent variable is a growth rate forecast error.

If the CRI or any probability of recession are transformed into a binary indicator, S_t, by choosing a threshold such that if the probability of recession increases beyond it then the indicator is assigned a value of one, the estimation method for the regression in (88) should be changed, since the dependent variable becomes discrete. In this case, a logistic or probit regression with appropriate corrections for the standard errors of the estimated coefficients would suit.

Contingency tables can also be used for a descriptive evaluation of the methodology in the case of binary forecasts and outcomes. They provide a summary of the percentage of correct predictions, missed signals (no prediction of slowdown when it takes place), and false alarms (prediction of slowdown when it does not take place). A more formal assessment can be based on a concordance index, defined as

$$I_{RS} = \frac{1}{T}\sum_{t=1}^{T}\bigl[R_t S_t + (1 - S_t)(1 - R_t)\bigr], \tag{89}$$

with values in the interval [0, 1], and 1 for perfect concordance. Under the assumption that S_t and R_t are independent, the estimate of the expected value of the concordance index is $2\overline{SR} = 1 - \overline{R} - \overline{S}$, where $\overline{R}$ and $\overline{S}$ are the averages of R_t and S_t. Subtracting this quantity from I_{RS} yields the mean-corrected concordance index [Harding and Pagan (2002, 2005)]:

$$I_{RS}^{*} = 2\frac{1}{T}\sum_{t=1}^{T}\bigl(S_t - \overline{S}\bigr)\bigl(R_t - \overline{R}\bigr). \tag{90}$$

AMP showed that under the null hypothesis of independence of S_t and R_t,

$$T^{1/2} I_{RS}^{*} \to N\bigl(0, 4\sigma^2\bigr), \quad \sigma^2 = \gamma_R(0)\gamma_S(0) + 2\sum_{\tau=1}^{\infty}\gamma_R(\tau)\gamma_S(\tau), \tag{91}$$

where $\gamma_S(\tau) = E[(S_t - E(S_t))(S_{t-\tau} - E(S_t))]$ and $\gamma_S(\tau)$ is defined accordingly. A consistent estimator of σ^2 is

$$\hat{\sigma}^2 = \hat{\gamma}_R(0)\hat{\gamma}_S(0) + 2\sum_{\tau=1}^{l}\left(1 - \frac{\tau}{T}\right)\hat{\gamma}_R(\tau)\hat{\gamma}_S(\tau), \tag{92}$$

where l is the truncation parameter and $\hat{\gamma}_R(\tau)$ and $\hat{\gamma}_S(\tau)$ are the sample counterparts of $\gamma_R(\tau)$ and $\gamma_S(\tau)$. As an alternative, Harding and Pagan (2002, 2005) proposed to regress R_t on S_t, and use a robust t-test to evaluate the significance of S_t.

Notice that since the predictive performance of the leading indicators can vary over expansions and recessions, and/or near turning points, it can be worth providing a separate evaluation of the models and the indicators over these subperiods, using any of the methods mentioned so far. The comparison should also be conducted at different forecast horizons, since the ability to provide early warnings is another important property for a leading indicator, though difficult to be formally assessed in a statistical framework.

A final comment concerns the choice of the loss function, that in all the forecast evaluation criteria considered so far is symmetric. Yet, when forecasting growth or a recession indicator typically the losses are greater in case of a missed signal than for a false alarm, for example, because policy-makers or firms cannot take timely counteracting measures. Moreover, false alarms can be due to the implementation of timely and effective policies as a reaction to the information in the leading indicators, or can signal major slowdowns that do not turn into recessions but can be of practical policy relevance. These considerations suggest that an asymmetric loss function could be a more proper choice, and in such a case using the methods summarized so far to evaluate a leading indicator based forecast or rank competing forecasts can be misleading. For example, a model can produce a higher loss than another model even if the former has a lower MSFE or MAE, the best forecast can be biased, or an indicator can be significant in (88) without reducing the loss; see, e.g., Artis and Marcellino (2001), Elliott, Komunjer and Timmermann (2003), Patton and Timmermann (2003), and Granger and Machina (2006) for an overview. More generally, the construction itself of the leading indicators and their inclusion in forecasting models should be driven by the loss function and, in case, take its asymmetry into proper account.

9.2. Examples

We now illustrate the methodology for model evaluation discussed in the previous subsection, using four empirical examples that involve some of the models reviewed in Sections 6 and 8.

The first application focuses on the use of linear models for the (one-month symmetric percent changes of the) $\mathrm{CCI}_{\mathrm{CB}}$ and the $\mathrm{CLI}_{\mathrm{CB}}$. We focus on the following six specifications. A bivariate VAR for the $\mathrm{CCI}_{\mathrm{CB}}$ and the $\mathrm{CLI}_{\mathrm{CB}}$, as in Equation (34). A univariate AR for the $\mathrm{CCI}_{\mathrm{CB}}$. A bivariate ECM for the $\mathrm{CCI}_{\mathrm{CB}}$ and the $\mathrm{CLI}_{\mathrm{CB}}$, as in Equation (39), where one cointegrating vector is imposed and its coefficient recursively estimated. A VAR for the four components of the $\mathrm{CCI}_{\mathrm{CB}}$ and the $\mathrm{CLI}_{\mathrm{CB}}$, as in Equation (29). A VAR for the $\mathrm{CCI}_{\mathrm{CB}}$ and the ten components of the $\mathrm{CLI}_{\mathrm{CB}}$. Finally, a VAR for the four components of the $\mathrm{CCI}_{\mathrm{CB}}$ and the ten components of the $\mathrm{CLI}_{\mathrm{CB}}$, as in Equation (21). Notice that most of these models are nonnested, except for the AR which is nested in some of the VARs, and for the bivariate VAR which is nested in the ECM.

The models are compared on the basis of their forecasting performance one and six month ahead over the period 1989:1–2003:12, which includes the two recessions of July 1990–March 1991 and March 2001–November 2001. The forecasts are computed recursively with the first estimation sample being 1959:1–1988:12 for one step ahead forecasts and 1959:1–1988:6 for six step ahead forecasts, using the final release of the indexes and their components. While the latter choice can bias the evaluation towards the usefulness of the leading indicators, this is not a major problem when the forecasting comparison excludes the ’70s and ’80s and when, as in our case, the interest focuses on the comparison of alternative models for the same vintage of data, see the next section for details. The lag length is chosen by BIC over the full sample. Recursive BIC selects smaller models for the initial samples, but their forecasting performance is slightly worse. The forecasts are computed using both the standard iterated method, and dynamic estimation (as described in Equation (25)).

The comparison is based on the MSE and MAE relative to the bivariate VAR for the $\mathrm{CCI}_{\mathrm{CB}}$ and the $\mathrm{CLI}_{\mathrm{CB}}$. The Diebold and Mariano (1995) test for the statistical significance of the loss differentials is also computed. The results are reported in the upper panel of Table 6.

Five comments can be made. First, the simple AR model performs very well, there are some very minor gains from the VAR only six step ahead. This finding indicates that the lagged behavior of the $\mathrm{CCI}_{\mathrm{CB}}$ contains useful information that should be included in a leading index. Second, taking cointegration into account does not improve the forecasting performance. Third, forecasting the four components of the $\mathrm{CCI}_{\mathrm{CB}}$ and then aggregating the forecasts, as in Equation (31), decreases the MSE at both horizons, and the difference with respect to the bivariate VAR is significant one-step ahead. Fourth, disaggregation of the $\mathrm{CLI}_{\mathrm{CB}}$ into its components is not useful, likely because of the resulting extensive parameterization of the VAR and the related increased estimation uncertainty. Finally, the ranking of iterated forecasts and dynamic estimation is not clear cut, but for the best performing VAR using the four components of the $\mathrm{CCI}_{\mathrm{CB}}$ the standard iterated method decreases both the MSE and the MAE by about 10%.

In the middle and lower panels of Table 6 the comparison is repeated for, respectively, recessionary and expansionary periods. The most striking result is the major improvement of the ECM during recessions, for both forecast horizons. Yet, this finding should be interpreted with care since it is based on 18 observations only.

The second empirical example replicates and updates the analysis of Hamilton and Perez-Quiros (1996). They compared univariate and bivariate models, with and without Markov switching, for predicting one step ahead the turning points of (quarterly) GNP using the $\mathrm{CLI}_{\mathrm{CB}}$ as a leading indicator, named $\mathrm{CLI}_{\mathrm{DOC}}$ at that time. They found a minor role for switching (and for the use of real time data rather than final revisions), and instead a positive role for cointegration. Our first example highlighted that cointegration is not that relevant for forecasting during most of the recent period, and we wonder whether the role of switching has also changed. We use monthly data on the $\mathrm{CCI}_{\mathrm{CB}}$ and the $\mathrm{CLI}_{\mathrm{CB}}$, with the same estimation and forecast sample as in the previous example. The turning point probabilities for the linear models are computed by simulations, as described at the end of Section 6.1, using a two consecutive negative growth rule to identify recessions. For the MS we use the filtered recession probabilities. We also add to the comparison a probit model where the NBER based expansion/recession indicator is regressed on six lags of the $\mathrm{CLI}_{\mathrm{CB}}$. The NBER based expansion/recession indicator is also the target for the linear and MS based forecasts, as in Hamilton and Perez-Quiros (1996).

