Journal of Economic and Social Measurement 25 (1998/1999) 183–233
IOS Press

An application of the Stock/Watson index methodology to the Massachusetts economy

Alan Clayton-Matthews[a,*] and James H. Stock[b]
[a] *University of Massachusetts, Boston, MA, USA*
[b] *John F. Kennedy School of Government, Harvard University, Cambridge, MA, USA*

The Stock/Watson index methodology is applied to the Massachusetts economy to estimate coincident and leading indexes for the state. A coincident index, calibrated to trend with gross state product, is estimated as a dynamic single factor, multiple indicator model, using the Kalman filter and smoother on a set of coincident indicators. The leading index is a six-month ahead forecast of the coincident index, based on a regression on recent growth in the coincident index and a set of leading indicators. Filtering of noisy data and model selection in the context of a short historical span of data are two issues common to index construction at the state and regional levels that the authors address.

Keywords: Coincident index, leading index, Kalman filter, dynamic single factor model, predictive least squares, Stock/Watson model

1. Introduction

Research on national composite indexes continues at a rapid pace. The official coincident index is re-estimated with a refinement of a frontier methodology or a new methodology at almost every journal cycle. The same cannot be said for research applied to state or regional levels of the economy. This is understandable. Data are more readily available at the national level, and working with a common set of data allows colleagues across the country to cooperate with each other through the research and publishing nexus in a manner that is much more efficient than if each focused on different regional indexes.

However, state and regional economy-watchers are perhaps more in need of applied work on indexes than are their national counterparts. If regional economies cycled in phase with, and with the same intensity as, the national economy, then national indexes would suffice. However, this is not the case. Moreover, regional economists are presented with several data problems. At the state and regional levels, there are fewer series available than at the national level. These series often have a much shorter historical record, are usually much noisier due to smaller survey sizes, and are often of poorer quality due to fewer data gathering resources.

*Corresponding author: Alan Clayton-Matthews, Public Policy Program, UMass Boston, 100 Morrissey Boulevard, Boston, MA 02125-3393, USA. Tel.: +1 617 287 6945; Fax: +1 617 287 6949; E-mail: acm@mediaone.net.

Finally, there is no timely measure of state and regional product that can be used to guide and evaluate the construction of economic indexes. Gross State Product (GSP), for example, is available with a lag of approximately a year and a half, and then is available only at an annual frequency. The next best proxy for product, state personal income and its components, is available at a quarterly frequency, with a lag of up to seven months. Often the most recent quarterly data point is subjected to substantial revision. The upshot is that the major burden of state and regional cyclical analysis falls on a single indicator, establishment employment. Although the measure is available at a monthly frequency and in a timely manner, the indicator has problems, most importantly its sampling design, which can lead to substantial errors in real time, especially at cyclical turning points. Also, it is merely a count of jobs, and does not capture changes in the intensity of work (e.g., weekly hours), or changes in aggregate productivity.

This paper applies one of the relatively new model based methods of national index construction, developed and refined by Stock and Watson in several papers, to the Massachusetts economy. The Stock/Watson method consists of constructing a coincident index as the estimated factor of a dynamic single-factor, multiple indicator model, using the Kalman filter. The leading index is then formed as a forecast of the six-month ahead growth rate of the coincident index using a set of leading indicators and recent growth in the coincident index. The method is especially applicable to state and regional economies because it does not require the existence of an observable series that represents the state of the economy, such as a regional equivalent to quarterly Gross Domestic Product (GDP). Rather, it assumes that the underlying state is unobservable, which is exactly the situation at hand.

There is nothing new or path-breaking about the methodology presented in this paper. Instead, the value offered here is the application of the Stock/Watson method in a state context where the data problems are more severe than at the national level, yet typical of those faced at the state and regional levels. Compared to the national level, there are fewer indicators. In Stock and Watson [17], their "short" list of national leading indicators consisted of fifty-five series; our short list for Massachusetts consists of only nine series. The length of historical availability is much shorter. Coincident indicator data are only available back to 1978 for Massachusetts, which means that there are effectively only one and a half substantial business cycles available for estimating the parameters of the models. (The double-dip national recession of the early eighties was little more than a pause in the expansion that began in the second half of the seventies.) The set of leading indicators in our short list begins even later, in 1981. Also, many of the data series are contaminated with very high-frequency noise, i.e., large month-to-month fluctuations, which make signal extraction methods like the Kalman filter absolutely necessary. We hope that our approach to these problems addressed in this paper will serve as a useful guide to analysts in other states.

The paper is organized as follows. Section 2 lays out the model used to formulate and estimate the coincident and leading indexes. Section 3 describes the pre-estimation filtering used to eliminate high-frequency noise in several series. Sections 4 and 5 document the estimation of the coincident and leading indexes respectively. Section 6 presents a brief analysis of several leading index models. Section 7 offers some conclusions.

2. The model

The basic model used here for the Massachusetts coincident and leading index was developed and applied at the national level by Stock and Watson [15–17], and has been used to construct coincident indexes in a total of at least eleven states by Crone [4], Clayton-Matthews, Kodrzycki and Swaine [3], Tsao [19], and Orr, Rich, and Rosen [11]. Crone [5] has also constructed coincident economic indexes for each of the 48 contiguous states. Leading indexes using the Stock/Watson approach have been constructed in a total of at least four states by Crone and Babyak [6], and Orr, Rich, and Rosen [12].

The form of the model used here is:

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

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

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

$$\Delta c_t = \begin{matrix} a_1 + b_1\Delta s_{t|T} & \text{for } t \in T_1 \\ a_2 + b_2\Delta s_{t|T} & \text{for } t \in T_2 \end{matrix}, \tag{4}$$

$$CEI_t = A\exp(c_t), \quad \text{where } A = 100/\exp\left(c_{\text{July 1987}}\right) \tag{5}$$

$$f_t(6) = \alpha + \lambda_c(L)\Delta c_t + \lambda_y(L)\Delta y_t + v_t, \quad \text{where } f_t(6) = c_{t+6} - c_t. \tag{6}$$

$$LEI_t = 200\hat{f}_t(6) \tag{7}$$

The observed data series consist of x and y, with the former being a vector of coincident indicators, and the latter begin a vector of leading indicators. For clarity of presentation, x and y are assumed to be measured in log form and stationary when differenced.[1] The logarithm of the state of the economy at time t is represented

[1] If the x series are cointegrated of order one with a single common trend, then the model in equations (1)–(3) can be estimated in levels, with all the accompanying benefits of superconsistency. In Clayton-Matthews, Kodrzycki, and Swaine, one of the indexes for Massachusetts involved six x series, three of which were cointegrated of order one with a single common trend. The model in (1)–(3) was modified accordingly to accommodate the presence of both stationary and nonstationary x variables. However, the resultant estimated index was not wholly satisfactory, as the cyclical frequencies of the output index were dominated by a single one of the nonstationary input series. In this paper, the input series x do not have a single common trend, and so all the data are differenced to achieve stationarity.

by the scalar s_t. The disturbances (ε_t, η_t) are assumed to be serially uncorrelated with a zero mean and a diagonal covariance matrix Σ. ε_t is a vector disturbance with the same dimension as x. η_t is a scalar disturbance. L is the lag operator, i.e., $L^k x_t = x_{t-k}$. The lag polynomial matrix $D(L)$ is assumed to be diagonal, so that the μ's in different equations in (1) are contemporaneously and serially uncorrelated with one another. v_t is the disturbance of the leading index model.

The first three equations comprise a dynamic single factor, multiple indicator model, first proposed by Sargent and Sims [14], where the growth in the unobserved state, Δs, represents the common comovements in the growth of the coincident indicators, Δx; and the autoregressive disturbances, μ, form the idiosyncratic portion of each observed coincident series. The first equation, except for its dynamics, is like the multiple indicator model familiar from factor analysis. Here, Δs represents a single, common, unobserved, factor, while the vector Δx constitutes the observed indicators, and the vector μ the unique components. If the x indicators move in tandem with the economy, then their common component s has the natural interpretation as the current state of the economy, and can serve as a composite coincident index. In the multiple indicator model, the observed variables are by convention expressed as deviations from their respective means. This is an identifying restriction which constrains the constant β to be zero. This identifying strategy is also used here. Each of the coincident series forming x is first-differenced (remember x is already logged) and normalized by subtracting its mean difference and dividing by the standard deviation of its differences.

The dynamic that distinguishes this dynamic factor model from non-dynamic factor models is given by the third equation. This equation, known as the "state equation" (Δs is also known as the "state variable"), the "transition equation", or the "law of motion", models the growth in the state of the economy as a stationary autoregressive process. Stationarity is assured because Δx is stationary by construction. Also, because Δx is constructed to have a mean of zero, the parameter δ is identified to be zero. The state of the economy is assumed to evolve by the accumulation of shocks, η. Each shock affects current period growth directly, and future growth through the autoregressive process, though with damped effect. The autoregressive structure allows above- and below trend growth rates to persist for some time, generating business cycles.

Growth in the state of the economy is revealed by the observable indicators through the set of equations in (1), also known as the "measurement" equations. Each coincident indicator can be expected to grow contemporaneously with the state of the economy, can lead the state, can lag the state, or can exhibit a more complex relationship to the state, depending on the set of factor loadings given by $\gamma(L)$ in equation (1).

The unique factors, or idiosyncratic components, μ, in the measurement equation are stationary, mean zero, autoregressive stochastic processes, as stated by equation (2). Again, stationarity is assured because Δx and Δs are stationary. These idiosyncratic factors are assumed to be uncorrelated with one another, which is another way of stipulating that there is only a single common factor among the indicators. If instead a subset of the μ's were correlated with one another, the correct way to model this would be to include another dynamic factor in addition to Δs that would have non-zero factor loadings in the corresponding measurement equations. However, then the interpretation of s as the state of the economy would be problematic. Which of the two common factors, if either, represents the true state? A simple specification test, described below, is used to verify this model assumption.

In estimating the coincident index, we follow the well-established procedure of Stock and Watson. In brief, each of the coincident series forming x is first-differenced (remember x is already logged) and normalized by subtracting its mean difference and dividing by the standard deviation of its differences. This identifies $\beta = 0$ and $\delta = 0$ for purposes of estimation. The scale of the $\gamma(L)$ coefficients is fixed by setting the variance of η to unity, and the timing of the coincident index is fixed by setting all but one of the elements of $\gamma(L)$ to zero in at least one of the equations in (1).

Quasi maximum likelihood[2] estimation of the parameters of the system in (1)–(3) and estimation of the filtered state is accomplished by representing the system in (1)–(3) in state space form and using the Kalman filter.[3] There are two ways to form the state space system. One way is to treat equation (1) as the measurement equation, and equations (2) and (3) as the state equation, with the state vector including both μ and Δs. The second way is to concentrate equation (2) out of the system by multiplying both sides of equation (1) by $D(L)$. Then the transformed equation (1) becomes the measurement equation, and equation (3) the state equation. The second method, which we used, has the advantage of a smaller dimension of the state vector, which decreases computation time significantly relative to the first method. Transformation of the system of equations (1)–(3) into state space form is described in detail by Stock and Watson [16].

Several outputs of the estimation procedure are used in the analysis. Once the parameter estimates are obtained, three versions of state vector are produced from the Kalman filter and smoother: (i) $\Delta s_{t|t-1}$, (ii) $\Delta s_{t|t}$, and (iii) $\Delta s_{t|T}$. Each version is an estimate of the state conditional on a different set of information, and may be called the "prediction", "filtered", and "smoothed" estimates respectively. In the first version, the state in each period is estimated with information available through the prior period; in the second version, with information available through the current period; and in the third version, with the entire set of information. The first version

[2] That is, we maximize a likelihood function that assumes the vector of disturbances ε and η are jointly normally distributed with mean zero and diagonal covariance matrix Σ. If their distribution is not normal, then the estimates are still consistent, although not asymptotically efficient.

[3] There are many references for the state space model and the Kalman filter. Anderson and Moore [1] and Hamilton [8] are two favorites of the authors. Software that estimates the parameters of state-space models is becoming more widely available. The software used by the authors to estimate the model in (1)–(3) is available upon request [2].

is used to form the "one-step ahead" prediction errors, $\hat{\varepsilon}_{t|t-1} = \Delta x_t - \Delta x_{t|t-1}$, used to calculate the likelihood conditional on the parameter estimates, and also used in the specification test below. These prediction errors are the fitted residuals from the measurement equation system (1) and (2), where the estimates $\Delta s_{t|t-1}$ are used in place of the latent Δs. The third version, which is "best" in that it uses the most information, is used to form the coincident index. To each version corresponds a linear (non-recursive) filter that, when applied to the observed indicators, yields the corresponding version of the state estimates. These filters may be labeled the Kalman "prediction", "filter", and "smoother" filters respectively. The latter two are commonly referred to by the shorter terms "Kalman filter" and "Kalman smoother" respectively. The first two are one-sided, and the third, the Kalman smoother, is two-sided.[4] The filters are as long as the data series, although the weights rapidly approach zero as one moves away from the current period, so in practice they could be used to directly calculate the state estimates from the observable (normalized) indicators each month.

The key to understanding the difference between the Stock/Watson methodology and the conventional US Bureau of Economic Analysis/The Conference Board (BEA/TCB) composite methodology is in the filter. The BEA/TCB methodology (Green and Beckman [7]) is equivalent to a filter with non-zero weights for the current period only. The resultant index smoothes across the indicators only. Furthermore, the BEA/TCB "equal weighting" scheme, although reasonable, is arbitrary.[5] The Kalman filter, in contrast, smoothes across both indicators and time, and since it is estimated by maximum likelihood, it is, in a sense, optimal. The resultant index value for a given month is based on more information than in the BEA/TCB method. Also, since noise is more effectively and completely filtered out, the output index is smoother, and the signal is clearer.

Because of the normalization of the input indicator series, and the identification strategy used, the output of the Kalman filter, $\Delta s_{t|T}$, is driftless with a unit-variance shock. It must be de-normalized before it can be integrated and de-logged to form the coincident economic index. This de-normalization is given in equation (4) where the a's and b's are constants derived so that the coincident economic index, CEI, has a trend rate of growth, and variance around the trend, equal to that of real Massachusetts Gross State Product. A good fit with GSP required two sets of denormalization constants, one for the period prior to 1988 (T_1), and the other for 1988 and later (T_2). The decision to use GSP as a target for denormilization is somewhat arbitrary. Gross State Product was chosen because Gross Domestic Product (GDP) is the standard measure for economic growth at the national level. At the state level, GSP is not

[4] The two-sided filter associated with $\Delta s_{t|T}$ of course changes as the ends of the data series are reached, and is identical to the one-sided filter associated with $\Delta s_{t|t}$ at the end of the data, i.e., when $t = T$.