Table 6

Forecast comparison of alternative VAR models for CCI_{CB} and CLI_{CB}

		1 step-ahead		6 step-ahead DYNAMIC		6 step-ahead ITERATED	
		Relative MSE	Relative MAE	Relative MSE	Relative MAE	Relative MSE	Relative MAE
		Whole sample					
CCI + CLI	VAR(2)	1	1	1	1	1	1
CCI	AR(2)	1.001	1.010	0.982	0.963*	1.063	1.032
CCI + CLI coint	VECM(2)	1.042	1.074*	1.067	1.052	1.115	1.100
4 comp. of CCI + CLI	VAR(2)	0.904**	0.976	0.975	0.973	0.854**	0.911**
CCI + 10 comp. of CLI	VAR(1)	1.158***	1.114***	1.035	1.017	1.133**	1.100***
4 comp. CCI + 10 comp. CLI	VAR(1)	0.995	1.029	1.090	1.035	0.913	0.967
	VAR(2)	0.075	0.186	0.079	0.216	0.075	0.201
		Recessions					
CCI + CLI	VAR(2)	1	1	1	1	1	1
CCI	AR(2)	0.988	0.975	0.949	0.940	1.303**	1.154**
CCI + CLI coint	VECM(2)	0.681***	0.774***	0.744	0.882	0.478***	0.626***
4 comp. of CCI + CLI	VAR(2)	0.703*	0.784**	0.825	0.879	0.504***	0.672***
CCI + 10 comp. of CLI	VAR(1)	1.095	1.009	1.151	1.131	1.274*	1.117
4 comp. CCI + 10 comp. CLI	VAR(1)	0.947	0.852	1.037	1.034	0.614***	0.714***
	VAR(2)	0.087	0.258	0.096	0.252	0.163	0.368
		Expansions					
CCI + CLI	VAR(2)	1	1	1	1	1	1
CCI	AR(2)	1.002	1.016	0.977	0.956*	0.997	1.005
CCI + CLI coint	VECM(2)	1.090*	1.123***	1.118	1.081	1.292***	1.206***
4 comp. of CCI + CLI	VAR(2)	0.931*	1.007	0.987	0.980	0.952	0.964
CCI + 10 comp. of CLI	VAR(1)	1.166***	1.132***	1.015	0.997	1.093*	1.096**
4 comp. CCI + 10 comp. CLI	VAR(1)	1.001	1.058	1.087	1.029	0.997	1.023
	VAR(2)	0.074	0.177	0.076	0.208	0.065	0.183

Note: Forecast sample is: 1989:1–2003:12. First estimation sample is 1959:1–1988:12 (for 1 step-ahead) or 1959:1–1988:6 (for 6 step-ahead), recursively updated. Lag length selection by BIC. MSE and MAE are mean square and absolute forecast error. VAR for CCI_{CB} and CLI_{CB} is benchmark.

*indicates significance at 10%,

**indicates significance at 5%,

***indicates significance at 1% of the Diebold–Mariano test for the null hypothesis of no significant difference in MSE or MAE with respect to the benchmark.

Table 7
Turning point predictions

Target	Model	Relative MSE	Relative MAE
NBER	univariate	1.0302	1.2685***
(1 step-ahead)	univariate MS	1.3417	1.0431
	bivariate	1.0020	1.0512
	bivariate MS	0.6095	0.4800***
	probit CLI_CB	1	1
	probit	0.0754	0.1711

Note: One-step ahead turning point forecasts for the NBER expansion/recession indicator. Linear and MS models [as in Hamilton and Perez-Quiros (1996)] for CCI_{CB} and CLI_{CB}. Six lags of CLI_{CB} are used in the probit model.
***indicates significance at 1% of the Diebold–Mariano test for the null hypothesis of no significant difference in MSE or MAE with respect to the benchmark.

In Table 7 we report the MSE and MAE for each model relative to the probit, where the MSE is just a linear transformation of the QPS criterion of Diebold and Rudebusch (1989, 1991a, 1991b) and the Diebold and Mariano (1995) test for the statistical significance of the loss differentials. The results indicate a clear preference for the bivariate MS model, with the probit a far second best, notwithstanding its direct use of the target series as dependent variable. The turning point probabilities for the five models are graphed in Figure 6, together with the NBER dated recessions (shaded areas). The figure highlights that the probit model misses completely the 2001 recession, while both MS models indicate it, and also provide sharper signals for the 1990–1991 recession. Yet, the univariate MS model also gives several false alarms.

Our third empirical application is a more detailed analysis of the probit model. In particular, we consider whether the other composite leading indexes discussed in Section 7.2, the CLI_{ECRI}, CLI_{OECD}, and CLI_{SW}, or the three-month ten-year spread on the treasury bill rates have a better predictive performance than the CLI_{CB}. The estimation and forecasting sample is as in the first empirical example, and the specification of the probit models is as in the second example, namely, six lags of each CLI are used as regressors (more specifically, the symmetric one month percentage changes for CLI_{CB} and the one month growth rates for the other CLIs). We also consider a sixth probit model where three lags of each of the five indicators are included as regressors.

From Table 8, the model with the five indexes is clearly favored for one-step ahead turning point forecasts of the NBER based expansion/recession indicator, with large and significant gains with respect to the benchmark, which is based on the CLI_{CB}. The second best is the ECRI indicator, followed by OECD and SW. Repeating the analysis for six month ahead forecasts, the gap across models shrinks, the term spread becomes the first or second best (depending on the use of MSE or MAE), and the combination of the five indexes remains a good choice. Moreover, the models based on these variables

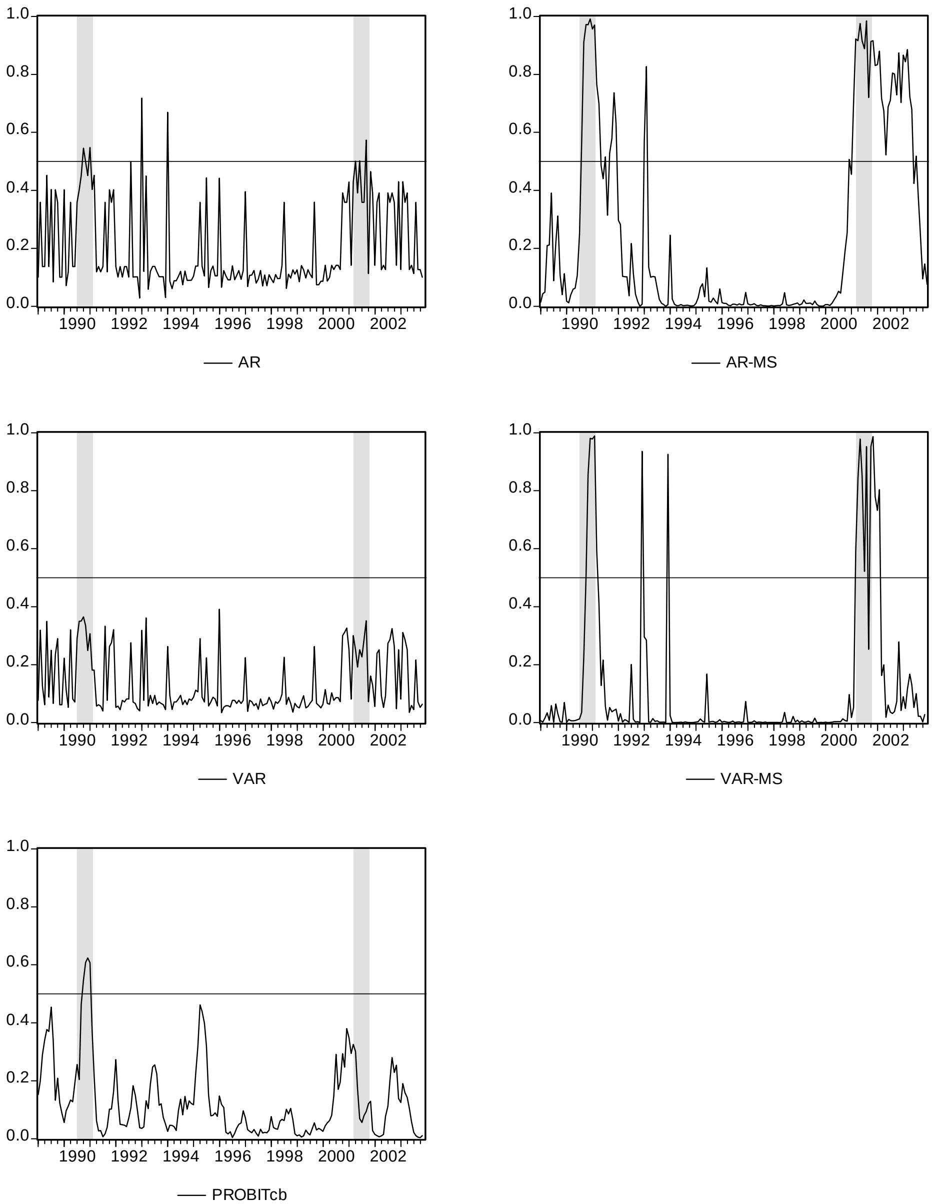

Figure 6. One month ahead recession probabilities. The models are those in Table 7. Shaded areas are NBER dated recessions.

Table 8
Forecasting performance of alternative CLIs using probit models for NBER recession/expansion classification

Target	Model	Relative MSE	Relative MAE
NBER	CLI_CB	1	1
(1 step-ahead)	CLI_SW	1.01	0.664***
	CLI_ECRI	0.588	0.597***
	CLI_OECD	0.719	0.714***
	termspread	0.952	0.937
	4 CLI + spread	0.565**	0.404***
NBER	CLI_CB	1	1
(6 step-ahead)	CLI_SW	1.085	0.956
	CLI_ECRI	0.888	0.948
	CLI_OECD	0.912	0.834**
	termspread	0.736**	0.726***
	4 CLI + spread	0.837**	0.692***
CLI_CB	1 step-ahead	0.073	0.169
	6 step-ahead	0.085	0.191

Note: Forecast sample is: 1989:1–2003:12. First estimation sample is 1959:1–1988:12, recursively updated. Fixed lag length: 6 lags for the first four models and 3 lags for the model with all four CLIs (see text for details). MSE and MAE are mean square and absolute forecast error. Probit model for CLI_{CB} is benchmark.
**indicates significance at 5%,
***indicates significance at 1% of the Diebold–Mariano test for the null hypothesis of no significant difference in MSE or MAE with respect to the benchmark.

(and also those using the ECRI and OECD indexes) provided early warnings for both recessions in the sample, see Figures 7 and 8.

The final empirical example we discuss evaluates the role of forecast combination as a tool for enhancing the predictive performance. In particular, we combine together the forecasts we have considered in each of the three previous examples, using either equal weights or the inverse of the MSEs obtained over the training sample 1985:1–1988:12. The results are reported in Table 9.

In the case of forecasts of the growth rate of the CCI_{CB}, upper panel, the pooled forecasts outperform most models but are slightly worse than the best performing single model, the VAR with the CLI_{CB} and the four components of the CCI_{CB} (compare with Table 6). The two forecast weighting schemes produce virtually identical results. For NBER turning point prediction, middle panel of Table 9, pooling linear and MS models cannot beat the best performing bivariate MS model (compare with Table 7), even when using the better performing equal weights for pooling or adding the probit model with the CLI_{CB} index as regressor into the forecast combination. Finally, also in the case of probit forecasts for the NBER turning points, lower panel of Table 9, a single model

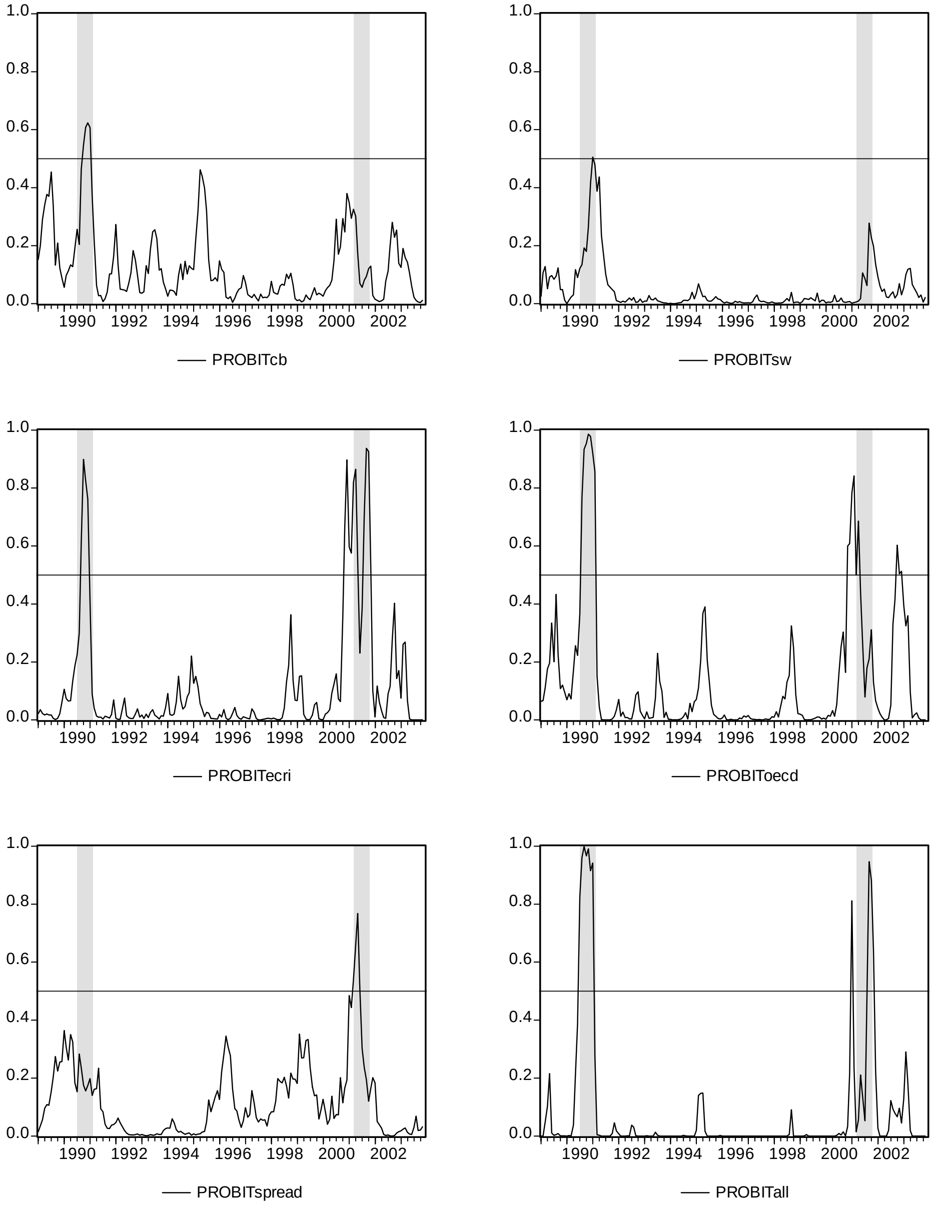

Figure 7. One month ahead recession probabilities for alternative probit models. The models are those in Table 8. Shaded areas are NBER dated recessions.

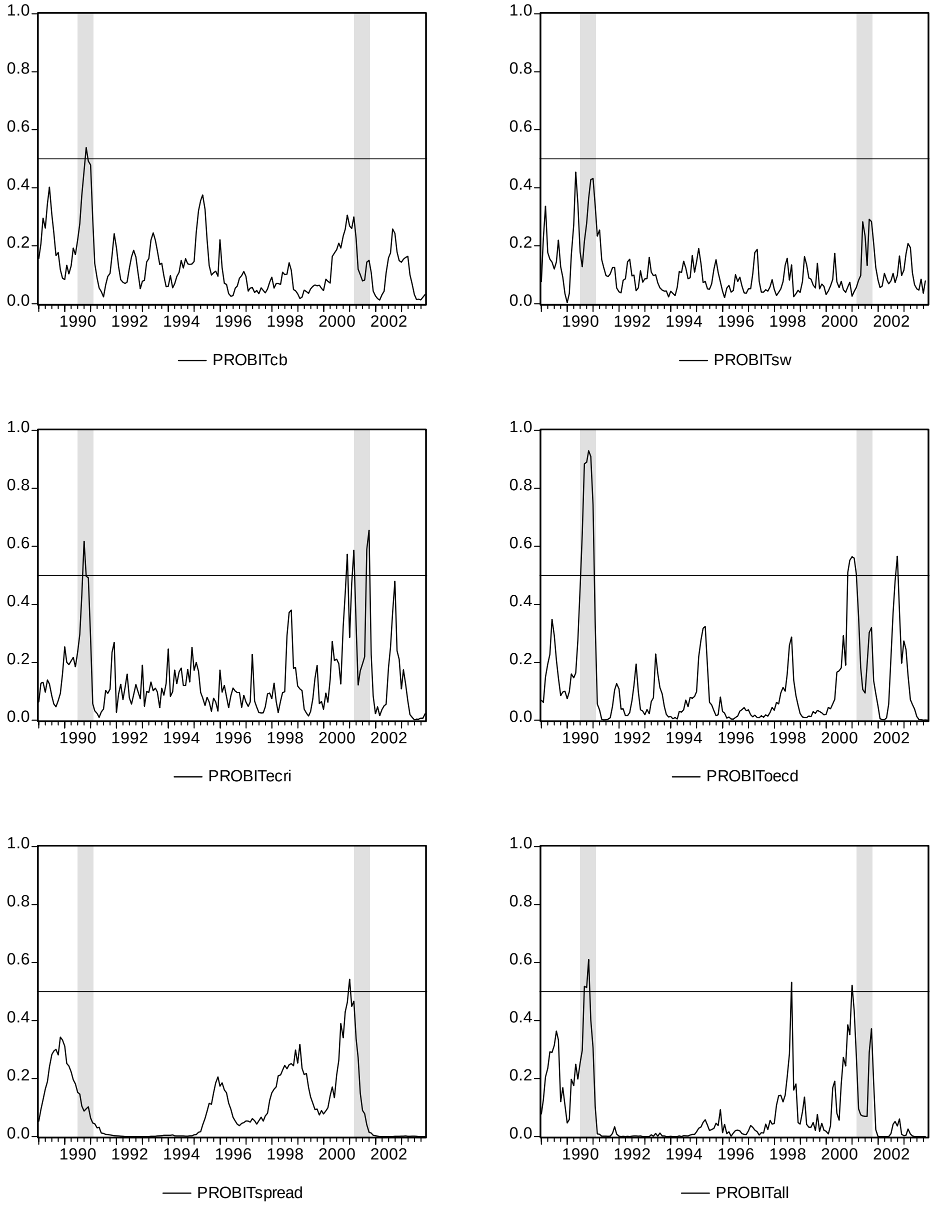

Figure 8. Six months ahead recession probabilities for alternative probit models. The models are those in Table 8. Shaded areas are NBER dated recessions.

Table 9
Evaluation of forecast pooling

Combine	Relative MSE	Relative MAE	Relative MSE	Relative MAE
	Predicting CCI_CB growth			
	MSE-weighted		Simple average	
6 linear models (1 month)	0.9474	0.9824	0.9418**	0.9781
6 linear models (6 month dynamic)	0.8873	0.9100	0.8863	0.9082
6 linear models (6 month iterated)	0.9352**	0.9776	0.9255**	0.9701
	Predicting NBER turning points			
	MSE-weighted		Simple average	
4 linear and MS models (1 month)	0.8683	1.1512	0.6676	0.9607
4 linear and MS models + probit (1 month)	0.8300	1.0989	0.6695	0.9686
	Predicting NBER turning points			
	MSE-weighted		Simple average	
5 single index PROBIT (1 month)	0.7423**	0.8028***	0.7014**	0.7844***
5 single index PROBIT + all (1 month)	0.6900**	0.7579***	0.6395**	0.7234***
5 single index PROBIT (6 months)	0.8863***	0.9069**	0.8667***	0.8956**
5 single index PROBIT + all (6 months)	0.8707***	0.8695***	0.8538***	0.8569***

Note: Forecast sample is 1989:1–2003:12. The forecasts pooled in the upper panel are from the six models in Table 6 and the benchmark is the VAR(2). The forecasts pooled in the middle panel are from the models in Table 7, including or excluding the probit, and the benchmark is the probit model with 6 lags of CLI_{CB} as regressor. The forecasts pooled in the lower panel are from the models in Table 8, including or excluding the probit with all indicators, and the benchmark is as in the middle panel.
**indicates significance at 5%,
***indicates significance at 1% of the Diebold–Mariano test for the null hypothesis of no significant difference in MSE or MAE with respect to the benchmark.

performs better than the pooled forecast for both one and six month horizons (compare Table 8), and equal weights slightly outperforms MSE based weights for pooling.

10. Review of the recent literature on the performance of leading indicators

Four main strands of research can be identified in the recent literature on the evaluation of the performance of leading indicators. First, the consequences of the use of real time information on the composite leading index and its components rather than the final releases. Second, the assessment of the relative performance of the new models for the coincident-leading indicators. Third, the evaluation of financial variables as leading indicators. Finally, the analysis of the behavior of the leading indicators during the two most recent US recessions as dated by the NBER, namely, July 1990–March 1991 and

March 2001–November 2001 [see, e.g., McNees (1991) for results on the previous recessions]. We now review in turn the main contributions in each field, grouping together the first two.

10.1. The performance of the new models with real time data

The importance of using real time data rather than final releases when evaluating the performance of the composite leading indicators was emphasized by Diebold and Rudebusch (1991a, 1991b). The rationale is that the composite indexes are periodically revised because of a variety of reasons including changes in data availability, timing or definition, modifications in the standardization factors, but also the past tracking performance of the index or some of its components; see Diebold and Rudebusch (1988), Swanson, Ghysels and Callan (1998) for an assessment of the revision process for the DOC-CB CLI, and Croushore (2006) for an updated overview on the use of real time data when forecasting. Therefore, an assessment of the usefulness of a composite leading index, even in a pseudo-real time framework but using the final release of the data, can yield biased results.