[5] The BEA/TCB weights are inversely proportional to the mean absolute deviation in symmetrical growth rates. They call this scheme "equal weighting" because, if each series of growth rates were normalized by dividing by its mean absolute deviation, then the weights would indeed be equal.

widely used because it is released only annually, and moreover, with a lag of about 18 months after the end of the year. Thus the choice of GSP as a de-normalization target can help fill a gap for regional economic analysts.

The de-normalization is actually performed in two steps. The first step produces an index that is analogous to the BEA/TCB methodology in that the resultant index is a weighted average of the observed coincident indicators, only instead of the "equal weighting" assumption of the BEA/TCB, the weights are given by the Kalman filter. The de-normalized growth rates from this first step are simply a linear transformation of the estimated state from the Kalman filter, $\Delta s_{t|T}$. This step is described in Stock and Watson [16]. Details are given in Appendix A. The second step simply adjusts the linear transformation from the first step so as to meet the goals of fitting the first two moments of GSP. Alternatively, the first step could be skipped, and the index could be directly calibrated to GSP. The details of this procedure are given in Appendix B. Finally, the index is set to 100 in July 1987 by dividing by the appropriate constant. The choice of July 1987 is purely arbitrary and is of no consequence. The date happens to be near the peak of the state's last expansion.

Equation (6) models the six-month ahead growth rate of the coincident index as a linear function of current and past values of the coincident index and leading indicators. The disturbance term v_t is not white noise, but is an ARMA process with 5th order moving-average components. This characteristic follows from the construction of the dependent variable as the six-month ahead growth in the CEI, which contains six future realizations of the η disturbance from equation (3), each of which is assumed to be independent of the right-hand side variables in (6). We estimate the coefficients of (6) with ordinary least squares. The OLS estimator gives consistent coefficient estimates under the assumptions of the model; however, the OLS standard error estimates are biased. Orr, Rich, and Rosen [12] use an estimator proposed by Newey and West [10] to calculate unbiased estimates of the standard errors. Since hypothesis tests on the estimated coefficients are only of minor interest to us, for convenience we simply use the OLS standard errors, and we don't rely on the OLS standard errors for model selection purposes.

Equations (3) and (6) are seemingly incompatible, since they both can be used to generate forecasts of the state. Equation (6) uses leading indicators as additional information. Presumably, this use of additional information would enable more accurate forecasts. However, the coefficients in (6) are not estimated efficiently, because they are estimated conditionally on the estimates of the state from the Kalman filter. Efficient estimation would involve replacing the state equation (3) by equation (6), in which case all relevant coefficients would be estimated jointly. [6] However, there are practical problems to joint estimation of equations (1), (2), and (6). This would substantially increase both the dimension of the transition matrix in the state equation and the number of parameters to be estimated by numerical

[6] The calibration and definitional equations (4) and (5) have no bearing on the analysis here.

optimization, resulting in a many-fold increase in the computational time to estimate a single specification of the model. Thus, there is a trade-off between parameter estimation efficiency and specification search. We felt that the loss in efficiency was small compared to the benefits of the more extensive specification search we would be able to perform if (6) were estimated separately from the model in (1) through (3).

Once the final specification of the model is chosen and its parameters are estimated, then the state, $\Delta s_{t|T}$, can be continually updated in real time using the Kalman filter and smoother. This provides updates of the coincident index, and updates of the six-month ahead forecast of the coincident index, $\hat{f}_t(6)$. The leading index can be presented in two ways. One is to form it as the estimate of the coincident index six months hence, i.e.,

$$CEI_{t+6|t} = A \exp\left\{c_t + \hat{f}_t(6)\right\}. \tag{8}$$

This traditional way focuses on the level of the index. We prefer to present the leading index instead as the expected six-month ahead growth in the CEI expressed as an annual rate, as in equation (7). We feel this focus on the growth rate has a more natural and thus easier interpretation. Because the coincident index is calibrated to grow at the same rate as real GSP, the growth-rate form of the leading index is directly comparable to other meaningful growth rates such as national GDP, historical GSP, employment, etc. Also, the monthly change in the leading index given by (7) consists solely of the change in the six-month outlook, while the monthly change in the leading index given by (8) also includes the change in the coincident index from the prior month. Economy-watchers who want an indication of what will happen in the future would naturally want to net out this latter part anyway, and so would prefer the form given in equation (7).

3. Pre-Estimation filtering

An issue that deserves attention is the noisiness of the coincident and leading indicators. The importance of this issue is best understood by noting the influence of noise on the traditional composite indexes. The BEA/TCB methodology weights each indicator series in inverse proportion to its monthly percent change. The reliance on the first difference magnifies the importance of very high frequencies, including high-frequency noise. However, the important information content of indicators for constructing indexes occurs at lower frequencies, corresponding to periods of several months to a year or two. Thus, there is a potential problem in the BEA/TCB weighting methodology. An indicator that is superior in its information content at the important frequencies may get unduly penalized if it has higher month-to-month noise than other series. This situation is even more problematic at the state level, where high-frequency noise is typically much more of a problem, due, for example, to smaller survey sizes than at the national level.

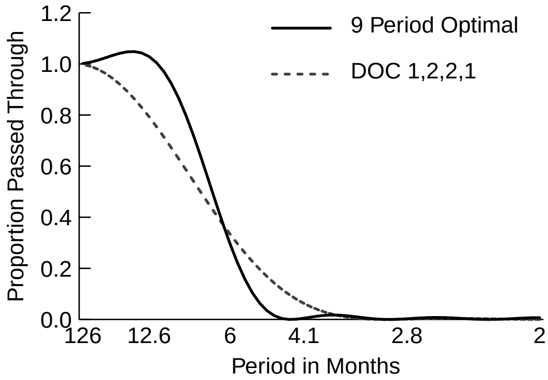

Fig. 1. Transfer functions, by period.

One solution to this problem is to pre-smooth those indicators that are contaminated with high-frequency noise. Commonly, a weighted moving-average, i.e., non-recursive filter, is used.[7] The term “filter” is apt, since each set of moving-average weights in the time domain corresponds to a transfer function in the frequency domain that specifies how each frequency of the input series is factored into the output series. Frequencies with low factors near zero are filtered out of the input series, while frequencies with factors near one are passed through to the output series. The use and specification of filters designed to pass through the frequencies critical for index formation while filtering out those dominated by noise is an important aspect of index construction. This is especially true for traditional composite methods that do not otherwise exploit time-series dynamics.

In the case of the Stock/Watson model that employs the Kalman filter, pre-estimation filtering is not theoretically necessary, because the procedure produces a set of moving-average weights that filter the series over time. Once again, however, pre-estimation filtering is preferred in practice because otherwise the lag polynomials $\gamma(L)$ and $D(L)$ in equations (1) and (2) would be of possibly much higher orders, making specification search and estimation much too time-consuming to be practical.

We use two filters for pre-estimation smoothing. One is an optimal band pass 9-period centered moving average filter, designed to filter out frequencies corresponding to periods of less than five months, while letting frequencies corresponding to periods of 10 months or more to pass. The filter is $\frac{1}{3.04}(-0.216L^{-4}, -0.007L^{-3}, 0.412L^{-2}, 0.831L^{-1}, 1, 0.831L, 0.412L^{2}, -0.007L^{3}, -0.216L^{4})$. Its transfer function is graphed in Fig. 1. This filter is used on the two tax-based coincident indicators, which are very noisy at high frequencies. Since we use this as a centered moving-average, we require a four-month

[7] An accessible text explaining the concepts used here, including the design of filters, is Hamming [9].

forecast of the input indicator for updating the index in real time. This is accomplished by the forecast of an ARIMA(4,1,0) model on the logarithm of the indicator. The information loss due to forecasting is not serious, as the filter weights on the second through fourth months of the forecast comprise less than 20 percent of the total sum of the filter weights (in terms of absolute values).

The other filter we employ is a four-period moving-average that was popular with the US Department of Commerce in the past, $\frac{1}{6}(1, 2L, 2L^2, 1L^3)$. Its transfer function is also graphed in Fig. 1. The filter passes less of the lower frequencies and filters less of the higher frequencies than the optimal band-pass filter, but has a narrower window of only four months. We use it as a one-sided filter for several of the leading indicators.

4. The concident index

4.1. The coincident indicators

The candidate coincident indicator series were chosen for the following criteria: comovement with regional economic activity, high frequency, timeliness of availability, length of historical record, reliability, low noise, and robustness to revisions. It is much more challenging to find a set of indicators that meet these criteria at the regional level than at the national level, and the set used here is far from ideal, resulting in compromises in the formation of the coincident index. The set of indicators considered included – all at the state level of observation – establishment employment, withholding taxes, sales taxes, the unemployment rate, household employment, and weekly hours in manufacturing. Household employment was dropped because of difficulties in estimating the model when it was included, i.e., the variance of its idiosyncratic component often vanished. Weekly hours in manufacturing was dropped because of its tiny impact on the estimated coincident series, and because it appeared to be a promising leading indicator.

Establishment employment is perhaps the most closely watched single economic indicator at the regional level. It is monthly, released about six weeks after the activity it measures, is closely related to in-state production, is relatively smooth, and is available for the entire post WWII span of history. Its major shortcoming is due to its sampling properties. It consists of a panel of approximately 10,000 firms. The voluntary panel is continually replenished with new firms, but with a considerable lag. The result is that it often fails to signal turning points in real time, so that there are sometimes major revisions that shift the employment turning points by several months, and occasionally, by a year or more.

The two state tax indicators have some appealing characteristics. They are monthly, and are released within two weeks after the end of the month for which they are collected. In the case of withholding taxes, the lag between the data release and the economic activity on which they are based is short. Large companies that remit withholding taxes weekly within days of the end of the pay period account for about 85 percent of total withholding taxes. The lag in sales tax collections is roughly a month longer. Prior to 1998, approximately 70 percent of monthly tax receipts were for sales made in the current month. However, due to a tax law change, sales tax receipts for a given month now reflect sales made in the prior month. Perhaps the best characteristic of these tax series is that they are not samples, by virtue of including all taxpayers.

There are a few problems with the tax data. First, the taxes do not reflect all the economic activity upon which they are based. The sales tax measure used here excludes food; most clothing; excise taxes on fuels, alcohol, cigarettes, and automobiles; and real estate taxes. (Sales taxes on motor vehicles are used as a leading indicator.) Some sales tax liabilities are underreported or not reported. Withholding taxes do not include self-employed earnings. Second, there have been several law changes affecting both taxes. In order to account for these, the series have been transformed into tax bases by adjusting the revenue amounts and dividing by the appropriate tax rates.[8] However, for a couple of periods of time, especially between early 1976 to late 1977 for sales taxes, and 1970 through 1972 for withholding taxes, tax revenues exhibit unexplainable fluctuations. Third, the data are noisy, partly due to large month-to-month fluctuations commonly referred to as "spill" within the Department of Revenue.[9] In order to reduce the noise in the tax series, the monthly data are smoothed by the filter described in the previous section. This markedly improves the data on the smoothness criterion. Fourth, changes in tax collection procedures, and perhaps accounting, have resulted in a changing pattern of month-to-month fluctuations in tax collections, with a clear break occurring after the end of fiscal year 1978. Fortunately, most – but not all – of the difference in noise patterns is at frequencies high enough to be eliminated by the filter.

The tax base indicators are deflated by the US consumer price index for all urban consumers (CPI-U). The pre- and post-filtered constant dollar tax bases are illustrated in Figs 2 and 3. The tax base series appear to be valid, except for the periods mentioned above. The constant-dollar withholding tax base is compared to BEA's quarterly state wage and salary disbursements, deflated by the US CPI-U, in Fig. 4. The sales tax base is compared to the now discontinued Census Bureau's retail sales for Massachusetts, deflated by the US CPI-U, in Fig. 5.

The unemployment rate has perhaps the weakest theoretical connection to the state of the economy, at least to the extent that the "state of the economy" represents current production. However, the historical time series profile of the unemployment rate fits the other coincident series well over several cycles, and so is included in the set. Recently the BLS developed a model-based approach for reporting the

[8] The construction of the tax bases is described in Clayton-Matthews, Kodrzycki, and Swaine [3].

[9] Spill occurs when tax revenues that arrive during a window bridging two months are randomly distributed between the two months.

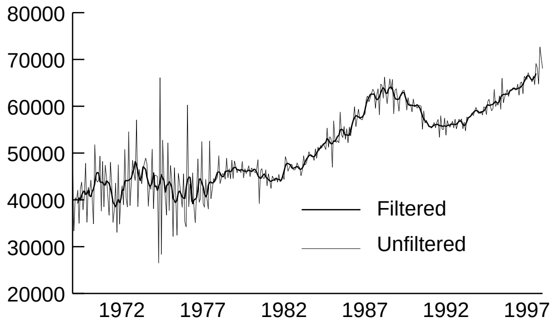

Fig. 2. Real withholding tax base, pre- and post-filtered.

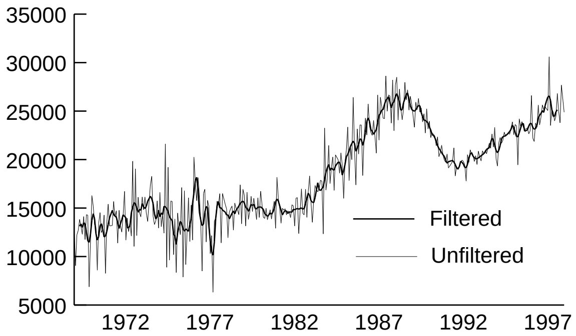

Fig. 3. Real sales tax base, pre- and post-filtered.

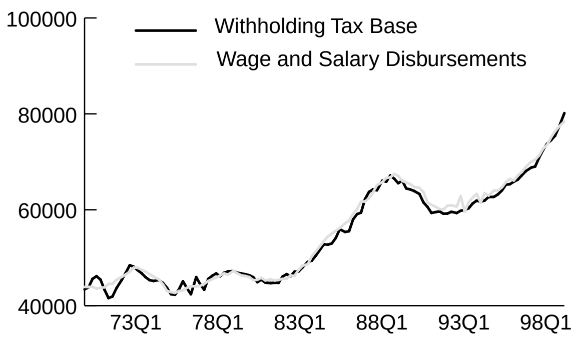

Fig. 4. BEA wage and salary disbursements vs. withholding tax, in millions of 1982–1984 dollars.

unemployment rate that has resulted in a series that is much smoother in the post model regime after 1977 than before. Unlike the other coincident indicators, this series was not logged before differencing. Also, even though differencing is not

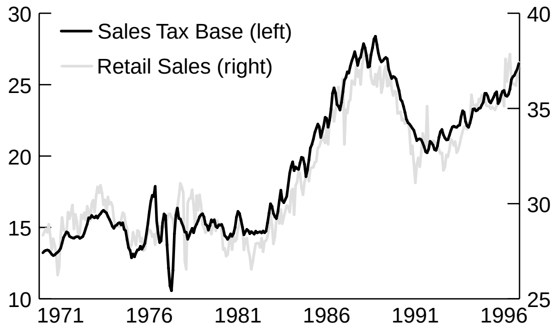

Fig. 5. Census Bureau retail sales for Massachusetts vs. sales tax base, in billions of 1982–1984 dollars.

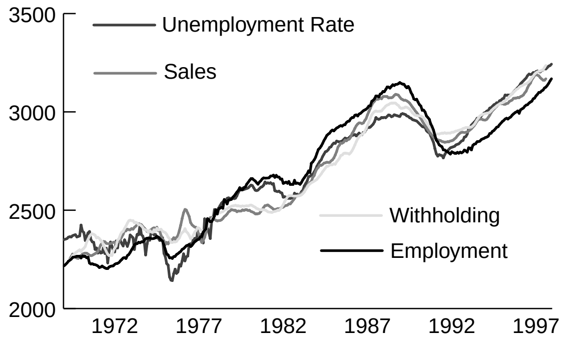

Fig. 6. Coincident indicators, retrended and scaled to employment.

needed to achieve stationarity (the unemployment rate must lie between 0 and 100), differencing makes the unemployment rate more like the other series in terms of its cyclical profile. This can be inferred from graphs of the undifferenced coincident indicators in Fig. 6.

The four coincident series are illustrated together in Fig. 6. In order to remove differences due to trend, intercept, and scale, which have no effect on the coincident index estimation, each series in this figure is represented as the fitted value of a regression of establishment employment on the series, a time trend, and a constant. The essential information for each of these four series, including its transformation for estimation purposes in the model (1)–(3), is summarized in Table 1.

4.2. Estimation of the coincident index

The time period for estimation was limited to the period beginning in January 1978. It is possible to extend the time period back to 1968 or 1969, but aside from historical interest, the pre 1978 data add little additional information, for the following reasons. First, some pre 1978 periods of the tax data are of questionable

Table 1
Coincident series

Series	Name	Transformation	Source	Availability	Notes
Establishment Employment	Emp	Log First Difference	BLS, DET	Third week following month	
Withholding Tax Base	With	Log First Difference	DOR	Second week following month	Formed into tax base; deflated by US CPI-U
Sales Tax Base	Sales	Log First Difference	DOR	Second week following month	Formed into tax base; deflated by US CPI-U
Unemployment Rate	UR	First Difference	BLS, DET	Third week following month	

BLS: Bureau of Labor Statistics.
DOR: Massachusetts Department of Revenue.
DET: Massachusetts Division of Employment and Training.

quality and should not be used in parameter estimation. Second, the BLS estimates of state unemployment rates are "model-based" from 1978 onward. This would require estimating an additional set of parameters for its idiosyncratic component prior to 1978. The main cost of excluding pre-1978 data is foregoing information from the steep mid 1970's recession and recovery.

We used the following criteria to select the specification for the coincident model in equations (1)–(3):

1. The estimated index should be consistent with knowledge of Massachusetts' recent economic history, and with other data series that should cycle with the economy, such as gross state product and personal income.
2. The estimated filter should be reasonable. This means that weights on the indicator series should be concentrated around a zero lag, and each indicator should contribute substantially to the output index.
3. The one-step ahead forecast errors of the model should pass a specification test for whiteness. The test assesses whether the white noise components in equation (2) are predictable. The null hypothesis is that the forecast error, $\hat{\varepsilon}_{t|t-1}$, for each indicator is uncorrelated with past values of: itself; forecast errors of other indicators; and each indicator. The test is implemented by a series of regressions. In each regression, the dependent variable is one of the one-step ahead forecast error series, and the independent variables consist of a constant and six lags of one of the forecast error series or indicator series. The null hypothesis is that these six coefficients are zero, and the null is tested by forming the appropriate F-statistic. In all, for k indicators, the tests require $2k^2$ regressions and F-statistics.
4. Smoothness of the estimated state, $\Delta s_{t|T}$. Presumably, the state of the economy is smooth, and this should be reflected in the output of the Kalman filter.
5. Characteristics of the index, such as its profile, timing of turning points, and contribution of coincident indicators, should be robust to minor changes in its specification.

Table 2
Estimated coincident index model

Parameter	Employment	Withholding Tax	Sales Tax	Unemployment Rate
γ_0		0.2351	0.1912	
		(0.0408)	(0.0343)	
γ_1	0.3819			−0.3552
	(0.0575)			(0.0560)
d_1	−0.1195	0.4153	0.3763	
	(0.0949)	(0.0686)	(0.0671)	
d_2	−0.0836	−0.1095	−0.1655	
	(0.0877)	(0.0754)	(0.0702)	
d_3	0.3088	−0.1506	−0.2377	
	(0.0927)	(0.0738)	(0.0698)	
d_4		−0.1799	−0.1825	
		(0.0676)	(0.0661)	
σ	0.6673	0.7564	0.7585	0.7561
	(0.0497)	(0.0394)	(0.0369)	(0.0424)
	Autoregressive coefficients for the state equation			
ϕ_1	0.1200			
	(0.0619)			
ϕ_2	0.7707			
	(0.0739)			
Logarithm of the likelihood		−1184.78		
Period of estimation		January 1978–June 1998		
Estimated standard errors are in parentheses.				

The final specification and parameter estimates are presented in Table 2. The parameters of the model in equations (1) through (3) can be expressed as follows, where G is the number of coincident indicators.

$$\gamma(L) = [\gamma_1(L), \gamma_2(L), \cdots, \gamma_G(L)]',$$

where $\gamma_g(L) = \gamma_{g0} + \gamma_{g1}L + \gamma_{g2}L^2 + \cdots.$

$$D(L) = diag[d_1(L), d_2(L), \cdots, d_G(L)]',$$

where $d_g(L) = 1 - d_{g1}L - d_{g2}L^2 - \cdots.$

$$\phi(L) = 1 - \phi_1 L - \phi_2 L^2 - \cdots.$$

$$\Sigma = diag\lfloor \sigma_1^2, \sigma_2^2, \cdots, \sigma_G^2, 1 \rfloor$$

The state of the economy enters the measurement equations for withholding taxes and sales taxes contemporaneously, and the employment and unemployment equations with a lag of one month. These relative timings were not imposed a priori, but rather empirically, as the result of trying many different specifications. The lag of employment and the unemployment rate is reasonable. Wages, on which withholding taxes are based, usually adjusts first in response to changes in labor input, as employers typically adjust hours first, before hiring or laying off workers. Consumption, on

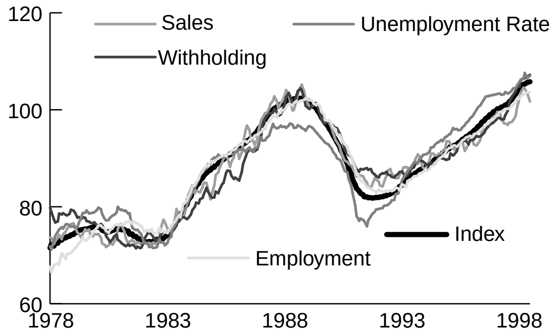

Fig. 7. Conincident indicators, retrended and scaled to the coincident index.

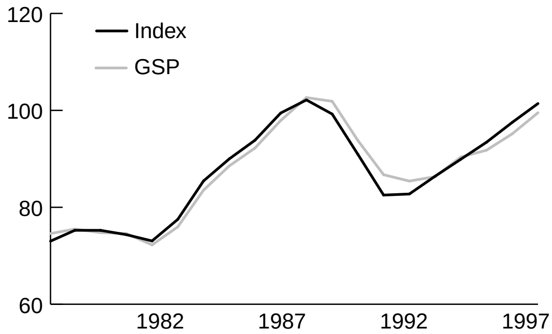

Fig. 8. Real gross state product, retrended to the coincident index, vs. the coincident index.

which sales tax receipts are based, might reasonably either lead, lag, or be contemporaneously related to wages. To the extent households make purchases after receipt of income, consumption would lag income, and to the extent that households make purchases in anticipation of income, consumption would lead income. The specification that both wages and consumption are contemporaneous is therefore reasonable. The second order autoregressive structure of the estimated state equation (3) implies a high degree of persistence in the economy. It suggests that the growth rate of the economy is the accumulation of a series of shocks whose effects diminish slowly over time. The specification search that led to the estimated autoregressive structures of the idiosyncratic portions of the indicators in equation (2) were aided by the white noise specification test and by testing down from fourth-order specifications. The high-order specifications for the tax series are not surprising given that they are the output from a 9-period filter. The "best" specification indicated no autoregression in the idiosyncratic portion of the unemployment rate.

The index is in accord with economists' knowledge of the history of business cycles in Massachusetts, and is in agreement with other state-level coincident indictors, as illustrated in Figs 7, 8, and 9. For purposes of comparison, the index used in

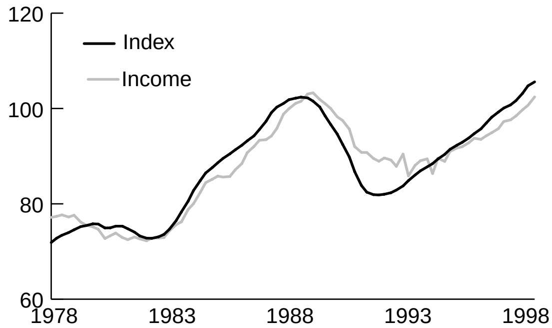

Fig. 9. Real state personal income, retrended to the coincident index, vs. the coincident index.

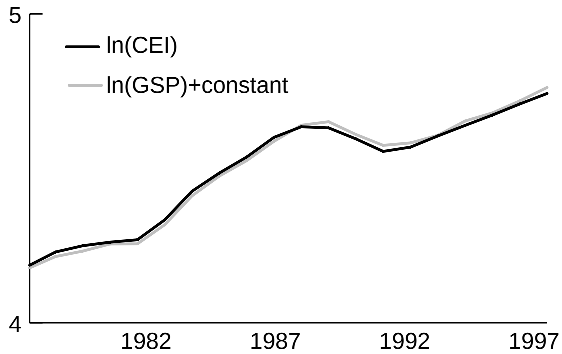

Fig. 10. Logarithm of the concident index vs. the logarithm of real gross state product plus a constant.

these figures is those resulting from the BEA/TCB-type denormalization, as given in equation (A1) of Appendix A. We want to assess the reasonableness of the index before it is rescaled to fit GSP. In Fig. 7, the index is graphed against its component indicators. In order to remove differences due to trend, intercept, and scale, each of the coincident indicators in this figure is the fitted value of a regression of the index on the corresponding indicator, a time trend, and a constant. This figure illustrates the use of the Stock/Watson model as a signal extraction device. The index provides a smooth and clear signal of the state of the economy from the several noisy input indicators. In Figs 8 and 9, the index is compared to GSP and personal income respectively, with GSP and personal income rescaled by the same regression technique.

Figure 10 shows the close agreement of the CEI with GSP after the trend and variance adjustments, as described in Appendix B and given by equations (4) and (5).[10] The choice of two regimes for rescaling the CEI provides a good fit with GSP

[10]The values calculated for the parameters in equation (4) are: $a_1 = 0.00290730$, $b_1 = 1/473.427$,

Table 3
The Kalman filter: Scoring coefficients or dynamic multipliers

Lag	Employment	Withholding Tax	Sales Tax	Unemployment Rate
0	0.1566	0.4788	0.3994	−0.1057
1	0.3835	−0.1855	−0.1287	−0.2606
2	0.1190	0.2205	0.2032	−0.0416
3	0.1117	0.0118	0.0563	−0.0870
4	−0.0521	0.1576	0.1438	−0.0282
5	0.0380	0.0203	0.0354	−0.0299
6	−0.0088	0.0510	0.0478	−0.0143
7	0.0091	0.0118	0.0164	−0.0111
Sum of Remaining Lags	0.0040	0.0393	0.0417	−0.0184
Sum of Coefficients	0.7611	0.8056	0.8153	−0.5967
Share of Impact (%)	25.6	27.0	27.4	20.0

Fig. 11. The Kalman filter.

over the 1978–1997 period, and so its growth may serve as an estimate of the growth in GSP.

The index agrees with the perception that: the minicomputer boom of the late 1970's was robust; the two early 1980's recessions were very mild in Massachusetts; the "Miracle" years of the 1980's exhibited rapid growth; the recession that ended the decade occurred earlier in Massachusetts than the nation, and also was much more severe; and was followed by a long expansion in the 1990's that continues to the present (as of early 2000). The dates of the turning points of the Massachusetts economy given by the CEI, and the corresponding dates for the national economy determined by the National Bureau of Economic Research, are listed in Table 9.

The Kalman filter and smoother are displayed in Figs 11 and 12, and in Tables 3 and 4. The Kalman smoother produces a set of filters for $\Delta s_{t|i}, t \leqslant i \leqslant T$. The one-sided filter for $\Delta s_{t|t}$ is used to update the CEI for the current period. This filter is presented

$a_2 = 0.00194050,\ b_2 = 1/637.332.$

Table 4
The Kalman smoother: Scoring coefficients or dynamic multipliers

Lag	Employment	Withholding Tax	Sales Tax	Unemployment Rate
−8	−0.0013	0.0080	0.0082	−0.0080
−7	0.0136	0.0081	0.0099	−0.0145
−6	−0.0154	0.0229	0.0224	−0.0176
−5	0.0578	0.0123	0.0190	−0.0398
−4	−0.0734	0.0676	0.0623	−0.0323
−3	0.1933	0.0082	0.0308	−0.1183
−2	0.0593	0.1130	0.1057	−0.0371
−1	0.5162	−0.0780	−0.0384	−0.3358
0	0.0593	0.2780	0.2391	−0.0371
1	0.1933	−0.0780	−0.0384	−0.1183
2	−0.0734	0.1130	0.1057	−0.0323
3	0.0578	0.0082	0.0308	−0.0398
4	−0.0154	0.0676	0.0623	−0.0176
5	0.0136	0.0123	0.0190	−0.0145
6	−0.0013	0.0229	0.0224	−0.0080
7	0.0033	0.0081	0.0099	−0.0056
8	0.0005	0.0080	0.0082	−0.0034
Sum of Remaining Lags and Leads	0.0080	0.0240	0.0264	−0.0216
Sum of Coefficients	0.9957	0.6261	0.7055	−0.9017
Share of Impact (%)	30.8	19.4	21.8	27.9

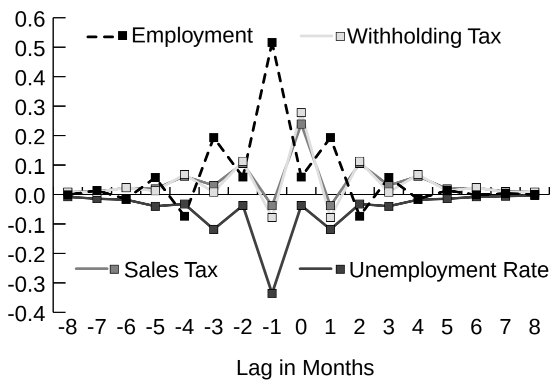

Fig. 12. The Kalman smoother.

in Fig. 11. For $t \ll i$ the filter is the two-sided symmetric smoother presented in Fig. 12. These filters exhibit the desired characteristics. They are concentrated around the zero lag and reflect that each indicator has substantial influence on the index. Summing the filter weights over time for each indicator gives the impact on the state, and therefore the CEI, of a unit change in the corresponding indicator. These show that the four indicators share roughly equally in their impact on the estimated CEI. For example, the cumulative weights from the one-sided filter vary from a 20% share for the unemployment rate, to a 27% share for sales taxes.