Diebold and Rudebusch (1991b) estimated a linear dynamic model for IP and the CLI, using dynamic estimation, and evaluated the marginal predictive content of the CLI in sample and recursively out of sample (for 1969–1988) using both finally and first released data for the CLI. While in the first two cases inclusion of the CLI in the model systematically reduces the MSFE, in the third one the results are not clear cut and depend on the lag-length and the forecast horizon. A similar finding emerges using the CCI instead of IP as the target variable, and when the Neftci's (1982) algorithm is adopted to predict turning points in IP [Diebold and Rudebusch (1991a)]. Instead, using an MS model for predicting turning points, Lahiri and Wang (1994) found the results to be rather robust to the use of historical or revised data on the DOC CLI.

Filardo (1999) analyzed the performance of simple rules of thumb applied to the CLI_{CB} and of the recession probabilities computed using Neftci's (1982) formula, a linear model, a probit model, and SW's CRI, using both final and first released data over the period 1977–1998. Overall, rules of thumb and the Neftci's formula applied to the CLI_{CB} performed poorly, better with ex-post data; probit and linear models were robust to the adoption of the real-time data, because of the use of mostly financial variables as regressors, while SW's CRI was not evaluated in real time. Since the models were not directly compared on the same grounds, a ranking is not feasible but, overall, the results point towards the importance of using real-time data for the CLI also over a different and more recent sample than Diebold and Rudebusch (1991a, 1991b).

Hamilton and Perez-Quiros (1996) evaluated the usefulness of the DOC-CB CLI using linear and MS VARs, with and without cointegration, finding that the best model for predicting GDP growth and turning points over the period 1975–1993 is the linear VAR (cointegration matters in sample but not out of sample), and in this framework the CLI appears to have predictive content also with real-time data. A similar conclusion emerged from the analysis of Camacho and Perez-Quiros (2002) for the period 1972–

1998, even though they found that nonlinearity matters, the MS model was the best in and out of sample. Even better is a combination of the MS model with the nonparametric forecast described in Section 8.2.

A few studies compared the models described in Sections 6 and 8 using the final release of the data. Notice that this is less problematic in comparative analyses than in single model evaluation since all the methods can be expected to be equally advantaged. Layton and Katsuura (2001) considered logit and probit models, and a Filardo (1994) type time-varying (static) MS model, using the ECRI coincident and leading indexes. The latter model performed best in a pseudo-real time evaluation exercise over the period 1979–1999, and was found to be quite useful in dating the business cycle in Layton (1998), confirming the findings in Filardo (1994). Instead, Birchenhall et al. (1999) found more support for the probit model than for the MS specification.

10.2. Financial variables as leading indicators

Though financial variables have a long history as leading indicators, e.g., Mitchell and Burns (1938) included the Dow Jones composite index of stock prices in their list of leading indicators for the US economy, a systematic evaluation of their forecasting performance started much later, in the ’80s, and since then attracted increased attention.

Stock and Watson (2003b) reviewed over 90 articles dealing with the usefulness of financial indicators for predicting output growth (and inflation), and we refer to them and to Kozicki (1997) and Dotsey (1998) for details on single studies. They also provided their own evaluation using several indicators for the G7 countries and, on the basis of the survey and of their results, concluded that some asset prices have significant predictive content at some times in some countries, but it is not possible to find a single indicator with a consistently good performance for all countries and time periods. While pooling provided a partial solution to the instability problem, Stock and Watson (2003a) suggested that “... the challenge is to develop methods better geared to the intermittent and evolving nature of these predictive relations” (p. 4).

The evidence reported in the previous and next subsection indeed points towards the usefulness of models with time-varying parameters, and also confirms the necessity of a careful choice of the financial variables to be used as leading indicators and of a continuous monitoring of their performance. A rapid survey of the literature on the interest rate spreads provides a clear and valuable illustration and clarification for this statement.

As mentioned in Section 7.2, Stock and Watson (1989) included two spreads into their CLI, a paper-bill spread (the difference between the 6-month commercial paper rate and the 6-month Treasury bill rate) and a term spread (the difference between the 10-year and the 1-year Treasury bond rates).

The paper-bill spread tends to widen before a recession reflecting expectations of business bankruptcies, corporations’ growing cash requirements near the peak of the business cycle, and tighter monetary policy (the paper rate rises because banks deny loans due to the restricted growth of bank reserves, so that potential borrowers seek funds in the commercial paper marker). Yet, the paper bill-spread could also change for other reasons unrelated to the business cycle, such as changes in the Treasury's debt management policy, or foreign central banks interventions in the exchange market since a large amount of their reserves in dollars are invested in Treasury bills; see, e.g., Friedman and Kutnner (1998), who found these reasons capable of explaining the bad leading performance of the paper-bill spread for the 1990–1991 recession, combined with the lack of a tighter monetary policy. The performance for the 2001 recession was also unsatisfactory, the spread was small and declining from August 2000 to the end of 2001, see also the next subsection.

The term spread has two components, expected changes in interest rates and the term premium for higher risk and/or lower liquidity. Therefore the commonly observed negative slope of the term structure prior to recession, i.e., long term rates becoming lower than short term ones, can be due either to lower expected short term rates (signaling expansionary monetary policy) or to lower term premia. Hamilton and Kim (2002) found both components to be relevant for forecasting output growth, with the former dominating at longer forecast horizons. The bad leading performance of the term spread for the 1990–1991 recession is also typically attributed to the lack of a tighter monetary policy in this specific occasion. The term spread became instead negative from June 2000 through March 2001, anticipating the recession of 2001, but the magnitude was so small by historical standards that, for example, SW's composite leading index did not signal the recession, see also the next subsection.

Gertler and Lown (2000) suggested to use the high-yield (junk)/AAA bond spread as a leading indicator, since it is less sensitive to monetary policy and provides a good proxy for the premium for external funds, i.e., for the difference between the costs of external funds and the opportunity costs of using internal funds. The premium for external funds moves countercyclically, since during expansions the borrowers' financial position typically improves, and this further fosters the aggregate activity; see, e.g., Bernanke and Gertler (1989) for a formalization of this final accelerator mechanism. Therefore, a widening high-yield spread signals a deterioration of economic conditions. Gertler and Lown (2000) found that after the mid '80s the high-yield spread had a better forecasting performance than both the paper-bill and the term spreads for the US GDP growth, also providing a warning for the 1990–1991 recession. Yet, as for the paper-bill spread, the high-yield spread can also change for reasons unrelated with the business cycle, such as confidence crises in emerging markets. In particular, Duca (1999) indicated that the widening of the spread prior to the 1990–1991 recession could be an accidental event related with the thrift crisis and the associated sale of junk bonds in an illiquid market.

A related question of interest is whether it is better to use a financial indicator in isolation or as a component of a composite index. Estrella and Mishkin (1998) ran probit regressions using the term-spread, the $\mathrm{CLI}_{\mathrm{CB}}$, the $\mathrm{CLI}_{\mathrm{SW}}$, and some of their components, concluding that both in sample and out of sample the spread yields the largest forecasting gains. Moreover, addition of other regressors is in general harmful, except for the NYSE index returns. Similar conclusions emerged from the analysis in Dueker

(1997), who also used more complicated versions of the probit model, allowing for dynamics and Markov switching parameters. Qi (2001) also obtained a similar finding using the neural network model described in Section 8.2. The $\mathrm{CLI}_{\mathrm{SW}}$ was best at 1-quarter forecast horizon, but the term spread at 2- to 6-quarter horizon. Yet, she also detected substantial instability of the results over different decades, namely, the ’70s, ’80s, and ’90s. Estrella, Rodrigues and Schich (2003) also found some instability for the US, more so when the dependent variable is the GDP growth rate than when it is a binary expansion/recession indicator.

Chauvet and Potter (2001a) detected substantial instability also in the probit model when it is estimated with the Gibbs sampler. Moreover, the date of the break has a major role in determining the predictive performance of the spread, for example, the probability of a future recession are about 45% in December 2000 when no break is assumed but increase to 90% imposing a break in 1984. Unfortunately, there is considerable uncertainty about the break date, so that the posterior mean probability of recession across all break dates is 32% with a 95% interval covering basically the whole $[0, 1]$ interval. Chauvet and Potter (2001b) extended the basic probit model to allow for parameter instability, using a time-varying specification, and also for autocorrelated errors. Though the more complicated models performed better, along the lines of Dueker (1997), they provided a weaker signal of recession in 2001 in a real-time evaluation exercise.

Finally, positive results on the leading properties of the term spread and other financial variables for other countries were reported, e.g., by Davis and Henry (1994), Davis and Fagan (1997), Estrella and Mishkin (1997), Estrella, Rodrigues and Schich (2003), and Moneta (2003). Yet, Moneta (2003) found also for the Euro area a deterioration in the relative leading characteristics of the spread after the ’80s, and an overall unsatisfactory performance in predicting the Euro area recession of the early ’90s.

10.3. The 1990–1991 and 2001 US recessions

Stock and Watson (1993) conducted a detailed analysis of possible reasons for the failure of their CRI to produce early warnings of the 1990–1991 recession. They could not detect any signs of model failure or mis-specification and therefore concluded that the major problem was the peculiar origin of this recession compared with its predecessors, namely, a deterioration in the expectations climate followed by a drop in consumption. In such a case, the treasury bill yield curve, exchange rates, and partly IP provided wrong signals. Only three other leading indicators in their set gave moderate negative signals, part-time work, building permits and unfilled orders, but they were not sufficiently strong to offset the other indicators.

Phillips (1998–1999) compared the performance of the $\mathrm{CRI}_{\mathrm{SW}}$, and of the $\mathrm{CLI}_{\mathrm{CB}}$ and the term spread, transformed into probabilities of recession using Neftci’s (1982) formula, for forecasting the 1990–1991 recession using real time data. He found that the $\mathrm{CLI}_{\mathrm{CB}}$ produced the best results. Moreover, the SW’s index modified to allow for longer lags on the term and quality spreads worked better in sample but not for this recession.

Chauvet (1998) also used a real time dataset to produce recession forecasts from her dynamic MS factor model, and found that the filtered probability of recession peaked beyond 0.5 already at the beginning of 1990 and then in May of that year.