Table 5
Whiteness specification test p-values

Regressors: A Constant and Six Lags of:		Dependent Variable: One-Step Ahead Forecast Error of Corresponding Indicator			
		Employment	Withholding Taxes	Sales Taxes	Unemployment Rate
One-Step Ahead Forecast Error	Employment	0.140	0.384	0.731	0.963
	Withholding Taxes	0.101	0.100	0.120	0.378
	Sales Taxes	0.119	0.406	0.583	0.507
	Unemployment Rate	0.913	0.135	0.673	0.018
Transformed Indicators (See Table 1 for the Transformation)	Employment	0.101	0.817	0.720	0.621
	Withholding Taxes	0.590	0.580	0.645	0.050
	Sales Taxes	0.232	0.213	0.856	0.246
	Unemployment Rate	0.004	0.150	0.071	0.066

The entries in the table are the p-values from a regression of the one-step ahead forecast error on a constant and six lags of the indicated regressor. The p-values correspond to the F-test of the null hypothesis that the coefficients (except for the constant) are all zero.

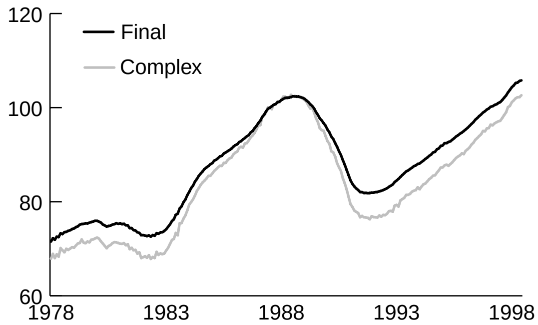

Fig. 13. Indexes of the final and more complex specifications..

The results of the whiteness specification test are given in Table 5. Only two out of the 32 p-values are less than 0.05, so the specification satisfies this test reasonably well. If the 32 tests were independent (they're not), then the probability of two or more p-values below 0.05 would be 0.48.

The resulting index appears to be fairly robust from specification to specification. As an example, the index was compared to the following more complex specification in which it is nested. In this more complex specification: 1) the state entered the measurement equations for withholding taxes, sales taxes, and the unemployment rate coincidentally and with lags from one to three months; 2) the autoregressive order of the state equation was set to four; and 3) the autoregressive orders of the measurement equations were each set to four. The two indexes (using the BEA/TCB-type denormalization) are displayed in Fig. 13 and their smoothed states, $\Delta s_{t|T}$, in Fig. 14. The two specifications result in indexes with similar profiles and turning points. The p-value for the null hypothesis that the 16 additional parameters of the

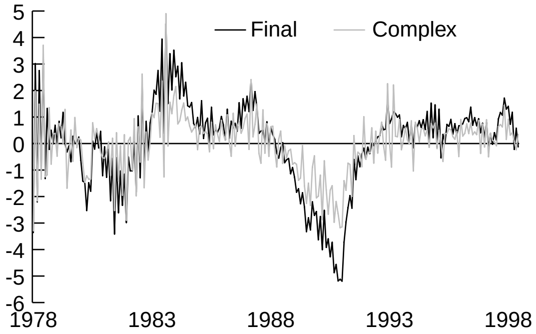

Fig. 14. States of the final and more complex specifications.

more complex model are all zero is 0.025. However, the state of the simpler model is smoother than the more complex one, and so we prefer it.

5. The leading index

5.1. The leading indicators

A short list of nine leading monthly indicator candidates was selected from a longer list of monthly series on the basis of timelines, expected performance, and historical availability. These nine are consumer confidence in New England, the spread between the 10-year Treasury Bond and 90-day Treasury Bill yields, help wanted advertising for Boston, the Bloomberg Massachusetts stock index, initial unemployment claims, housing permits, construction employment, motor vehicle sales tax collections, and weekly hours in manufacturing. All series except for interest rates, consumer confidence, and help wanted advertising are state-level data. The Bloomberg Massachusetts stock index was only available from December 1994 to the present. Prior values for the index back to January 1968 were constructed by Tsao [18]. Some examples of series that were not included include exports, which are quarterly and not available in a timely manner; and several indices whose historical record is too short, includeing the Case Shiller/Weiss housing price index, BankBoston's Instant Reading index (like a purchasing manager's index), and the Associated Industries of Massachusetts' business confidence index.

The nine candidates are graphed against the current economic index in Figs 15 through 19. The transformations performed on each of the leading indicators in estimating the model in (6) are given in Table 6. All but two of the indicators were smoothed with the filter $\frac{1}{6}(1, 2L, 2L^2, 1L^3)$, and all series were differenced except the interest rate spread. The choice of whether a series was filtered or differenced was determined by which of the four alternatives had the best fit in a simple version

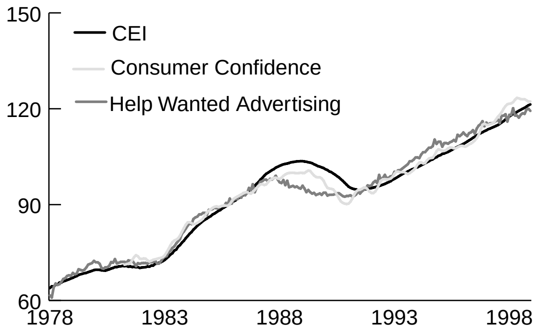

Fig. 15. Consumer confidence and help wanted leading indicators, retrended to the CEI, vs. the CEI.

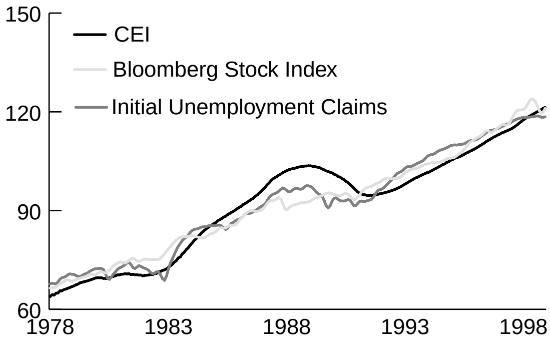

Fig. 16. Bloomberg stock index and initial unemployment claims leading indicators, retrended to the CEI, vs. the CEI.

of the model in (6) in which the explanatory variables included the growth in the coincident index and the single leading indicator under consideration. [11]

5.2. Estimation of the leading index

A search was conducted over several thousand specifications of equation (6). The specifications can be enumerated as follows. All specifications include the logged, first-differenced coincident index entering with lags of zero, one, and two, i.e., the LEI enters all forms of equation (6) in the following manner: $\lambda_c(L)\Delta c_t = \lambda_{c0}\Delta c_t + \lambda_{c1}\Delta c_{t-1} + \lambda_{c2}\Delta c_{t-2}$.[12] This set of terms is highly significant in all

[11] In the notation of the next section, the model specifications used were I(#)0, where # indicates which of the nine leading indicators was being evaluated.

[12] Note that $\Delta c_t = \ln\left(\frac{CEI_t}{CEI_{t-1}}\right)$, i.e., the growth rate of the CEI.

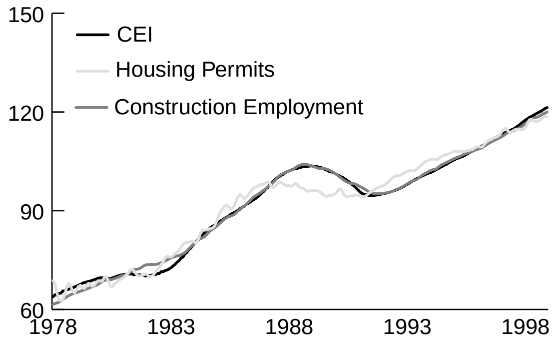

Fig. 17. Housing permits and construction employment leading indicators, retrended to the CEI, vs. the CEI.

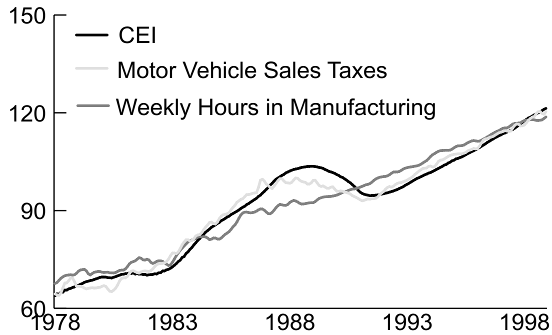

Fig. 18. Motor vehicle sales taxes and weekly hours in manufacturing leading indicators, retrended to the CEI, vs. the CEI.

specifications because the autoregressive structure of the state in equation (3) indicates a strong degree of persistence. In addition, all specifications include a constant.

The specifications are best described as belonging to one of four sets, I, II, III, and IV. Each set contains 1533 specifications, for a total of 6132 specifications in all. The 1533 specifications in each set were formed in the following manner. All 511 combinations of the nine leading indicator series were tried, where $511 = \binom{9}{1} + \binom{9}{2} + \cdots + \binom{9}{9}$. Three specifications were tried for each of the 511 combinations. In the first, each leading indicator was entered with a lag of zero months. In the second, each leading indicator was entered with lags of zero and one month. In the third, each was entered with lags of zero, one, and two.[13] Thus, there were $511 \times 3 = 1533$ specifications in each set.

[13]By lags of zero, one, and two months, we mean relative to the currently available information set. For example, if the current month of the coincident index is June, this corresponds to observations for

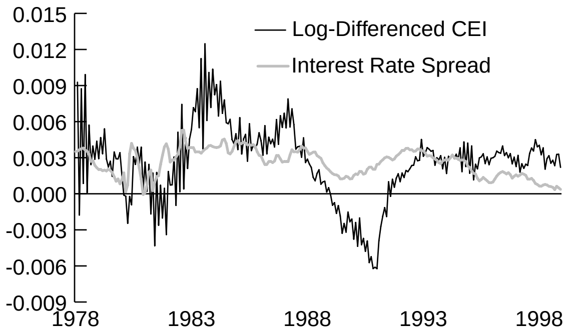

Fig. 19. Interest rate spread leading indicator, retrended to the CEI, vs. the log-differenced CEI.

In the first set, set I, the specifications include a constant, the three coincident indicator terms, and the set of leading indicator terms as described above.

At this point it is useful to develop a convenient notation to refer to a specific specification of equation (6). Numbers in parentheses shall indicate which combination of leading indicators is used in the model, where the numbering corresponds to the order in Table 6. The number after the parentheses indicates the lag structure. The Roman numeral preceding the parentheses indicates the set to which the specification belongs. For example, I(126)2 refers to the model in the first set in which consumer confidence, the interest rate spread, and housing permits enter, each with lags of zero, one, and two. The model that includes only the CEI, and does not include any of the candidate leading series, shall be called ()0.

In set II, symmetry between the implicit parameters for the coincident indicators and those for the leading indicators was imposed. Because $\Delta s_{t|T}$ is a linear combination of the coincident indicators, Δx, one could imagine using this linear combination and equation (4) to substitute for the Δc terms in equation (6), resulting in terms involving Δx instead. Then, because of the rescaling in (4), there would be two sets of parameters for each coincident indicator, one corresponding to the regime T_1, and the second corresponding to the regime T_2. These parameters would be related by the constant of proportionality b_1/b from (4). Thus, in set II, we specify two sets of $\lambda_y(L)$ parameters, one for T_1, the second for T_2, and constrained by the constant of proportionality b_1/b_2. This constraint is easy to implement for the OLS estimator and requires no additional parameters to estimate. This symmetry specification also requires separate constant parameters for the two regimes. The relationship between the two constants is known in terms of the previously determined parameters from (4) and the $\lambda_y(L)$ parameters from (6), so no new parameters need be estimated. However, not imposing the constraint on the constant involves

May, June, and July for the Bloomberg stock index and the interest rate spread; March, April, and May for help wanted and housing starts; and April, May, and June for all other leading indicators.

Table 6
Leading indicators

#	In Final Model	Series	Transformation	Filtered	Relative Lead
	*	Coincident Index	growth rate		0
1	*	Consumer Confidence for New England	first difference	*	0
2	*	U.S. Treasury Spread: 10-Year Minus 3-Month	none		1
3		Help Wanted Advertising, Boston	growth rate		−1
4	*	Bloomberg Stock Index for Massachusetts	growth rate	*	1
5	*	Initial Unemployment Claims	growth rate	*	0
6		Housing Permits	growth rate	*	−1
7	*	Construction Employment	growth rate	*	0
8	*	Motor Vehicle Sales Taxes	growth rate	*	0
9		Average Weekly Hours in Manufacturing	first difference	*	0

Unless otherwise indicated, the series is for Massachusetts.
The Bloomberg Stock Index and Motor Vehicle Sales taxes are deflated by the US Consumer Price Index for all Urban Consumers.
The growth rate is the the first difference of the logarithm.
The filter is the one-sided moving average filter (1, 2, 2, 1) given in the text.
A relative lead of 1 (one) indicates that the data are available a month prior to other data, and so enter the model with a lead of one month relative to other data; a relative lead of −1 (minus one) indicates that the data are available a month later than other data, and so enter the model with a lag of one month relative to other data.

estimating only one additional parameter, so we estimated separate unconstrained constants for the two regimes. The model specification for set II can be described in the following manner:

$$f_t(6) = \begin{matrix} \alpha + \alpha' + \lambda_c(L)\Delta c_t + \frac{b_1}{b_2}\lambda_y(L)\Delta y_t + v_t & \text{for } t \in T_1 \\ \alpha + \lambda_c(L)\Delta c_t + \lambda_y(L)\Delta y_t + v_t & \text{for } t \in T_2 \end{matrix}. \tag{9}$$

Each model in set II therefore involves the estimation of one additional parameter relative to the corresponding model from set I.

Set III involves the same specifications as in set I, but estimated according to a different methodology – one that has proved successful by out of sample tests in other applications. The idea is to estimate separate models, each with a single leading indicator, and then average them to yield a model with a given combination of leading indicators. For example, the coefficient estimates of the model III(126)2, which includes three of the leading indicators, is formed by a simple average of the coefficient estimates of the three models I(1)2, I(2)2, and I(6)2. In forming these averages, excluded leading indicators implicitly have a coefficient of zero.[14] Averaging the coefficients of the models is equivalent to averaging their predicted or fitted values.

Set IV applies the same estimation strategy used in set III to the specifications of set II. For example, the coefficient estimates of the model IV(126)2, is formed by a simple average of the coefficient estimates of the three models II(1)2, II(2)2, and II(6)2.

Each specification was ranked on the basis of several criteria. These included the following:

1. Predictive least squares (PLS) sum of squared residuals (lower is better). This criterion ranks specifications on the basis of the cumulative sum of squares of six-step ahead forecast errors, also known as recursive residuals. For the linear regression model $y_t = \boldsymbol{b}'\boldsymbol{x}_t + \varepsilon_t$, the criterion selects the regressor x that minimizes $PLS(\boldsymbol{x}) = \sum_{t=s}^{T}(y_t0 - \boldsymbol{b}'_{t-k}\boldsymbol{x}_t)^2$, where b_i is the least squares estimate based on $t \leqslant i$, and where s is great enough so that the estimated in sample residuals from the estimation of b_s have positive degrees of freedom.[15] We set $k = 6$ (rather than $k = 1$) because the dependent variable is the six-month ahead growth rate of the CEI. The PLS procedure we used reflected, as much as conveniently possible, real time updating of the coefficient estimates of b and the estimation of y_t. For example, in constructing the tax bases each month, which required applying a filter centered on the current month, appropriate forecasts were made, rather than using actual historical data. We did not, however, attempt to reproduce the actual, pre-revised, data that would have been available each month. Nor did we reestimate the parameters of the coincident index model (1)–(4) each month.
2. Lowest value of the Bayesian Information Criterion (BIC) (lower is better). The measure is $BIC = \ln\left(\frac{SSR}{T}\right) + K\left(\frac{\ln T}{T}\right)$, where SSR is the sum of squared residuals of the regression, T is the number of observations, and K is the number of parameters in the model. The BIC rewards fit but penalizes for complexity. Wei [20] showed that, for ergodic models, PLS and BIC are

[14] We also tried a variant of this method where an additional regression was run that only included the coincident index terms, and the regressions for each leading indicator did not include the coincident index terms. This variant did not perform as well, because the averaging of coefficients diluted the effect of the best predictor, the coincident index.

[15] The choice of s is arbitrary, but does affect the result. A straightforward standard is to select the minimum s possible. However, as Wei [20] notes, the early prediction errors in this case may be quite large and have too much influence on the result. We chose a value of s that was six years (72 observations) into the observation range, which allowed recursive residuals to be calculated over a period of time that spanned the last two turning points. yet allowed at least 40 degrees of freedom for the most parameterized model.

asymptotically equivalent, although in practice the two measures can yield quite different results.

3. Highest share of contribution due to the leading indicators (higher is better). The share of an indicator is defined to be the weighted sum of its coefficients from (6), divided by the weighted sum for all indicators, including the coincident index. The sign for initial unemployment claims is appropriately reversed. The weight for an indicator's coefficients is the standard deviation of the year-over-year change in the indicator. The purpose of this weighting is to standardize the coefficients for typical changes in the indicators at business cycle frequencies. Because of the high persistence of the state in equation (3), the coincident indicator consistently accounts for the largest share, which was well over half in all 6132 specifications. This criterion favors those specifications in which the leading indicators, as a group, have the most influence on the forecast of the coincident index.
4. Equality of shares of the leading indicators (more equal is better). Considering only the leading indicators, this measure is the average absolute deviation of the indicator share from the average share of all the leading indicators, where the shares are defined as above.[16]
5. The number of leading indicators (more is better).
6. The number of "wrong" signs on the leading indicators (fewer are better). The sign of an indicator is defined to be the sign of the sum of its coefficients when it enters with one or more lags. The sign is expected to be positive for all of the indicators except initial unemployment claims.

Interestingly, the rankings based on the PLS and the BIC criteria are different in several respects, according to some basic characteristics of the top 100 models according to each of the two criteria. First, only two models are in both top 100 rankings. Second, half the top 100 PLS models are from set I and half are from set II, while all expect one of the top 100 BIC models are from set II. The extra constant parameter for the regime dummy is almost always significant, but that doesn't necessarily lead to the best out of sample performance. The top models by the PLS criterion are more complex than those that score highest on the BIC. All the PLS models include two lags of the leading indicators, while only a few of the BIC models include any lags at all (except two lags of the CEI, which all models include). The top models by PLS also contain slightly more leading indicators on average, 4.96 versus 4.44 for the top models by the BIC. The leading indicators contribute a larger share, on average, of the variation in the index (criterion number three above) in the top PLS models than in the top BIC models, 29 percent versus 18 percent. This is not surprising given their relative complexity. However, there are more wrong signs in the top PLS models. Weekly hours in manufacturing have the wrong sign in all

[16]In calculating the average the divisor is the number of leading indicators in the specification less one-half to offset a bias in this measure in favor of specifications with few indicators.

66 of the top 100 PLS models in which they enter, and help wanted advertising has the wrong sign in all 41 of the top 100 PLS models in which they enter. Among the top 100 BIC models, 22 have weekly hours entering with the wrong sign (in only one model does it enters with the right sign), and 12 have construction employment entering with the wrong sign (9 have construction employment with the right sign). No other indicators enter any of these 198 models with the wrong signs. Three leading indicators occur frequently in both lists. Consumer confidence appears in all top 100 PLS models and 66 of the top 100 BIC models. The Bloomberg stock index appears in all top 100 BIC models and in 60 of the top PLS models. Motor vehicle sales taxes appear in all PLS models and 54 BIC models. No other indicators enter more than half the models of both lists. However, the interest rate spread enters 63 of the BIC models, and 30 of the PLS models; and weekly hours in manufacturing enters 66 of the PLS models, but with the wrong sign as previously mentioned.

None of the models from sets III or IV were in either of the top 100 lists, and therefore did not fit or predict as well as their multivariate counterparts.

These distinctly different rankings by the PLS and BIC are part of the reason for including the other criteria in model selection. However, an even more important consideration is the short time period over which the leading indicator models are fit. Given data availability, the estimation period is July 1981 through April 1998. Although this consists of 202 monthly observations, it includes essentially only one full cycle and part of another: the rapid expansion of the “Massachusetts Miracle” years, the steep and long recession that followed, and the long and steady expansion of the 1990’s. Models that either fit well or that perform well in out of sample forecasts over this short period of time are not guaranteed to perform better in the future than models that do not score as well on the PLS or BIC criteria.

The additional four selection criteria are based on the common sense idea that the more indicators contributing to the index, the better. Having a large set of indicators where none is dominant serves as insurance against a mistake in model selection due to peculiarities of the short historical period on which the models were estimated. Of course, the criterion that favors more indicators introduces the risk of including series that are not reliable leading indicators. Our way of guarding for this risk was to initially narrow down the list to a reasonable set of nine candidate indicators.

To aid us in applying these multidimensional criteria, we eliminated those models from consideration that contained wrong signs according to the last criterion, and weighted the other criteria equally, except for double weighting the PLS criterion. This served as guide to help narrow down the models into a manageable handful where application of the criteria by subjective judgment, i.e., expert opinion, led to choosing a “best” model.

6. Analysis of the leading index models

The fitted and actual six-month-ahead growth rates of the CEI for several leading index specifications are displayed in Figs 20 through 39, and a guide to these figures

Table 7
Guide to figures of selected models

Figure	Model #	Specification	Comments
20	1	I(1)0	consumer confidence and CEI only
21	2	I(2)0	interest rate spread and CEI only
22	3	I(3)0	help wanted advertising and CEI only
23	4	I(4)0	Bloomberg stock index and CEI only
24	5	I(5)0	initial unemployment claims and CEI only
25	6	I(6)0	housing permits and CEI only
26	7	I(7)0	construction employment and CEI only
27	8	I(8)0	motor vehicle sales taxes and CEI only
28	9	I(9)0	average weekly hours in manufacturing and CEI only
29	2960	II(124578)2	chosen model
30	2987	II(145678)2	scores high on evaluation criteria
31	2828	II(12458)2	scores high on evaluation criteria
32	2694	II(1248)2	scores high on evaluation criteria, ranked 52nd by PLS
33	1206	I(1689)2	ranked 1st by PLS, wrong sign on one indicator
34	1095	I(189)2	ranked 2nd by PLS, wrong sign on one indicator
35	1672	II(1248)0	ranked 1st by BIC
36	1580	II(124)0	ranked 2nd by BIC
37	1806	II(12458)0	ranked 3rd by BIC
38	511	I(123456789)0	all indicators included
39	3066	II(123456789)2	all indicators, most parameters

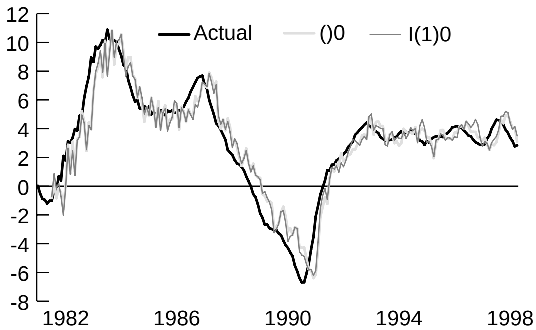

Fig. 20. Fitted LEI models vs. actual: Fitted models using only the lagged growth in CEI, ()0, and including consumer confidence, I(1)0.

is given in Table 7. Values are displayed as percentage growth at annual rates, accomplished by multiplying the actual and fitted growth rates by 200 as stated in equation (7). Since the CEI is calibrated to trend with GSP, the real time predictions of the models can be interpreted as the expected growth, at annual rates, of GSP over the next six months. The zero line separates expansions and contractions. When the leading index is negative, future output six months hence is expected to be less than in the present month. Conversely, when the index is positive, future output six months from now is expected to be greater than in the present month. The figures highlight the dual purposes of the index. One, as in traditional leading indexes, is to

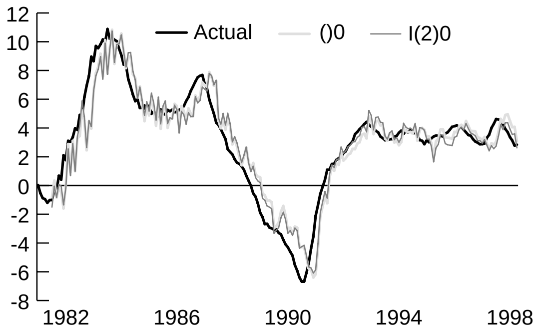

Fig. 21. Fitted LEI models vs. actual: Fitted models using only the lagged growth in CEI, ()0, and including the interest rate spread, I(2)0.

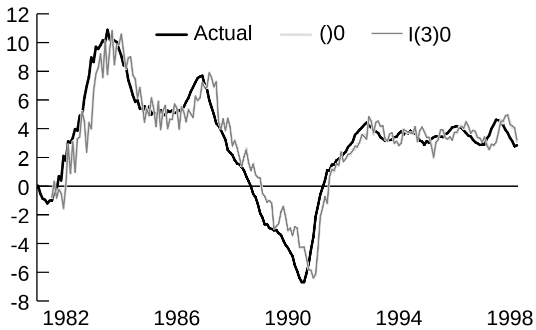

Fig. 22. Fitted LEI models vs. actual: Fitted models using only the lagged growth in CEI, ()0, and including help wanted advertising, I(3)0.

give a forewarning of turning points. The second is to predict the level of growth. If the model predicted growth of the CEI perfectly, its fitted values would lie right on the actual curve in the figure, and the model would also perfectly lead turning points. Note however, that perfect prediction of the CEI does not mean perfect prediction of the economy, as the CEI itself is only an estimate of the unobservable state of the economy.

Vertical deviations between the fitted and actual curves measure the model's error in predicting future changes in the rate of growth. Alternatively, the model's performance in predicting cyclical turning points may be measured by the horizontal displacement of the fitted curve from the actual. No horizontal displacement where the actual curve crosses the zero line indicates prefect prediction of a turning point. If the fitted curve lies to the right of the actual, but by less than by six months, the index is leading, but by less than six months. If the fitted curve lies to the left of the actual, the index is prematurely announcing turning points.

Each graph in Figs 20 through 39 includes two fitted models, one of them always

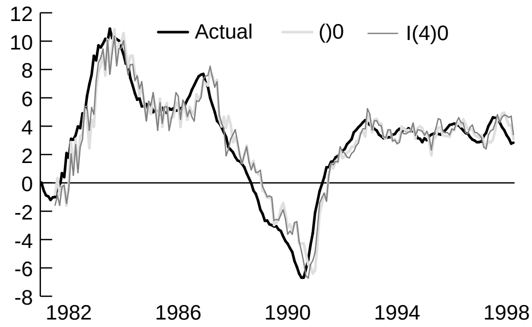

Fig. 23. Fitted LEI models vs. actual: Fitted models using only the lagged growth in CEI, ()0, and including the Bloomberg stock index for Massachusetts, I(4)0.

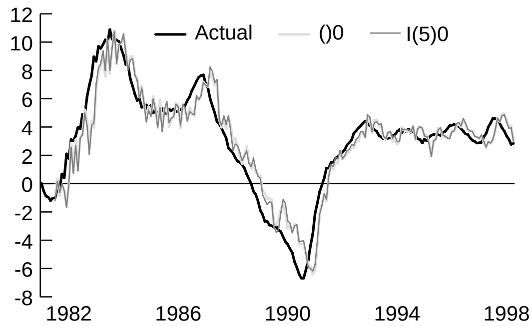

Fig. 24. Fitted LEI models vs. actual: Fitted models using only the lagged growth in CEI, ()0, and including initial unemployment claims, I(5)0.

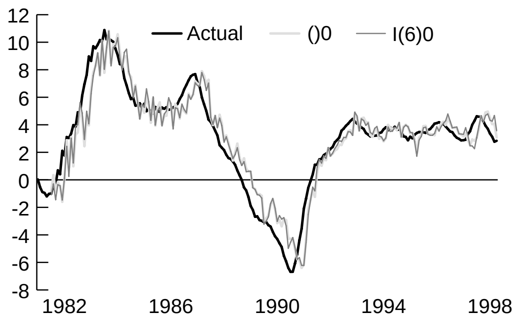

Fig. 25. Fitted LEI models vs. actual: Fitted models using only the lagged growth in CEI, ()0, and including housing permits, I(6)0.

being the specification that includes only the current and lagged values of the growth in the CEI, i.e., specification ()0 – call this the minimal specification. Because of the

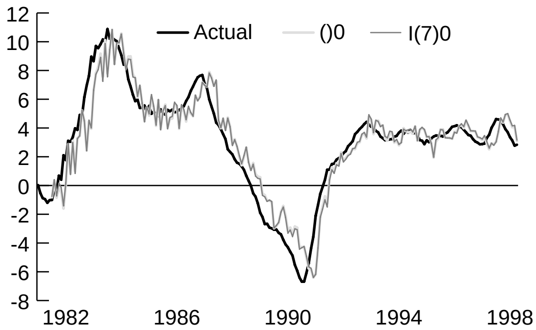

Fig. 26. Fitted LEI models vs. actual: Fitted models using only the lagged growth in CEI, ()0, and including construction employment, I(7)0.