Filardo and Gordon (1999) contrasted a linear VAR model, an MS model with time-varying parameters, the SW's model, and an MS factor model with time-varying parameters, along the lines of Chauvet (1998). All models were estimated using Gibbs sampling techniques, and compared on the basis of the marginalized likelihoods and Bayes factors in 1990, as suggested by Geweke (1994), since these quantities are easily computed as a by-product of the estimation. They found that all models performed comparatively over the period January–June, but in the second part of the year, when the recession started, the MS model was ranked first, the VAR second, and the factor model third, with only minor differences between the two versions.

Filardo (2002), using the same models as in Filardo (1999) found that the two-month rule on the $\mathrm{CLI}_{\mathrm{CB}}$ worked well in predicting the 2001 recession, but sent several false alarms in the '90s. A probit model with a 3-month forecast horizon and the term spread, corporate spread, S&P500 returns and the $\mathrm{CLI}_{\mathrm{CB}}$ as regressors also worked well, predicting the beginning of the recession in January 2001 using a 50% rule. Instead, the $\mathrm{CRI}_{\mathrm{SW}}$ did not perform well using a 50% rule, while SW's CRI-C (contemporaneous) worked better but was subject to large revisions.

Stock and Watson (2003a) analyzed in details the reasons for the poor performance of the CRI, concluding that is was mostly due to the particular origin of the recession (coming from the decline in stock prices and business investment), which is not properly reflected by most of the indicators in their CRI. In particular, the best indicators for the GDP growth rate were the term spread, the short term interest rate, the junk bond spread, stock prices, and new claims for unemployment. Notice that most of these variables are included in Filardo's (2002) probit models. Moreover, they found that pooled forecasts worked well, but less well than some single indicators in the list reported above.

Dueker (2003) found that his Qual-VAR predicted the timing of the 2001 recession quite well relative to the professional forecasters, while the evidence in Dueker and Wesche (2001) is more mixed. Dueker (2002) noticed that an MS-probit model with the $\mathrm{CLI}_{\mathrm{CB}}$ as regressor worked also rather well in this occasion, providing a 6-month warning of the beginning of the recession (but not in the case of the previous recession).

Overall, some differences in the ranking of models and usefulness of the leading indicators emerged because of the choice of the specific coincident and leading variables, sample period, criteria of evaluation, etc. Yet, a few findings are rather robust. First, indicator selection and combination methods are important, and there is hardly a one fits all choice, even though financial variables and the equal weighted $\mathrm{CLI}_{\mathrm{CB}}$ seem to have a good average performance. Second, the model that relates coincident and leading indicators also matters, and an MS feature is systematically helpful. Finally, pooling the forecasts produced good results whenever applied, even though there is only limited evidence as far as turning points are concerned.

11. What have we learned?

The experience of the last two recessions in the US confirmed that these are difficult events to predict, because the generating shocks and their propagation mechanism change from time to time, and there is a very limited sample to fit the more and more complex models that try to capture these time-varying features. Nonetheless, the recent literature on leading indicators provided several new useful insights for the prediction of growth rates and turning points of a target variable.

The first set of improvements is just in the definition of the target variable. In Section 5 we have seen that several formal procedures were developed to combine coincident indicators into a composite index, which is in general preferable to monitoring a single indicator because of its narrower coverage of the economy. In practice, the new model based CCIs are very similar to the old-style equal averages of the (standardized) coincident indicators, such as the $\mathrm{CCI}_{\mathrm{CB}}$, but they provide a sounder statistical framework for the use and evaluation of the CCIs. More sophisticated filtering procedures were also developed to emphasize the business cycle information in a CCI, as detailed in Section 3, even though substantial care should be exerted in their implementation to avoid phase shifts and other distortions. New methods were also developed for dating the peaks and troughs in either the classical or the deviation cycle. They closely reproduce the NBER dating for the US and the CEPR dating for the Euro area, but are more timely and can also provide a probabilistic measure of uncertainty around the dated turning points.

The second set of advances concerns the construction of leading indicators. While there was general agreement on the characteristics of a good leading indicator, such as consistent timing or conformity to the general business cycle, in Section 2 we have seen that there are now better methods to formally test the presence of these characteristics and assess their extent. Moreover, there were several developments in the construction of the composite leading indexes, ranging from taking into explicit account data problems such as missing values or measurement error, to an even more careful variable selection relying on new economic and statistical theories, combined with sounder statistical procedures for merging the individual leading indicators into a CLI, as described in Sections 6 and 7.

The third, and perhaps most important, set of enhancements is in the use of the leading indicators. In Sections 6 and 8 we have seen that simple rules to transform a CLI into a turning point forecast have been substituted with sophisticated nonlinear and time-varying models for the joint evolution of the coincident and leading indicators. Moreover, mainly using simulation-based techniques, it is now rather easy to use a model to produce both point and probability and duration forecasts.

The final set of improvements is in the evaluation of leading indicators. In Section 9 we have seen that formal statistical methods are now available to assess the forecasting performance of leading indicators, possibly combined with the use of real time data to prevent biased favorable results due to revisions in the composition of the CLIs. Moreover, the overview in Section 10 of the forecasting performance over the two most recent recessions in the US has provided some evidence in favor of the forecasting capabilities of CLIs, in particular when simple weighting procedures are applied to a rather large set of indicators, combined with sophisticated models for the resulting CLI and the target variable.

Notwithstanding the substantial progress in the recent years, there is still considerable scope for research in this area. For example, it might be useful to achieve a stronger consensus on the choice of the target variable, and in particular on whether the classical cycle is really the target of interest or a deviation cycle could provide more useful information. The collection of higher quality monthly series and the development of better methods to handle data irregularities also deserve attention. But the crucial element remains the selection of the leading variables, and of the weighting scheme for their combination into a CLI. Both choices should be made endogenous and frequently updated to react to the changing shocks that hit the economy, and further progress is required in this area. Forecast pooling could provide an easier method to obtain more robust predictions, but very limited evidence is available for turning point and duration forecasts. It is also worth mentioning that while in this chapter we have focused on real activity as the target variable, other choices are possible such as inflation or a stock market index [see, e.g., the contributions in Lahiri and Moore (1991)] and most of the developments we have surveyed could be usefully applied in these related contexts.

Uncited references

[Simpson, Osborn and Sensier (2001)]

References

Albert, J., Chib, S. (1993a). "Bayesian analysis via Gibbs sampling of autoregressive time series subject to Markov mean and variance shifts". Journal of Business and Economic Statistics 11, 1–15.

Albert, J., Chib, S. (1993b). "Bayesian analysis of binary and polychotomous response data". Journal of the American Statistical Association 88, 669–679.

Altissimo, F., Bassanetti, A., Cristadoro, R., Forni, M., Lippi, M., Reichlin, L., Veronese, G. (2001). "Euro-Coin: A real time coincident indicator of the Euro area business cycle". CEPR Working Paper 3108.

Anderson, H.M., Vahid, F. (2001). "Predicting the probability of a recession with nonlinear autoregressive leading-indicator models". Macroeconomic Dynamics 5, 482–505.

Artis, M.J., Banerjee, A., Marcellino, M. (2005). "Factor forecasts for the UK". Journal of Forecasting. In press.

Artis, M.J., Galvao, A.B., Marcellino, M. (2003). "The transmission mechanism in a changing world". CEPR Working Paper No. 4014.

Artis, M.J., Krolzig, H.-M., Toro, J. (2004). "The European business cycle". Oxford Economic Papers 56 (1), 1–44.

Artis, M.J., Marcellino, M. (2001). "Fiscal forecasting: The track record of IMF, OECD and EC". Econometrics Journal 4, s20–s36.

Artis, M.J., Marcellino, M., Proietti, T. (2004). "Dating the Euro area business cycle". Oxford Bulletin of Economics and Statistics 66, 537–565.

Artis, M.J., Osborn, D., Bladen-Hovell, R., Smith, G., Zhang, W. (1995). "Turning point prediction for the UK using CSO leading indicators". Oxford Economic Papers 4, July.

Artis, M.J., Canova, F., Galí, J., Giavazzi, F., Portes, R., Reichlin, L., Uhlig, H., Weil, P. (2003). "Business cycle dating committee of the Centre for Economic Policy Research". CEPR.

Auerbach, A.J. (1982). "The index of leading indicators: "Measurement without theory", thirty-five years later". Review of Economics and Statistics 64 (4), 589–595.

Bai, J. (2003). "Inferential theory for factor models of large dimensions". Econometrica 71, 135–171.

Bai, J., Ng, S. (2002). "Determining the number of factors in approximate factor models". Econometrica 70 (1), 191–221.

Bai, J., Ng, S. (2003). "Confidence intervals for diffusion index forecasts with a large number of predictors". Working Paper, University of Michigan.

Banerjee, A.N., Marcellino, M. (2005). "Are there any reliable leading indicators for US inflation and GDP growth?" International Journal of Forecasting. In press.

Banerjee, A.N., Marcellino, M., Masten, I. (2005). "Forecasting macroeconomic variables for the accession countries". In: Artis, M., Banerjee, A., Marcellino, M. (Eds.), The European Enlargement: Prospects and Challenges. Cambridge University Press, Cambridge. In press.

Bates, J.M., Granger, C.W.J. (1969). "The combination of forecasts". Operations Research Quarterly 20, 451–468.

Baxter, M., King, R.G. (1999). "Measuring business cycles: Approximate band-pass filters for macroeconomic time series". Review of Economics and Statistics 81 (4), 575–593.

Bernanke, B.S., Gertler, M.L. (1989). "Agency costs, net worth, and business fluctuations". American Economic Review 79, 14–31.

Bernard, H., Gerlach, S. (1998). "Does the term structure predict recessions? The international evidence". International Journal of Finance and Economics 3, 195–215.

Birchenhall, C.R., Osborn, D.R., Sensier, M. (2001). "Predicting UK business cycle regimes". Scottish Journal of Political Economy 48 (2), 179–195.

Birchenhall, C.R., Jessen, H., Osborn, D.R., Simpson, P. (1999). "Predicting US business cycle regimes". Journal of Business and Economic Statistics 17 (3), 313–323.

Boehm, E.A. (2001). "The contribution of economic indicator analysis to understanding and forecasting business cycles". Indian Economic Review 36, 1–36.