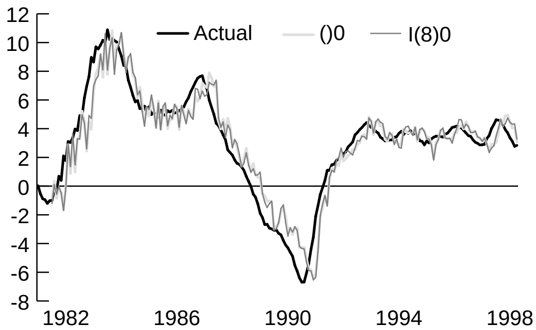

Fig. 27. Fitted LEI models vs. actual: Fitted models using only the lagged growth in CEI, ()0, and including motor vehicle sales taxes, I(8)0.

persistence of the state as measured by the autoregressive parameters in equation (3), current and lagged values of the CEI are excellent predictors. In other words, since the economy exhibits a high degree of momentum, recent growth rates are the best predictors of growth in the near future. Inclusion of the minimal specification illustrates this. Moreover, its inclusion allows one to gauge the impact of the leading indicators on the forecasts. To the extent that the leading indicators add useful information, the fitted values of the model including them should lie closer to the actual curve than those of the minimal specification; and near turning points they should also lie to the left of the minimal specification.

Figures 20 through 39 support several observations about the estimated models. All the models fit well. The proportion of the variance in actual six-month ahead growth rates explained by the models ranges from 85.8 percent for the minimal specification to 92.6 percent for the most highly parameterized specification. The preponderance of the variance is explained by recent growth in the CEI. Individual leading indicators add little additional predictive power, as illustrated by Figs 20

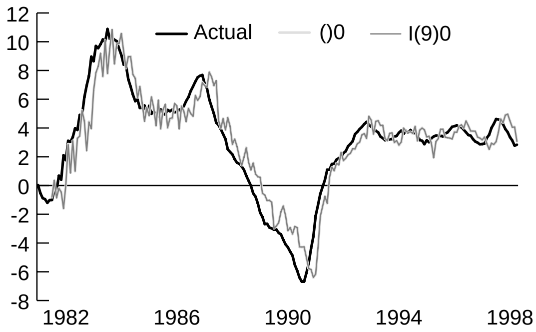

Fig. 28. Fitted LEI models vs. actual: Fitted models using only the lagged growth in CEI, ()0, and including average weekly hours in manufacturing, I(9)0.

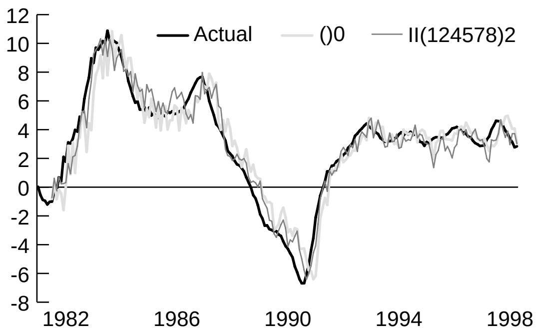

Fig. 29. Fitted LEI models vs. actual: Fitted models using only the lagged growth in CEI, ()0, and the chosen model, II(124578)2.

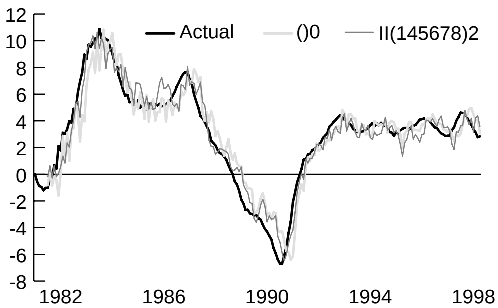

Fig. 30. Fitted LEI models vs. actual: Fitted models using only the lagged growth in CEI, ()0, and an alternative to the chosen model, II(145678)2.

through 28. Nevertheless, the single leading indicators in all but three of these models – help wanted advertising, construction employment, and average weekly

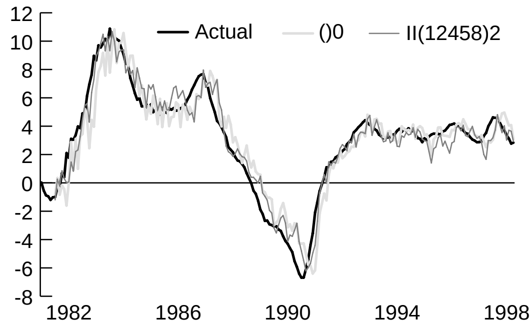

Fig. 31. Fitted LEI models vs. actual: Fitted models using only the lagged growth in CEI, ()0, and an alternative to the chosen model, II(12458)2.

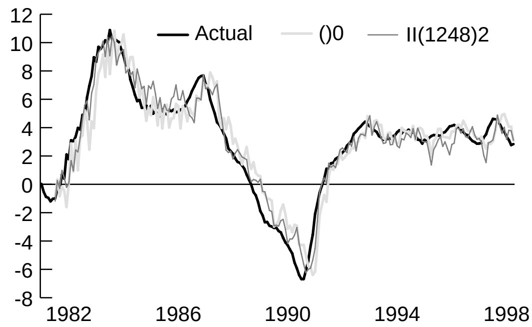

Fig. 32. Fitted LEI models vs. actual: Fitted models using only the lagged growth in CEI, ()0, and an alternative to the chosen model, II(1248)2.

hours in manufacturing – are statistically significant with coefficients several times their estimated standard errors.[17]

In combination, the leading indicators add noticeable predictive power. Several of the better models are displayed in Figs 29 through 37, and two models with all nine leading indicators are graphed in Figs 38 and 39. The models in Figs 29 through 37 were selected because they performed well overall on the six model selection criteria, or were one of the few top models according to the PLS or BIC criteria. They are difficult to distinguish from one another on the basis of their fitted values, and in all cases, the specifications show a marked improvement over the minimal model. They are similar primarily because four of the same leading indicators appear in most of them. Of the nine specifications in Figs 29 through 37, all include

[17]The hypothesis tests given here or implied are based on OLS estimates of the standard errors. As such, they may be unreliable and are biased because the residuals of the model in (6) are not white noise. Therefore they are not to be taken literally. They are used here to suggest relative significance.

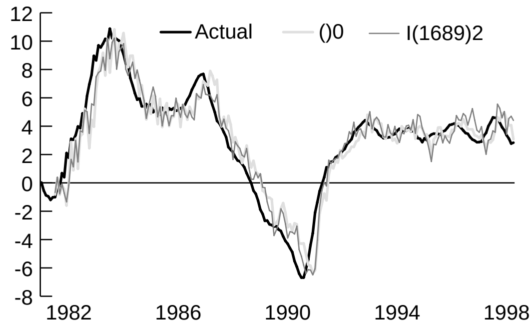

Fig. 33. Fitted LEI models vs. actual: Fitted models using only the lagged growth in CEI, ()0, and the best by predictive least squares, I(1689)2.

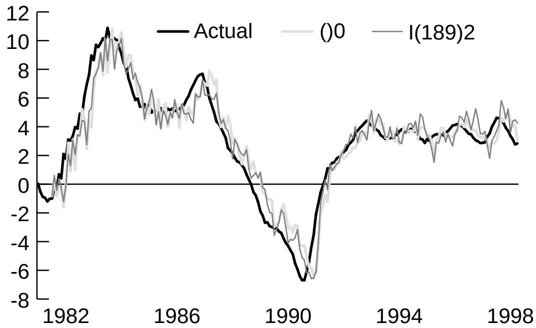

Fig. 34. Fitted LEI models vs. actual: Fitted models using only the lagged growth in CEI, ()0, and the second best by predictive least squares, I(189)2.

consumer confidence, eight include motor vehicle sales taxes, seven include the Bloomberg stock index, and six include the interest rate spread. Almost any of the nine specifications would make a good index. Only the two best-performing models on the basis of the PLS criterion are easy to eliminate from consideration as the "best" specification. This is because one of the indicators, average weekly hours in manufacturing, enters with the "wrong" sign.

The model that was chosen as the best overall according to the six specification evaluation criteria is model # 2960 with specification II(124578)2. The fitted values are displayed in Fig. 29, and the parameter estimates and model statistics are given in Table 8. This model contains six leading indicators; the four often-occurring ones mentioned earlier, plus initial unemployment claims and construction employment. The first four indicators are statistically significant at the five percent level or better, and the latter two are marginally significant, with p-values of 0.15 and 0.17

Table 8
Model # 2960 estimates

Dependent Variable		ln[CEI($t+6$)/CEI(t)]			
Estimation Period		202 months from July 1981–April 1998			
Independent Variable		Coefficient	Standard Error	T-Statistic	P-Value for Variable Group
Constant		0.0012	0.0011	1.1	
Regime Dummy (pre 1988)		0.0044	0.0011	4.1	
CEI	t	2.4227	0.4061	6.0	
	$t-1$	1.6803	0.2484	6.8	0.0000
	$t-2$	−0.8025	0.4069	−2.0	
Consumer	t	0.0006	0.0002	3.1	
Confidence	$t-1$	−0.0006	0.0003	−2.2	0.0020
	$t-2$	0.0006	0.0002	3.0	
Interest Rate	$t+1$	0.0013	0.0012	1.2	
Spread	t	−0.0004	0.0018	−0.2	0.0307
	$t-1$	0.0003	0.0011	0.3	
Bloomberg Stock	$t+1$	0.1336	0.0284	4.7	
Index	t	−0.1190	0.0428	−2.8	0.0000
	$t-1$	0.0723	0.0286	2.5	
Initial	t	−0.0571	0.0282	−2.0	
Unemployment	$t-1$	0.0556	0.0393	1.4	0.1516
Claims	$t-2$	−0.0424	0.0275	−1.5	
Construction	t	0.0034	0.1588	0.0	
Employment	$t-1$	−0.1353	0.2479	−0.5	0.1704
	$t-2$	0.2165	0.1493	1.4	
Motor Vehicle	t	0.0812	0.0255	3.2	
Taxes	$t-1$	−0.0426	0.0304	−1.4	0.0000
	$t-2$	0.0951	0.0248	3.8	
R-Square	0.9214		F(22,179)	95.3559	
R-Square Bar	0.9117		BIC	−9.9598	
RMSE	0.0054				

The coefficients and standard errors are ordinary least squares (OLS) estimates.
The p-values test the null hypotheses that the corresponding set of three coefficients are zero.

respectively.[18] A specification including these two relatively doubtful indicators was chosen because, like the other four included indicators, their effects were of the expected sign and of a reasonable magnitude. The chosen specification is from set II, which involves separate parameters on the leading indicators for the regimes T_1 before 1988 and T_2 for 1988 and later, the coefficients of the leading indicators in the first time period constrained to be proportional to those in the first time period, with the factor of proportionality being b_1/b_2 from equation (4). The leading indicator coefficient estimates reported in Table 8 correspond to the second time period. No constraint is imposed on the estimates of the constants of the two regimes. The root mean square error (RMSE), i.e., standard deviation of the residuals, is 0.0054, which suggests that a typical (in sample) error in predicting a six-month growth is about one percentage point of growth at annual rates.

[18] See the prior footnote.

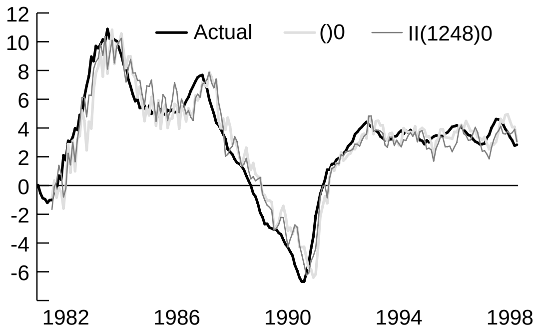

Fig. 35. Fitted LEI models vs. actual: Fitted models using only the lagged growth in CEI, ()0, and the best by the BIC, II(1248)0.

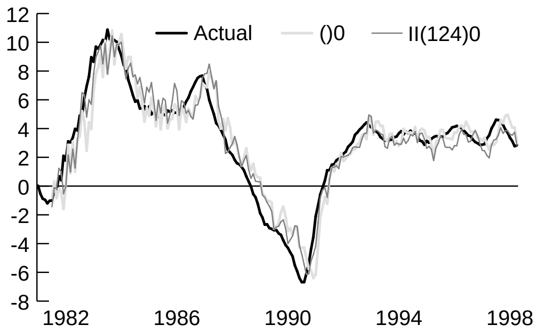

Fig. 36. Fitted LEI models vs. actual: Fitted models using only the lagged growth in CEI, ()0, and the second best by the BIC, II(124)0.

The leading indicators in this specification add about 5 percentage points to the proportion of explained variance of the six-month ahead growth rate of the CEI. This is enough to improve the fit and timing of the index relative to the minimal specification. These improvements are illustrated in Fig. 29 by the leftward shift, in the neighborhood of the turning points, of the fitted values of the model relative to the minimal specification.

6.1. Turning point analysis

Only three turning points occurred during the span of the LEI: the trough at the beginning of the 1980's expansion, the peak near the end of the 1980's, and the trough at the beginning of the current expansion. The accuracy of LEI in predicting these turning points can be evaluated in sample for these three turning points, and out of sample for the last two turning points. The in sample LEI is the fitted value of the model in equation (6) using the full set of historical data. The out of sample LEI is

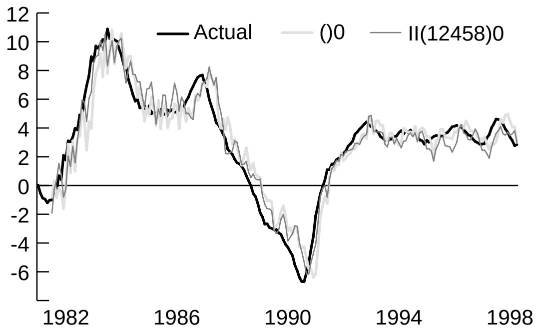

Fig. 37. Fitted LEI models vs. actual: Fitted models using only the lagged growth in CEI, ()0, and the third best by the BIC, II(12458)0.Fig. 37.

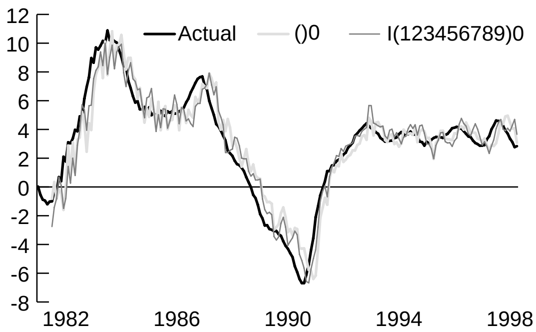

Fig. 38. Fitted LEI models vs. actual: Fitted models using only the lagged growth in CEI, ()0, and a model with all leading indicators, I(123456789)0.

the predicted value of the model using the data up to and including the month the six-month ahead forecast is made, which is given by the PLS forecast. In this section we evaluate the in and out of sample performance of the chosen LEI model in predicting these turning points. The results for other "best" model candidates are similar, as can be verified for the in sample results in Figs 29 through 37. The similarity carries over to the out of sample results for the same set of models,[19] although the models that scored highest on the PLS criteria have – not surprisingly – somewhat better out of sample results.

There are three reasonable measures of the LEI's turning-point performance. One is when the LEI crosses zero relative to when the CEI indicates a turning point. A second is when the LEI crosses zero relative to the turning point of the actual six-month ahead growth in the CEI. A third is the error in the predicted six-month

[19]The figures for the out of sample forecasts are not provided here, but are available by request from the authors.

Table 9
Turning points of the national economy, and the Massachusetts CEI, LEI, and establishment employment

	Peak	Trough	Peak	Trough	Peak	Trough
NBER	Jan-80	Jul-80	Jul-81	Nov-82	Jul-90	Mar-91
CEI	Jan-80	Jun-80	May-81	Dec-81	Dec-88	Jun-91
Turning Points of 6-Month Ahead Growth Rates						
Actual						
Month				Oct-81	Sep-88	May-91
Growth				0.7	0.0	0.4
In Sample Forecast						
Month				Oct-81[1]	Feb-89[2]	Jul-91[3]
Growth				0.7	0.3	0.4
Out of Sample Forecast (PLS)						
Month					Jul-89[4]	Sep-91
Growth					2.7	−1.2
Employment	Mar-80	Jul-80	Jun-81	Jun-82	Dec-88	Dec-91

The NBER dates are for the national economy, and are determined by the Business Cycle Dating Committee of the National Bureau of Economic Reasearch.
The LEI results are for model # 2960.
Turning point months of the 6-month ahead growth rates are the month in which the LEI or actual 6-month ahead growth rates in the CEI changed sign.
Growth rates are the actual or predicted values of the LEI at the actual 6-month ahead growth turning point.
[1]The LEI first turned positive in August, but dipped back below zero in September before turning and remaining positive in October.
[2]The LEI was below 0.5 percent in September 1988 through January 1989 before turning negative in February.
[3]The LEI turned positive in May, but dipped back below zero in June before turning and remaining positive in July.
[4]The LEI turned negative in April and May 1989, but turned positive in June 1989.

ahead growth at the turning point of the actual six-month ahead growth in the CEI. These measures for the chosen LEI model are presented in Table 9.

6.1.1. In sample performance

The in sample record of the ability of the LEI to predict turning points is mixed. For the trough in the CEI that occurred in December 1981, the LEI performed well on all three measures. The LEI turned positive in October, and the actual six-month ahead growth in the CEI also turned positive in October. The LEI first turned positive in August, but dipped back below zero in September before turning positive again in October. The difference between actual and predicted growth in October was less than 0.1 percent at annual rates.

The LEI missed the peak of December 1988. It turned negative two months later in February, while the actual six-month ahead growth turned negative in September 1988. In terms of the level of growth, rather than its sign, the performance of the LEI was not bad. The error in the growth rate in September was only 0.3 percent at annual rates. Also, the LEI was predicting six-month ahead growth of less than

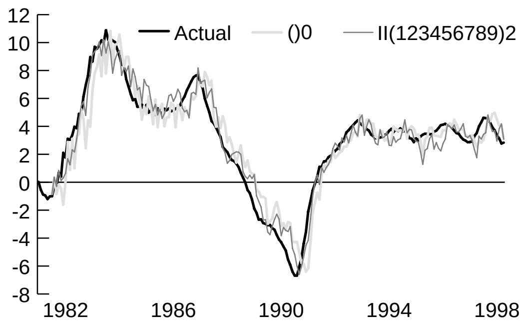

Fig. 39. Fitted LEI models vs. actual: Fitted models using only the lagged growth in CEI, ()0, and a model with all leading indicators and the most parameters, II(123456789)2.

one-half percent (at annual rates) as early as September, and the LEI remained below one-half percent for five consecutive months before turning negative in February.

The LEI also did not lead the trough in June 1991. It turned positive one month later in July, while the actual six-month ahead growth in the CEI turned positive in May. The LEI also turned positive in May, but dipped back down below zero in June. The growth error again was very small, at less than 0.1 percent in May. However, despite the mediocre performance of the index on usual turning point measures, in real time the index might have been perceived to work very well. This is because most economy watchers at the time believed that Massachusetts was still in a recession as late as the fall of 1992, before the dramatic upward revision in establishment employment, released in March 1994, revised the turning point in employment by over a year earlier. Since the index relies on nine other indicators, its revision may have been substantially less than that of employment.

6.1.2. Out of sample performance

The out of sample LEI forecasts missed both of the last two turning points by substantial margins on all three measures. For the peak that occurred in December 1988, the LEI did not turn negative until July 1989. It did turn negative three months earlier in April and May, but then moved back in to positive territory in June. In September 1988, when the actual six-month ahead growth in the CEI turned negative, the LEI predicted positive growth of 2.7 percent at annual rates.

For the trough of June 1991, the LEI did not turn positive until September. In May, when the actual six-month ahead growth in the CEI turned positive, the LEI predicted a decline of 1.2 percent at annual rates.

Three factors account for differences between the in and out of sample results: data, coefficient estimates, and updates of the CEI. The out of sample data differs from the in sample situation in that the filtered withholding and sales taxes rely on four month forecasts in order to provide the future data needed for the nine month centered filter; and the real Bloomberg stock index relies on a one month forecast of the CPI for deflation purposes. These data differences are minor, and account for a tiny fraction – roughly 5 percent – of the difference between the in sample and out of sample growth predictions.

The major portion of the deviation between the in and out of sample predictions is due to differences in the parameter estimates of the LEI model. They account for 56 percent of the difference in growth rates in September 1988 and 68 percent in May 1991. These models were estimated on essentially only one expansion phase and one full cycle respectively, so it is not surprising that the PLS parameter estimates diverged significantly from the full sample estimates.

Virtually all of the remaining difference is accounted for by updates in the CEI due to the Kalman smoother. The estimate of the CEI for a given month is updated as future data become available, with most of the revision occurring in the first few months. Revisions in the CEI translate directly into revisions in the LEI. Thus, the model is formally mimicking the “Monday morning quarterbacking” that goes on in practice, and that is sometimes instituted in ad hoc rules such as moving averages.

One might ask why we did not undertake an extensive analysis of turning points, perhaps even including the ability to anticipate the CEI turning points as an additional criterion for choosing the preferred leading index model. The main reason we did not do this is that turning point criteria are essentially captured by the PLS and BIC criteria for fit. If the model fits well, it is likely to do a better job of leading. Also, according to the linear structure of the model, there is no reason to believe that specifications that perform marginally better in terms of anticipating turning points, but which are equivalent in terms of other criteria like fit and number of indicators, would perform any better in predicting turning points in real time. Of course, this is a conjecture that should be tested, but unfortunately not with our data. The reason is that, because of the short time-span of our data, there are only two turning points that could be evaluated with out of sample techniques, and the first would be based on parameter estimates that, in effect, are based on only a single expansion phase of one cycle. Finally, one of the features of the leading index is that it is used far more often these days to forecast near-term growth in the economy than it is to assess the probability of entering a recession. For this task, forecast fit is the primary criterion.

The most closely watched single indicator is establishment employment, so it is interesting to compare its turning points to those of the indexes. Table 9 shows that both the in sample LEI and CEI led the two troughs by many months. This is not surprising, since the trend of the CEI is higher than that of employment. In the long run, their trends differ by productivity growth. The three CEI and employment peaks were roughly coincident, with the CEI leading by two, one, and zero months. This is interesting given that the higher trend rate of growth in the CEI would tend to result in the CEI lagging employment at peaks. The result may be explained by two factors. First, hours of work may decline before employment, so that employment peaks lag labor input peaks. Second, two of the coincident indicators, withholding and sales taxes, may lead employment.

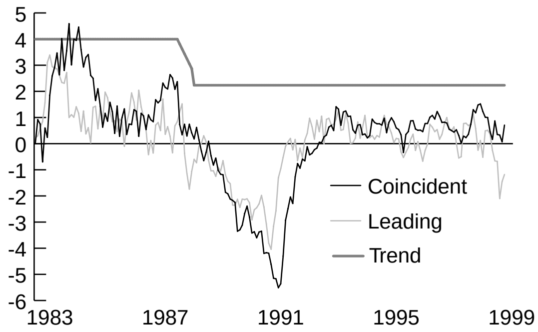

Fig. 40. Decompostion of the LEI into coincident, leading, and trend contributions.

6.2. Decomposition of the leading index

It is useful in analyzing current economic conditions to decompose the leading index in such a way as to measure the contribution of each indicator to the overall index value each month. Equation (6) suggests a straightforward methodology for decomposing the leading index into the sum of a trend rate of growth and the contribution of each indicator to greater or less than trend rate of growth. The trend rate of growth is the fitted or predicted value of the model evaluated at the mean of each independent variable. The contribution of each indicator is the deviation of its value from its mean times its coefficient. Thus, mathematically, these components sum to the value of the index in each month. One of the right-hand side variables is the growth in the CEI. Rather than report the contribution of the CEI, it is more useful to decompose its contribution into the separate contributions of each coincident indicator. This is possible by chaining back through equation (4) and the Kalman smoother. Equation (4) expresses the growth in the CEI as a linear function (or rather, a linear function for each regime) of the growth in the state of the economy, $\Delta s_{t|T}$, which in turn is a linear function, given by the Kalman smoother, of the normalized growths of the coincident indicators, Δx_t. One consequence of this substitution is that the trend rate of growth has a step when the regime changes from T_1 to T_2. In our chosen model, the regime dummy parameter also contributes to the step in the trend rate of growth.[20]

The contributions of the indicators to the chosen leading index model are displayed in Figs 40 through 43. In Fig. 40, the contributions of the coincident and leading indicators are separately aggregated, so that the leading index is decomposed into trend, coincident, and leading components. The leading index is simply the sum of the three curves in the graph.

[20]The regime dummy parameter does not switch in one step, but instead has a six-month linear phase in, because the dependent variable is the six-month ahead growth rate.

Fig. 41. Contributions of the coincident indicators.

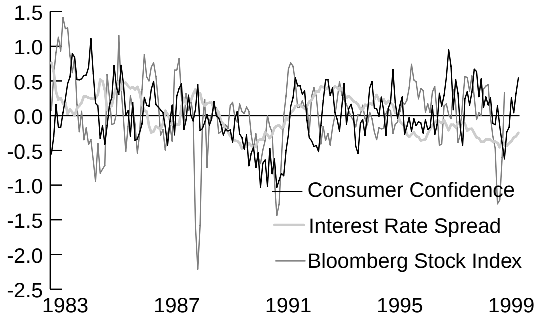

Fig. 42. Contributions of the leading indicators.

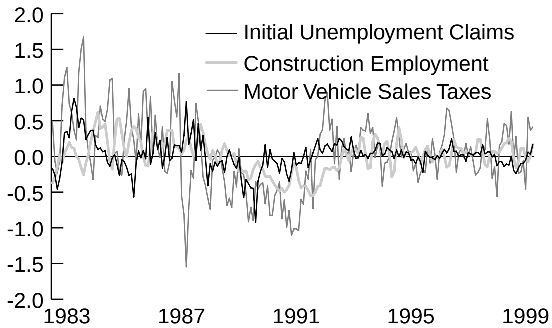

Fig. 43. Contributions of the leading indicators, continued.

The trend rate of growth is substantially higher in the earlier regime, which roughly coincides with the "Massachusetts Miracle" years. This suggests that growth in GSP during these years was greater than that associated with either the coincident or leading indicators. This may be the result of some real phenomenon, or it may simply be due to data anomalies in the measurement of GSP or several of the indicators.

The zero line separates the contributions of the indicators into those that increase the predicted rate of growth above trend (positive values), versus those that decrease the predicted rate of growth below trend (negative values). An indicator contributes to above trend growth if its growth is above its own average value. (The converse applies to the unemployment rate and initial unemployment claims.)

Figure 40 does a good job of showing the contribution made by the leading indicators in improving the timing of the leading index relative to the minimal specification. In the recoveries of 1982 and 1991, the leading indicators contributed to above trend growth before the coincident indicators did. In late 1987, the leading indicators fell sharply below trend, preceding the rapid decline in the economy's growth rate prior to the 1989 recession.

Figures 41 through 43 present the contributions for each of the individual indicators, giving a richer explanation of movements in the leading index. Thus, for example, the leading indicators' recovery in 1991 was broad-based, including consumer confidence, stock prices, the interest rate spread, and initial unemployment claims. The recent sharp drop in the leading indicators during the short-lived financial scare in the summer of 1998 was composed of stock prices, consumer confidence, the interest rate spread, and initial unemployment claims. The movements in these indicators were consistent with the economic news of the time, and the resulting leading index, which was positive but close to zero, was consistent with reports of sharply increased probabilities of recession, and with the aggressive action by the Fed to support and calm financial markets. On the other hand, the "correct" indication of leading components in the fall of 1987 is revealed to be dominated by a single indicator, stock prices. The stock market crash of October 1987 did have an impact on the Massachusetts economy, but the cause of the subsequent steep recession is widely attributed to other factors, such as the bursting of the housing market bubble and the collapse of the minicomputer industry.

7. Conclusion

The development of state and regional indexes in general, and the application of the Stock/Watson method to their construction in particular, is beginning to move forward at a quickening pace. The computer power and speed necessary to estimate dynamic factor models is now exceeded by even low-end personal computers, and the software needed for such estimation is becoming more widely available.

Our application of the model in this paper suggests several conclusions that may be helpful to other researchers who are developing state and regional indexes.

First, the dynamic factor approach to estimating coincident indexes appears to be a fruitful signal extraction technique. The resulting coincident index for Massachusetts is much smoother than the indicators from which it is formed, and is also smoother than the index that would have resulted from the application of the BEA/Conference Board composite index methodology. This smoothness characteristic is not unique to Massachusetts; similar results have been obtained for nearly every reported state or regional index for which a Stock/Watson model has been estimated.

Second, the coincident index is in accord with our knowledge of the cyclical history of the Massachusetts economy over the last quarter of the century. The index appears to be capturing what it was intended to estimate.

Third, the leading index model appears to have a good chance of working satisfactorily. The leading indicators add predictive power over and above that provided by the CEI alone, and forecast reasonably well in sample. The relatively poor out of sample forecast performance may improve as information accumulates over time. Parameter estimation and model selection can be periodically updated. Also, at least in terms of in sample forecasts, the experience in other states where the Stock/Watson model approach has been tried has been more successful overall. See Crone and Babyak [6], and Orr, Rich, and Rosen [12].