Boivin, J., Ng, S. (2003). "Are more data always better for factor analysis". NBER Working Paper No. 9829. Journal of Econometrics. In press.

Boldin, M.D. (1994). "Dating turning points in the business cycle". Journal of Business 67 (1), 97–131.

Boschan, C., Banerji, A. (1990). "A reassessment of composite indexes". In: Klein, P.A. (Ed.), Analyzing Modern Business Cycles. M.E. Sharpe, Armonk, NY.

Brillinger, D.R. (1981). Time Series Data Analysis and Theory. Holt, Rinehart, and Winston, New York.

Bry, G., Boschan, C. (1971). "Cyclical analysis of time series: Selected procedures and computer programs". Technical Paper 20, NBER. Columbia University Press.

Burns, A.F., Mitchell, W.C. (1946). "Measuring business cycles". NBER Studies in Business Cycles No. 2, New York.

Camacho, M. (2004). "Vector smooth transition regression models for US GDP and the composite index of leading indicators". Journal of Forecasting 23 (3), 173–196.

Camacho, M., Perez-Quiros, G. (2002). "This is what the leading indicators lead". Journal of Applied Econometrics 17 (1), 61–80.

Camba-Mendez, G., Kapetanios, G., Smith, R.J., Weale, M.R. (2001). "An automatic leading indicator of economic activity: Forecasting GDP growth for European countries". Econometrics Journal 4 (1), S56–S90.

Canova, F. (1999). "Reference cycle and turning points: A sensitivity to detrending and classification rules". Economic Journal 112, 117–142.

Canova, F. (2004). Methods for Applied Macroeconomic Research. Princeton University Press. In press.

Canova, F., Ciccarelli, M. (2001). "Forecasting and turning point predictions in a Bayesian panel VAR model". Journal of Econometrics 120 (2), 327–359.

Canova, F., Ciccarelli, M. (2003). “Panel index VAR models: Specification, estimation, testing and leading indicators”. CEPR Discussion Paper No. 4033.

Carriero, A., Marcellino, M. (2005). “Building composite coincident indicators for European countries”. Mimeo, Bocconi University.

Carter, C.K., Kohn, R. (1994). “On Gibbs sampling for state space models”. Biometrika 81, 541–555.

Chauvet, M. (1998). “An econometric characterization of business cycle dynamics with factor structure and regime switching”. International Economic Review 39 (4), 969–996.

Chauvet, M., Piger, J.M. (2003). “Identifying business cycle turning points in real time”. Federal Reserve Bank of St. Louis Review 85 (2), 13–26.

Chauvet, M., Potter, S. (2001a). “Predicting a recession: Evidence from the yield curve in the presence of structural breaks”. Economic Letters 77 (2), 245–253.

Chauvet, M., Potter, S. (2001b). “Forecasting recessions: Using the yield curve”. Staff Reports 134, Federal Reserve Bank of New York. Journal of Forecasting. In press.

Chen, C.W.S., Lee, J.C. (1995). “Bayesian inference of threshold autoregressive models”. Journal of Time Series Analysis 16, 483–492.

Chin, D., Geweke, J., Miller, P. (2000). “Predicting turning points”. Federal Reserve Bank of Minneapolis Staff Report No. 267.

Christiano, L., Fitzgerald, T. (2003). “The band pass filter”. International Economic Review 44 (2), 435–465.

Clark, T.E., McCracken, M.W. (2001). “Test of equal forecast accuracy and encompassing for nested models”. Journal of Econometrics 105, 85–100.

Clements, M.P., Hendry, D.F. (1996). “Multi-step estimation for forecasting”. Oxford Bulletin of Economics and Statistics 58, 657–684.

Clements, M.P., Hendry, D.F. (1999). Forecasting Non-Stationary Economic Time Series. MIT Press, Cambridge, MA.

Corradi, V., Swanson, N.R. (2006). “Density and interval forecasting”. In: Elliott, G., Granger, C.W.J., Timmermann, A. (Eds.), Handbook of Economic Forecasting. Elsevier, Amsterdam. Chapter 5 in this volume.

Croushore, D. (2006). “Real-time data”. In: Elliott, G., Granger, C.W.J., Timmermann, A. (Eds.), Handbook of Economic Forecasting. Elsevier, Amsterdam. Chapter 17 in this volume.

Davis, E.P., Fagan, G. (1997). “Are financial spreads useful indicators of future inflation and output growth in EU countries?” Journal of Applied Econometrics 12, 701–714.

Davis, E.P., Henry, S.G.B. (1994). “The use of financial spreads as indicator variables: Evidence for the United Kingdom and Germany”. IMF Staff Papers 41, 517–525.

Del Negro, M. (2001). “Turn, turn, turn: Predicting turning points in economic activity”. Federal Reserve Bank of Atlanta Economic Review 87.

Diebold, F.X., Lee, J.-H., Weinbach, G.C. (1994). “Regime switching with time-varying transition probabilities”. In: Hargreaves, C. (Ed.), Non-Stationary Time-Series Analyses and Cointegration. Oxford University Press, Oxford, pp. 283–302.

Diebold, F.X., Mariano, R. (1995). “Comparing predictive accuracy”. Journal of Business and Economic Statistics 13, 253–263.

Diebold, F.X., Rudebusch, G.D. (1988). “Stochastic properties of revisions in the index of leading indicators”. In: Proceedings of the Business and Economic Statistics Section. American Statistical Association, pp. 712–717.

Diebold, F.X., Rudebusch, G.D. (1989). “Scoring the leading indicators”. The Journal of Business 62 (3), 369–391.

Diebold, F.X., Rudebusch, G.D. (1990). “A nonparametric investigation of duration dependence in the American business cycle”. Journal of Political Economy 98, 596–616.

Diebold, F.X., Rudebusch, G.D. (1991a). “Turning point prediction with the composite leading index: An ex ante analysis”. In: Lahiri, K., Moore, G.H. (Eds.), Leading Economic Indicators: New Approaches and Forecasting Records. Cambridge University Press, Cambridge, UK, pp. 231–256.

Diebold, F.X., Rudebusch, G.D. (1991b). “Forecasting output with composite leading index: A real-time analysis”. Journal of the American Statistical Association 86 (415), 603–610.

Diebold, F.X., Rudebusch, G.D. (1996). "Measuring business cycles: A modern perspective". The Review of Economics and Statistics 78 (1), 67–77.

Diebold, F.X., Rudebusch, G.D., Sichel, D.E. (1993). "Further evidence on business cycle duration dependence". In: Stock, J.H., Watson, M.W. (Eds.), Business Cycles, Indicators, and Forecasting. The University of Chicago Press, Chicago, pp. 255–280.

Doan, T., Litterman, R., Sims, C. (1984). "Forecasting and conditional projection using realistic prior distributions". Econometric Reviews 3, 1–100.

Dotsey, M. (1998). "The predictive content of the interest rate term spread for future economic growth". Federal Reserve Bank of Richmond Economic Quarterly 84 (3), 31–51.

Duca, J.V. (1999). "What credit market indicators tell us". Federal Reserve Bank of Dallas Economic and Financial Review 1999:Q3, 2–13.

Dueker, M.J. (1997). "Strengthening the case for the yield curve as a predictor of US recessions". Federal Reserve Bank of St. Louis Review 79 (2), 41–51.

Dueker, M.J. (2002). "Regime-dependent recession forecasts and the 2001 recession". Federal Reserve Bank of St. Louis Review 84 (6), 29–36.

Dueker, M.J. (2003). "Dynamic forecasts of qualitative variables: A Qual VAR model of US recessions". Journal of Business and Economic Statistics. In press.

Dueker, M.J., Wesche, K. (2001). "Forecasting output with information from business cycle turning points: A qualitative variable VAR". Federal Reserve Bank of St. Louis Working Paper 2001-019B.

Durland, J.M., McCurdy, T.H. (1994). "Duration-dependent transitions in a Markov model of US GNP growth". Journal of Business and Economic Statistics 12, 279–288.

Eichengreen, B., Watson, M.W., Grossman, R.S. (1985). "Bank rate policy under the interwar gold standard". Economic Journal 95, 725–745.

Elliott, G., Komunjer, I., Timmermann, A. (2003). "Estimating loss function parameters". CEPR Discussion Papers No. 3821. Submitted Review of Economic Studies.

Emerson, R.A., Hendry, D.F. (1996). "An evaluation of forecasting using leading indicators". Journal of Forecasting 15, 271–291.

Estrella, A., Mishkin, F.S. (1997). "The predictive power of the term structure of interest rates in Europe and the United States: Implications for the European Central Bank". European Economic Review 41, 1375–1401.

Estrella, A., Mishkin, F.S. (1998). "Predicting US recessions: Financial variables as leading indicators". The Review of Economics and Statistics 80 (1), 45–61.

Estrella, A., Rodrigues, A.P., Schich, S. (2003). "How stable is the predictive power of the yield curve? Evidence from Germany and the United States". Review of Economics and Statistics 85 (3), 629–644.

Fair, R. (1993). "Estimating event probabilities from macroeconomic models using stochastic simulation". In: Stock, J.H., Watson, M.W. (Eds.), Business Cycles, Indicators, and Forecasting. The University of Chicago Press, Chicago, pp. 157–178.

Filardo, A.J. (1994). "Business cycle phases and their transitional dynamics". Journal of Business and Economic Statistics 12 (3), 299–308.

Filardo, A.J. (1999). "How reliable are recession prediction models?" Federal Reserve Bank of Kansas City Economic Review 84 (2), 35–55.

Filardo, A.J. (2002). "The 2001 US recession: What did recession prediction models tell us?" Paper prepared for a book honoring Geoffrey H. Moore, Bank for International Settlements.

Filardo, A.J., Gordon, S.F. (1994). "International co-movements of business cycles". Research Working Paper 94-11, Federal Reserve Bank of Kansas.

Filardo, A.J., Gordon, S.F. (1998). "Business cycle durations". Journal of Econometrics 85, 99–123.

Filardo, A.J., Gordon, S.F. (1999). "Business cycle turning points: Two empirical business cycle model approaches". In: Rothman, P. (Ed.), Nonlinear Time Series Analysis of Economic and Financial Data, vol. 1. Kluwer Academic, pp. 1–32. Chapter I.

Forni, M., Hallin, M., Lippi, M., Reichlin, L. (2000). "The generalized factor model: Identification and estimation". The Review of Economics and Statistics 82 (4), 540–554.