Fourth, pre-estimation filtering appears to be useful, especially because many state and regional series are inherently noisy. A rational approach to pre-estimation filtering may be a fruitful topic of future research. Transforming a set of indicators so that high-frequency noise is uniformly filtered out could make specification and estimation of index models easier. Such pre-filtering should be especially useful for the more traditional BEA/Conference Board method, which averages only across indicators, not time.

Fifth, a systematic approach to model selection is probably a good thing. It has several advantages to a search over a limited number specifications where the path to the final specification is determined by intuition and trial and error. Given today's computing power, systematic approaches are practical. For example, our specification search for the leading index model entailed running over two million regressions, yet involved only a few hours of computer time on a relatively slow processor. Nevertheless, model selection is still subjective, especially when the historical record is short as in Massachusetts, which limits the power of techniques such as PLS.

Last, a probability of recession index would be a useful addition to the leading index. It would help the analyst quantify the likelihood of a turning point occurrence when the signal is ambiguous, e.g., when the index changes sign temporarily, or forecasts near zero growth. Several alternatives are available, including Stock and Watson [16,17], Phillips [13], and Orr, Rich, and Rosen [12].

Appendix A. De-normalization of the Kalman state: A comparison with the BEA/TCB methodology

The state that is output from the Kalman filter in the model given in equations (1)–(3) is, by construction, normalized to be driftless, with unit-variance input shocks.

This appendix explains the Stock/Watson approach for de-normalizing the state, $\Delta s_{t|t}$. Their approach is analgous to the BEA/TCB method, with one major difference. The Bureau of Economic Analysis/The Conference Board (BEA/TCB) methodology imposes an equal-weighting assumption on the input indicators, while the weights in the Stock/Watson methodology are given by the estimated Kalman filter. The BEA/TCB method is described in Green and Beckman [7]. In addition to providing an additional convenient source of the Stock/Watson procedure for the reader, the purpose of this appendix is to highlight the relationship between the BEA/TCB and Stock/Watson approaches.

The de-normalized growth rate of the index in the Stock/Watson approach, Δc_t, is shown to be a linear function of the state from the Kalman filter, given by:

$$\Delta c_t = \mu + F\Delta s_{t|t}. \tag{A1}$$

The constants μ and F have analogues in the BEA/TCB method. The former is a drift formed as a weighted average of the drifts of the coincident indicators, and the latter rescales the variance of the monthly change in the index. In the following analysis the parameters μ and F are derived from the Kalman filter for $\Delta s_{t|t}$. Alternatively, they could be derived from the Kalman smoother filter for $\Delta s_{t|T}$, with little difference in the calculated values for μ and F.

First, some definitions and notation need to be introduced, which are slightly different in the BEA/TCB and Stock/Watson methodologies. These minor differences are of no practical consequence in the comparison of the denormalization methodologies.

The first way in which the two approaches are slightly different is in the definition of the first difference of an input series. In the BEA/TCB method, this is defined as the symmetric percentage change:

$$\Delta X_{jt} = 200\frac{X_{jt} - X_{j,t-1}}{X_{jt} + X_{j,t-1}} \tag{A2a}$$

where X_{jt} is the value of the jth series in month t. The series X_j is not logged. In the Stock/Watson method, the first difference is defined as the first difference of the natural logarithm of the series:

$$\Delta X_{jt} = \ln(X_{jt}) - \ln(X_{j,t-1}). \tag{A2b}$$

The two measures are essentially identical in practice.[21]

The second minor difference concerns the average growth deviation measure for an input series. In the BEA/TCB method, this is defined as the average absolute change:

$$\sigma_j = \frac{\sum\limits_t |\Delta X_{jt}|}{T} \tag{A3a}$$

[21] For some series, for example, unemployment rates, the first difference is calculated as the simple first difference: $\Delta X_{jt} = X_{jt} - X_{j,t-1}$.

where T is the number of observations for the series. In the Stock/Watson method, the measure is the standard deviation:

$$\sigma_j = \sqrt{\frac{\sum_t (\Delta X_{jt} - \mu_j)^2}{T - 1}} \tag{A3b}$$

where the mean of the series, μ_j, is defined the same in both approaches:

$$\mu_j = \frac{\sum_t \Delta X_{jt}}{T}. \tag{A4}$$

The difference in the deviation measures imparts a very minor difference in the calculated weights between the two approaches that would otherwise not be present.

We will need to use the standardized first difference of an input series, defined as:

$$\Delta x_{jt} = \frac{\Delta X_{jt} - \mu_j}{\sigma_j} \tag{A5}$$

The main difference between the two approaches is the manner in which the input series are filtered to produce the estimated state of the economy. In the Stock/Watson approach, the state is calculated according to:

$$\Delta s_{t|t} = \sum_j m_j(L) \Delta x_{jt}, \tag{A6}$$

where

$$m_j(L) = m_{j0}L^0 + m_{j1}L^1 + m_{j2}L^2 + \cdots \tag{A7a}$$

is the Kalman filter.[22] In factor analysis terminology, the filter would be called the "scoring coefficients", since they transform observations on several variables (for example, answers to a battery of questions) into a factor score (for example, an intelligence or attitude index). Stock and Watson call the filter the set of dynamic multipliers, which reflects the way the filter is calculated in practice, i.e., by observing how the series $\Delta s_{k|k}, k \geqslant t$, changes in response to a one-time unit change in Δx_{jt} (a pulse equal to one in period t and zero before and after period t).

In the BEA/TCB methodology, the Kalman filter is effectively replaced by the following:

$$m_j(L) = m. \tag{A7b}$$

[22] L is the lag operator, i.e., $L^k \Delta x_t = \Delta x_{t-k}$.

This is the essence of the "equal weighting" methodology. Each normalized series enters the index equally. It doesn't matter what the value of m is, so the BEA/TCB method implicitly uses $m = 1$. Furthermore, the composite is formed only from the contemporaneous values of the indicators.

We now show how the BEA/TCB de-normalized index is formed. This closely follows the procedure given in Green and Beckman [7], with slight differences in notation needed to highlight the relationship with the Stock/Watson method. First, define

$$B_j = \frac{m_j(1)}{\sigma_j}, \text{ where, from (A7a), } m_j(1) = \sum_{k\geqslant 0} m_{jk}. \tag{A8a}$$

Note that, from (A7b), this simplifies to

$$B_j = \frac{m}{\sigma_j} \tag{A8b}$$

for the BEA/TCB methodology. Next, define

$$F = \frac{1}{\sum_j B_j}, \text{ and} \tag{A9}$$

$$w_j = FB_j. \tag{A10}$$

Note that $\sum_j w_j = 1$. Also define a weighted mean growth rate:

$$\mu = \sum_j w_j \mu_j. \tag{A11}$$

In the BEA/TCB methodology, the growth of the denormalized index can now be expressed as:

$$\Delta c_t = \mu + \sum_j W_j(\Delta X_{jt} - \mu_j) = \sum_j w_j \Delta X_{jt}. \tag{A12}$$

We now show how this is equivalent to (A1), when one substitutes the Kalman filter for the equal weighting scheme of the BEA/TCB. Begin by using (A5) to replace ΔX_{jt} in the middle portion of (A12) by $\sigma_j \Delta x_{jt} + \mu_j$.

$$\Delta c_t = \mu + \sum_j w_j \sigma_j \Delta x_{jt}. \tag{A13}$$

Using (A10), this becomes

$$\Delta c_t = \mu + F \sum_j B_j \sigma_j \Delta x_{jt}. \tag{A14}$$

Using (A8b), this simplifies to

$$\Delta c_t = \mu + F \sum_j m \Delta x_{jt}. \tag{A15}$$

Now the crucial substitution of the Kalman filter is made, using (A7b):

$$\Delta c_t = \mu + F \sum_j m_j(L) \Delta x_{jt}. \tag{A16}$$

Finally, substituting (A6) for the summation gives us (A1):

$$\Delta c_t = \mu + F \Delta s_{t|t}. \tag{A17}$$

So the denormalized growth in the index in the Stock/Watson methodology is a linear function of the state from the state space model in equations (1)–(3). Furthermore, the only essential difference between the Stock/Watson and BEA/TCB methodologies is the use of the Kalman filter in place of the equal weighting assumption.

Appendix B. Calibrating the Kalman state to Gross State Product

This appendix describes how the state vector from the model in equations (1)–(3) was calibrated to Massachusetts Gross State Product, i.e., it describes how the coefficients a_1, b_1, a_2, and b_2 in equation (4) were calculated. The procedure described below can be used to fit the index to the first two moments of the growth – the average growth rate and average deviation from trend – of any observable series.

Consider an index constructed from the state output by the Kalman filter so that its growth is a linear function of the state:

$$\Delta c'_t = \mu + F \Delta s_{t|T}. \tag{B1}$$

The construction of such an index is described in Appendix A, but this appendix is also applicable to calibrating the state vector directly, i.e., when $\mu = 0$ and $F = 1$. $\Delta c'_t$ represents the growth in the index, or the first difference of the log of the index. Suppose that the calibration is to be performed to fit Gross State Product over the period of time $t \in T$. This appendix describes how to choose the parameters a and b so that the linear function of the state

$$\Delta c_t = a + b \Delta s_{t|T} \tag{B2}$$

matches the first two moments of Gross State Product growth during the period T.

Suppose that the first two moments of $\Delta c'_t$ from (B1) are g_c and d_c respectively, i.e., they are respectively, the mean and standard deviation of $\Delta c'_t$ over the period

T.[23] Suppose also that the first two moments of the growth of Gross State Product over the period T are g_{GSP} and d_{GSP} respectively.[24]

The desired parameters in (B2) can be derived in two simple steps. First, detrend and standardize (B1) over the period T by subtracting its growth and dividing by the average deviation:

$$\Delta c_t'' = \frac{\mu + F\Delta s_{t|T} - g_c}{d_c}. \quad \text{(B3)}$$

Next, reconstitute $\Delta c_t''$ so that is has an average growth of g_{GSP} and an average deviation of d_{GSP}:

$$\Delta c_t = \Delta c_t'' d_{GSP} + g_{GSP} = fF\Delta s_{t|T} + f(\mu - g_c) + g_{GSP}, \quad \text{where } f = \frac{d_{GSP}}{d_c}. \quad \text{(B4)}$$

Then

$$\begin{aligned} a &= f(\mu - g_c) + g_{GSP} \\ b &= fF \end{aligned} . \quad \text{(B5)}$$

In order to get a satisfactory fit, it may be necessary to apply this procedure piecewise by dividing the entire span of the index and Gross State Product into two or more periods. For our index, two periods, 1978–1987 and 1988–1996, were deemed to be satisfactory.

Acknowledgements

We gratefully acknowledge helpful comments from Robert Rich and an anonymous referee. We also thank Douglas Woodward and Charles Renfro for their support and patience throughout the project, as well as the University of Minho, Braga, Portugal and their faculty and staff for hosting the Workshop on Regional Economic Indicator Models, at which an earlier version of this paper was presented. This research was supported by a Public Service Grant from the University of Massachusetts Boston. Software for estimating the coincident index model is available from Alan Clayton-Matthews (acm@mediaone.net), who is also responsible for remaining errors and omissions.

---

[23] We used the average absolute deviation, but the same procedure would apply to the standard deviation. Also, because the index is monthly but Gross State Product is annual, $\Delta c_t'$ was annualized before calculating the moments. The growth rate g_c was expressed as a monthly growth rate in order to be consistent with μ from (B1). The average deviation, d_c , was retained in annual units.

[24] The growth rate g_{GSP} is expressed as a monthly growth rate, in order to be consistent with g_c and μ.

References

[1] B.D.O. Anderson and J.B. Moore, *Optimal Filtering,* Prentice Hall, Englewood Cliffs, New Jersey, 1979.

[2] A. Clayton-Matthews, DSFM manual: version 2/23/99, mimeo, 1999.

[3] A. Clayton-Matthews, Y.K. Kodrzycki and D. Swaine, Indexes of economic indicators: what can they tell us about the New England economy, *New England Economic Review* (November/December 1994), Federal Reserve Bank of Boston, 17–41.

[4] T.M. Crone, New indexes track the state of the states, *Business Review* (January/February 1994), Federal Reserve Bank of Philadelphia, 19–31.

[5] T.M. Crone, Using state indexes to define economic regions in the US, paper prepared for the Workshop on Regional Economic Indicator Models, Braga, Portugal, June 8–10, 1998.

[6] T.M. Crone and K.J. Babyak, Looking ahead: leading indexes for Pennsylvania and New Jersey, *Business Review* (May/June 1996), Federal Reserve Bank of Philadelphia, 3–14.

[7] G.R. Green and B.A. Beckman, Business cycle indicators: upcoming revision of the composite indexes, *Survey of Current Business* (October 1993), 44–51.

[8] J.D. Hamilton, *Time Series Analysis,* Princeton University Press, Princeton, New Jersey, 1994.

[9] R.W. Hamming, *Digital Filters,* (Third ed.), Dover Publications, Mineola, New York, 1998.

[10] W.K. Newey and K.D. West, A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix, *Econometrica* **55**(3) (May 1987), 703–708.

[11] J. Orr, R. Rich and R. Rosen, Two new indexes offer a broad view of economic activity in the New York – New Jersey region, *Current Issues in Economics and Finance* **5**(14) (October 1999), Federal Reserve Bank of New York.

[12] J. Orr, R. Rich and R. Rosen, Leading economic indicators for New York and New Jersey, mimeo, Federal Reserve Bank of New York, September 30, 1999.

[13] K.R. Phillips, The Texas index of leading economic indicators: a revision and further evaluation, *Economic Review* (July 1990), Federal Reserve Bank of Dallas.

[14] T.J. Sargent and C.A. Sims, Business cycle modeling without pretending to have too much a priori economic theory, in: *New Methods in Business Cycle Research: Proceedings from a Conference,* Federal Reserve Bank of Minneapolis, 1977.

[15] J.H. Stock and M.W. Watson, New indexes of coincident and leading economic indicators, *NBER Macroeconomics Annual* (1989), 351–394.

[16] J.H. Stock and M.W. Watson, A probability model of the coincident economic indicators, in: *Leading Economic Indicators: New Approaches and Forecasting Records,* K. Lahiri and G.H. Moore, eds., Cambridge University Press, Cambridge, 1991, pp. 63–89.

[17] J.H. Stock and M.W. Watson, A procedure for predicting recessions with leading indicators: econometric issues and recent experience, in: *Business Cycles, Indicators, and Forecasting,* J.H. Stock and M.W. Watson, eds., The University of Chicago Press, Chicago and London, 1993, pp. 95–156.

[18] C.K.H. Tsao, The Massachusetts stock index: methodology in construction using historical data, mimeo, University of Massachusetts Boston, 1998.

[19] C.K.H. Tsao, Regional coincident economic indexes: an investigation of the New England states' "state of the economy", Ph.D. Dissertation, University of Massachusetts Boston, 1999.

[20] C.Z. Wei, On predictive least squares principles, *The Annals of Statistics* **20**(1) (1992), 1–42.