Forni, M., Hallin, M., Lippi, M., Reichlin, L. (2001). "Coincident and leading indicators for the Euro area". The Economic Journal 111 (May), C62–C85.
Forni, M., Hallin, M., Lippi, M., Reichlin, L. (2003a). "The generalized dynamic factor model: One-sided estimation and forecasting". Journal of the American Statistical Association. In press.
Forni, M., Hallin, M., Lippi, M., Reichlin, L. (2003b). "Do financial variables help forecasting inflation and real activity in the Euro area?" Journal of Monetary Economics 50, 1243–1255.
Friedman, B.M., Kutnner, K.N. (1998). "Indicator properties of the paper-bill spread: Lessons from recent experience". The Review of Economics and Statistics 80 (1), 34–44.
Fuhrer, J.C., Rudebusch, G.D. (2004). "Estimating the Euler equation for output". Journal of Monetary Economics 51 (September), 1133–1153.
Fuhrer, J.C., Schuh, S. (1998). "Beyond shocks: What causes business cycles? An overview". Proceedings from the Federal Reserve Bank of Boston Conference Series No. 42, pp. 1–31.
Gertler, M., Lown, C.S. (2000). "The information in the high yield bond spread for the business cycle: Evidence and some implications". NBER Working Paper Series No. 7549.
Geweke, J. (1977). "The dynamic factor analysis of economic time series". In: Aigner, D.J., Goldberger, A.S. (Eds.), Latent Variables in Socio-Economic Models. North-Holland, Amsterdam. Chapter 19.
Geweke, J. (1994). "Variable selection and model comparison in regression". In: Berger, J.O., Bernardo, J.M., Dawid, A.P., Smith, A.F.M. (Eds.), Proceedings of the Fifth Valencia International Meetings on Bayesian Statistics.
Geweke, J., Terui, N. (1993). "Bayesian threshold autoregressive models for nonlinear time series". Journal of Time Series Analysis 14, 441–455.
Geweke, J., Whiteman, C.H. (2006). "Bayesian forecasting". In: Elliott, G., Granger, C.W.J., Timmermann, A. (Eds.), The Handbook of Economic Forecasting. Elsevier, Amsterdam. Chapter 1 in this volume.
Granger, C.W.J., Machina, M.J. (2006). "Decision theory and forecasting". In: Elliott, G., Granger, C.W.J., Timmermann, A. (Eds.), The Handbook of Economic Forecasting. Elsevier, Amsterdam. Chapter 2 in this volume.
Granger, C.W.J., Ramanathan, R. (1984). "Improved methods of combining forecasts". Journal of Forecasting 3, 197–204.
Granger, C.W.J., Teräsvirta, T., Anderson, H.M. (1993). "Modelling nonlinearity over the business cycle". In: Stock, J.H., Watson, M.W. (Eds.), Business Cycles, Indicators, and Forecasting. The University of Chicago Press, Chicago, pp. 311–325.
Hall, R.E., Feldstein, M., Frankel, J., Gordon, R., Romer, C., Romer, D., Zarnowitz, V. (2003). "The NBER's recession dating procedure". Business Cycle Dating Committee, National Bureau of Economic Research.
Hamilton, J.D. (1989). "A new approach to the economic analysis of nonstationary time series and the business cycle". Econometrica 57, 357–384.
Hamilton, J.D. (1994). Time Series Analysis. Princeton University Press, Princeton.
Hamilton, J.D. (2003). "Comment on a comparison of two business cycle dating methods". Journal of Economic Dynamics and Control 27 (9), 1691–1694.
Hamilton, J.D., Kim, D.H. (2002). "A re-examination of the predictability of economic activity using the yield spread". Journal of Money, Credit and Banking 34, 340–360.
Hamilton, J.D., Perez-Quiros, G. (1996). "What do the leading indicators lead?" The Journal of Business 69 (1), 27–49.
Hamilton, J.D., Raj, B. (2002). "New directions in business cycle research and financial analysis". Empirical Economics 27 (2), 149–162.
Harding, D., Pagan, A.R. (2002). "Dissecting the cycle: A methodological investigation". Journal of Monetary Economics 49 (2), 365–381.
Harding, D., Pagan, A.R. (2003). "A comparison of two business cycle dating methods". Journal of Economic Dynamics and Control 27 (9), 1681–1690.
Harding, D., Pagan, A.R. (2005). "Synchronisation of cycles". Journal of Econometrics. In press.
Härdle, W., Vieu, P. (1992). "Kernel regression smoothing of time series". Journal of Time Series Analysis 13, 209–232.

Hendry, D.F. (1995). Dynamic Econometrics. Oxford University Press, Oxford.

Hendry, D.F., Krolzig, H.-M. (1999). "Improving on 'data mining reconsidered' by K.D. Hoover and S. Perez". Econometrics Journal 2, 167–191.

Hendry, D.F., Krolzig, H.-M. (2001). "Computer automation of general-to-specific model selection procedures". Journal of Economic Dynamics and Control 25 (6–7), 831–866.

Hodrick, R.J., Prescott, E.C. (1997). "Postwar US business cycles: An empirical investigation". Journal of Money, Credit and Banking 29, 1–16.

Hornick, K., Stinchcombe, M., White, H. (1989). "Multilayed feedforward networks are universal approximators". Neural Networks 2, 359–366.

Johansen, S. (1988). "Statistical analysis of cointegration vectors". Journal of Economic Dynamics and Control 12, 231–254.

Kapetanios, G., Marcellino, M. (2003). "A Comparison of Estimation Methods for Dynamic Factor Models of Large Dimensions". Working Paper No. 489, Queen Mary, University of London, Department of Economics.

Keilis-Borok, V., Stock, J.H., Soloviev, A., Mikhalev, P. (2000). "Pre-recession pattern of six economic indicators in the USA". Journal of Forecasting 19, 65–80.

Kim, C.-J. (1994). "Dynamic linear models with Markov switching". Journal of Econometrics 60, 1–22.

Kim, C.-J., Murray, C. (2002). "Permanent and transitory components of recessions". Empirical Economics 27, 163–183.

Kim, C.-J., Nelson, C.R. (1998). "Business cycle turning points, a new coincident index, and tests of duration dependence based on a dynamic factor model with regime switching". The Review of Economics and Statistics 80, 188–201.

Kim, M.-J., Yoo, J.-S. (1995). "New index of coincident indicators: A multivariate Markov switching factor model approach". Journal of Monetary Economics 36, 607–630.

Kling, J.L. (1987). "Predicting the turning points of business and economic time series". Journal of Business 60 (2), 201–238.

Koch, P.D., Rasche, R.H. (1988). "An examination of the commerce department leading-indicator approach". Journal of Business and Economic Statistics 6 (2), 167–187.

Koopmans, T.C. (1947). "Measurement without theory". Review of Economics and Statistics 29, 161–179.

Kozicki, S. (1997). "Predicting real growth and inflation with the yield spread". Federal Reserve Bank of Kansas City Economic Review, Fourth Quarter.

Krolzig, H.-M. (1997). Markov Switching Vector Autoregressions. Modelling, Statistical Inference and Application to Business Cycle Analysis. Springer, Berlin.

Krolzig, H.-M. (2004). "Predicting Markov-switching vector autoregressive processes". Journal of Forecasting. In press.

Krolzig, H.-M., Marcellino, M., Mizon (2002). "A Markov-switching vector equilibrium correction model of the UK labour market". Empirical Economics 27 (2), 233–254.

Lahiri, K., Moore, G.H. (1991). Leading Economic Indicators: New Approaches and Forecasting Records. Cambridge University Press, Cambridge.

Lahiri, K., Wang, J. (1994). "Predicting cyclical turning points with leading index in a Markov switching model". Journal of Forecasting 13, 245–263.

Layton, A.P. (1996). "Dating and predicting phase changes in the US business cycle". International Journal of Forecasting 12, 417–428.

Layton, A.P. (1998). "A further test of the leading indicators on the probability of US business cycle phase shifts". International Journal of Forecasting 14, 63–70.

Layton, A.P., Katsuura, M. (2001). "Comparison of regime switching, probit and logit models in dating and forecasting US business cycles". International Journal of Forecasting 17, 403–417.

Li, D.T., Dorfman, J.H. (1996). "Predicting turning points through the integration of multiple models". Journal of Business and Economic Statistics 14 (4), 421–428.

Lieli, R.P. (2004). "A flexible framework for predicting binary variables". Job Market Paper, UC San Diego.

Lindgren, G. (1978). "Markov regime models for mixed distributions and switching regressions". Scandinavian Journal of Statistics 5, 81–91.

Lubrano, M. (1995). "Bayesian tests for co-integration in the case of structural breaks: An application to the analysis of wages moderation in France". G.R.E.Q.A.M. 95a19, Universite Aix-Marseille III.

Lütkepohl, H. (1987). Forecasting Aggregated Vector ARMA Processes. Springer-Verlag, Berlin.

Lütkepohl, H. (2006). "Forecasting with VARMA models". In: Elliott, G., Granger, C.W.J., Timmermann, A. (Eds.), Handbook of Economic Forecasting. Elsevier, Amsterdam. Chapter 6 in this volume.

Marcellino, M. (2003). "Forecasting EMU macroeconomic variables". International Journal of Forecasting 20, 359–372.

Marcellino, M. (2004). "Forecast pooling for short time series of macroeconomic variables". Oxford Bulletin of Economics and Statistics 66, 91–112.

Marcellino, M., Mizon, G. (2004). Progressive Modelling: Encompassing and Hypothesis Testing. Oxford University Press, Oxford. In press.

Marcellino, M., Stock, J.H., Watson, M.W. (2003). "Macroeconomic forecasting in the Euro area: Country specific versus Euro wide information". European Economic Review 47, 1–18.

Marcellino, M., Stock, J.H., Watson, M.W. (2005). "A comparison of direct and iterated AR methods for forecasting macroeconomic series h-steps ahead". CEPR WP 4976. Journal of Econometrics. In press.

McCulloch, R., Tsay, R. (1994). "Statistical analysis of economic time series via Markov switching models". Journal of Times Series Analysis 15, 523–539.

McGuckin, R.H., Ozyildirim, A., Zarnowitz, V. (2003). "A more timely and useful index of leading indicators". Working Paper.

McNees, S.K. (1991). "Forecasting cyclical turning points: The record in the past three recessions". In: Lahiri, K., Moore, G.H. (Eds.), Leading Economic Indicators: New Approaches and Forecasting Records. Cambridge University Press, Cambridge.

Mitchell, W., Burns, A.F. (1938). "Statistical indicators of cyclical revivals". NBER, New York. Reprinted in: Moore, G.H. (Ed.), Business Cycle Indicators. Princeton University Press, Princeton, 1961. Chapter 6.

Moneta, F. (2003). "Does the yield spread predict recessions in the Euro area?" ECB Working Paper No. 294.

Moore, G.H. (1983). Business Cycles, Inflation and Forecasting, second ed. Ballinger Publishing, Cambridge, MA.

Moore, G.H., Shiskin, J. (1967). "Indicators of business expansions and contractions". NBER Occasional Paper No. 103.

Moore, G.H., Zarnowitz, V. (1986). "The development and role of NBER's business cycle chronologies". In: Gordon, R.J. (Ed.), The American Business Cycle: Continuity and Change. University of Chicago Press, Chicago, pp. 735–779.

Neftci, S.N. (1982). "Optimal prediction of cyclical downturns". Journal of Economic Dynamics and Control 4, 225–241.

Niemira, M.P., Klein, P.A. (1994). Forecasting Financial and Economic Cycles. Wiley.

Osborn, D., Sensier, M., Simpson, P.W. (2001). "Forecasting UK industrial production over the business cycle". Journal of Forecasting 20 (6), 405–424.

Otrok, C., Whiteman, C.H. (1998). "Bayesian leading indicators: Measuring and predicting economic conditions in Iowa". International Economic Review 39, 997–1014.

Pagan, A. (2005). "Some econometric analysis of constructed binary time series". CAMA WP 7/2005.

Patton, A., Timmermann, A. (2003). "Properties of optimal forecasts". CEPR DP4037.

Phillips, K.R. (1998–1999). "The composite index of leading economic indicators: A comparison of approaches". Journal of Economic and Social Measurement 25 (3–4), 141–162.

Potter, S.M. (1995). "A nonlinear approach to US GNP". Journal of Applied Econometrics 10, 109–125.

Proietti, T., Moauro, F. (2004). "Dynamic factor analysis with nonlinear temporal aggregation constraints". Econometrics 0401003, Economics Working Paper Archive at WUSTL.

Qi, M. (2001). "Predicting US recessions with leading indicators via neural network models". International Journal of Forecasting 17, 383–401.

Raj, B. (2002). "Asymmetries of business cycles: The Markov-switching approach". In: Ullah, A., Wan, A., Chaturvedi, A. (Eds.), Handbook of Applied Econometrics and Statistical Inference. Marcel Dekker, pp. 687–710. Chapter 31.

Ravn, M.O., Sola, M. (1999). "Business cycle dynamics: Predicting transitions with macrovariables". In: Rothman, P. (Ed.), Nonlinear Time Series Analysis of Economic and Financial Data, vol. 1. Kluwer Academic, pp. 231–265. Chapter 12.
Sargent, T.J., Sims, C.A. (1977). "Business cycle modeling without pretending to have too much a priori economic theory". In: Sims, C., et al. (Eds.), New Methods in Business Cycle Research. Federal Reserve Bank of Minneapolis.
Shepard, N. (1994). "Partial non-Gaussian state space". Biometrika 81, 115–131.
Sichel, D.E. (1994). "Inventories and the three phases of the business cycle". Journal of Business and Economic Statistics 12 (July), 269–277.
Simpson, P.W., Osborn, D.R., Sensier, M. (2001). "Modelling business cycle movements in the UK economy". Economica 68 (270), 243–267.
Sims, C.A. (1989). "Comment on Stock and Watson (1989)". In: Blanchard, O., Fischer, S. (Eds.), NBER Macroeconomics Annual. MIT Press, Cambridge, MA, pp. 395–397.
Sims, C.A., Stock, J.H., Watson, M.W. (1990). "Inference in linear time series models with some unit roots". Econometrica 58, 113–144.
Stock, J.H., Watson, M.W. (1989). "New indexes of coincident and leading economic indicators". In: Blanchard, O., Fischer, S. (Eds.), NBER Macroeconomics Annual. MIT Press, Cambridge, MA, pp. 351–394.
Stock, J.H., Watson, M.W. (1991). "A probability model of the coincident indicators". In: Lahiri, K., Moore, G.H. (Eds.), Leading Economic Indicators: New Approaches and Forecasting Records. Cambridge University Press, Cambridge, UK.
Stock, J.H., Watson, M.W. (1992). "A procedure for predicting recessions with leading indicators: Econometric issues and recent experience". NBER Working Paper Series No. 4014.
Stock, J.H., Watson, M.W. (1993). "A procedure for predicting recessions with leading indicators: Econometric issues and recent experience". In: Stock, J.H., Watson, M.W. (Eds.), Business Cycles, Indicators, and Forecasting. The University of Chicago Press, Chicago, pp. 95–153.
Stock, J.H., Watson, M.W. (1999a). "Business cycle fluctuations in US macroeconomic time series". In: Taylor, J.B., Woodford, M. (Eds.), Handbook of Macroeconomics, vol. IA. North-Holland, Amsterdam.
Stock, J.H., Watson, M.W. (1999b). "A comparison of linear and nonlinear univariate models for forecasting macroeconomic time series". In: Engle, R.F., White, H. (Eds.), Cointegration, Causality, and Forecasting: Festschrift in Honour of Clive W.J. Granger. Oxford University Press.
Stock, J.H., Watson, M.W. (2002a). "Macroeconomic forecasting using diffusion indexes". Journal of Business and Economic Statistics 20, 147–162.
Stock, J.H., Watson, M.W. (2002b). "Forecasting using principal components from a large number of predictors". Journal of the American Statistical Association 97, 1167–1179.
Stock, J.H., Watson, M.W. (2003a). "Forecasting output and inflation: The role of asset prices". Journal of Economic Literature 41 (3), 788–829.
Stock, J.H., Watson, M.W. (2003b). "How did the leading indicator forecasts perform during the 2001 recession". Federal Reserve Bank of Richmond Economic Quarterly 89, 71–90.
Stock, J.H., Watson, M.W. (2006). "Forecasting with large datasets". In: Elliott, G., Granger, C.W.J., Timmermann, A. (Eds.), Handbook of Economic Forecasting. Elsevier, Amsterdam. Chapter 10 in this volume.
Swanson, N.R., Ghysels, E., Callan, M. (1998). "A multivariate time series analysis of the data revision process for industrial production and the composite leading indicator". In: Engle, R.F., White, H. (Eds.), Cointegration, Causality, and Forecasting: Festschrift in Honour of Clive W.J. Granger. Oxford University Press, Oxford, pp. 47–75.
Swanson, N.R., White, H. (1997). "A model selection approach to real-time macroeconomic forecasting using linear models and artificial neural networks". The Review of Economics and Statistics 79, 540–550.
Teräsvirta, T. (2006). "Forecasting with nonlinear models". In: Elliott, G., Granger, C.W.J., Timmermann, A. (Eds.), Handbook of Economic Forecasting. Elsevier, Amsterdam. Chapter 8 in this volume.
Timmermann, A.G. (2006). "Forecast combination". In: Elliott, G., Granger, C.W.J., Timmermann, A. (Eds.), Handbook of Economic Forecasting. Elsevier, Amsterdam. Chapter 4 in this volume.
Vaccara, B.N., Zarnowitz, V. (1978). "How good are the leading indicators?" In: Proceedings of the Business and Economic Statistics Section. American Statistical Association, pp. 41–51.

van Dijk, D., Teräsvirta, T., Franses, P.H. (2002). "Smooth transition autoregressive models – A survey of recent developments". Econometric Reviews 21, 1–47.

Walsh, C.E. (2003). Monetary Theory and Policy, second ed. The MIT Press.

Watson, M. (1991). "Using econometric models to predict recessions". Economic Perspectives, Research Periodical of the Chicago Federal Reserve Bank, September/October.

Wecker, W.E. (1979). "Predicting the turning points of a time series". Journal of Business 52 (1), 35–50.

West, K.D. (1996). "Asymptotic inference about predictive ability". Econometrica 64, 1067–1084.

West, K. (2006). "Forecast evaluation". In: Elliott, G., Granger, C.W.J., Timmermann, A. (Eds.), Handbook of Economic Forecasting. Elsevier, Amsterdam. Chapter 3 in this volume.

Zarnowitz, V., Boschan, C. (1975a). "Cyclical indicators: An evaluation and new leading indexes". Business Condition Digest May 1975. Reprinted in: US Department of Commerce. (1977). Handbook of Cyclical Indicators, pp. 170–183.

Zarnowitz, V., Boschan, C. (1975b). "New composite indexes of coincident and lagging indicators". Business Condition Digest November 1975. Reprinted in: US Department of Commerce. (1977). Handbook of Cyclical Indicators, pp. 185–198.

Zarnowitz, V., Braun, P. (1990). "Major macroeconomic variables and leading indicators: Some estimates of their interrelations, 1886–1982". In: Klein, P.A. (Ed.), Analyzing Business Cycles: Essays Honoring Geoffrey H. Moore. Armonk, New York, pp. 177–205.

Zellner, A., Hong, C. (1991). "Bayesian methods for forecasting turning points in economic time series: Sensitivity of forecasts to asymmetry of loss structures". In: Lahiri, K., Moore, G.H. (Eds.), Leading Economic Indicators: New Approaches and Forecasting Records. Cambridge University Press, Cambridge, UK, pp. 129–140.

Zellner, A., Hong, C., Gulati, G.M. (1990). "Turning points in economic time series loss structures and Bayesian forecasting". In: Geisser, S., Hodges, J., Press, S.J.A., Zellner, A. (Eds.), Bayesian and Likelihood Methods in Statistics and Econometrics: Essays in Honor of George A. Barnard. North-Holland, Amsterdam, pp. 371–393.

Zellner, A., Hong, C., Min, C. (1991). "Forecasting turning points in international output growth rates using Bayesian exponentially weighted autoregression, time-varying parameter and pooling techniques". Journal of Econometrics 49, 275–304.

Zellner, A., Min, C. (1993). "Bayesian and non-Bayesian methods for combining models and forecasts with applications to forecasting international growth rates". Journal of Econometrics 56, 89–118.

Zellner, A., Rossi, P.E. (1984). "Bayesian analysis of dichotomous quantal response models". Journal of Econometrics 25, 365–394.

Zha, T. (1998). "A dynamic multivariate model for use in formulating policy". Economic Review 83 (First Quarter 1998).

Proof of Raw Subject Index