Coincident and leading indicators of the stock market

Marcelle Chauvet *, Simon Potter¹

Department of Economics, University of California, Riverside, CA 92521-0247, USA
Federal Reserve Bank of New York, 33 Liberty St., New York, NY 10045-0001, USA

Accepted 15 October 1999

Abstract

In this paper we have two goals: first, we want to represent monthly stock market fluctuations by constructing a non-linear coincident financial indicator. The indicator is constructed as an unobservable factor whose first moment and conditional volatility are driven by a two-state Markov variable. It can be interpreted as the investors' real-time belief about the state of financial conditions. Second, we want to explore an approach in which investors may use their perceptions of the state of the economy to form forecasts of financial market conditions and possibly of excess returns. To investigate this, we build leading indicators as forecasts of the estimated coincident financial index. The leading indicators yield better within and out-of-sample performance in forecasting, not only the state of the stock market but also of excess stock returns, as compared with the performance obtained using linear methods that have been proposed in the existing literature. © 2000 Elsevier Science B.V. All rights reserved.

JEL classification: G1; C32
Keywords: Stock Market; Excess returns; Risk premia; Dynamic factor; Markov-switching

* Corresponding author. Tel.: +1-909-787-5037 ext. 1587; fax: +1-909-787-5685; e-mail: chauvet@mail.ucr.edu
¹ Tel.: +1-212-720-6309; fax: +1-212-720-1844; e-mail: simon.potter@ny.frb.org.

0927-5398/00/$ - see front matter © 2000 Elsevier Science B.V. All rights reserved.
PII: S0927-5398(99)00015-8
1. Introduction

A great deal of evidence has recently been uncovered concerning the predictability of excess stock returns using financial and macroeconomic variables that convey publicly available information about business conditions. Most of this evidence is based on linear regression methods, which show that lagged values of the dividend yield, earning-price ratios, interest rates, and measures of the term premium or default risk are statistically significant explanatory variables. Some researchers have interpreted these regressors as proxies for unobserved time-varying risk premia. \(^2\) One conclusion has been that the unobserved risk premia vary systematically over the business cycle.

Most related empirical studies assume a constant linear relationship between stock returns and proxies for market risk. Although economic theory does not provide unequivocal implications concerning the relationship between risk and excess returns, equilibrium asset pricing models lead one to expect that it may be non-linear. \(^3\) The empirical presence of non-linearity has been illustrated by the distinct behavior of excess returns grouped according to the state of the business cycle. For example, Chauvet and Potter (1998) find a time-varying relationship between risk and return around business cycle turning points. On the other hand, Fama and French (1989), Whitelaw (1994), Pesaran and Timmermann (1995), Perez-Quiros and Timmermann (1998), and Chauvet (2000) find evidence of significant state dependence in the conditional distribution of stock returns. In particular, the last two papers show that financial variables proxying for risk forecast NBER business cycle turning points.

This paper has two goals: first, we want to represent stock market fluctuations by constructing a non-linear coincident financial indicator. The indicator is based on a broad information set about market conditions including the overall state of financial markets, which is generated from expectations about changes in future economic activity. Second, we build leading financial indicators to predict stock market fluctuations as forecasts of the estimated coincident index.

In particular, we propose modeling swings in the stock market at the monthly frequency as a function of investors’ reactions towards changes in unobserved market risk factors. The non-linear proxy for the market risk premia is constructed as a latent factor whose first moment and conditional volatility are driven by a


\(^3\) Optimality conditions yield a non-linear function between excess returns and the stochastic discount factor capturing intertemporal consumption risk. In addition, when using conditional variance as a proxy for risk, Backus and Gregory (1993) find that theoretical models are consistent with any sort of relationship between excess return and risk, depending on model preferences and the probability structure across states.
two-state Markov variable. Excess returns and their volatility may reflect changes in average opinions about financial information as investors learn more about the data. However, the econometrician cannot directly observe the information used by the market participants to assess the effect on stock returns of changes in systematic risk. Moreover, given a broad information set, investors’ responses may be asymmetric, depending on their perception of the state of business conditions. For example, changes in interest rates may be interpreted as bad or good news depending on whether the economy is in a recession or a boom. The Markov-switching introduced in the unobserved factor can capture the asymmetric behavior of investors toward risk and their inclination to hedge against noise, and is capable of generating higher levels of volatility in certain states of the economy.

Unlike previous regression methods, the dynamic factor does not have an immediate interpretation as a forecast of excess returns, because it is produced using contemporaneous information on a set of financial variables. The estimated indicator can be interpreted as the investors’ real-time belief about the state of financial conditions. That is, it is a coincident indicator of movements in the stock market reflecting common assessments of the implications of given sets of financial information.

We explore an approach in which investors may use their perceptions about the state of stock markets to form forecasts of financial conditions. To investigate this, we build leading financial indicators as forecasts of the estimated coincident financial index in two ways: by projecting it forward based on its history and by constructing a leading indicator using the history of the dynamic factor and other variables in an autoregressive system. These predicted values can then be used not only as a forecast of the state of stock market but also of excess returns.

The empirical evidence of asymmetries in financial variables as a function of the underlying state of the economy has been based mainly on recessions and expansions as measured by the NBER. Although the NBER business cycle dating is generally recognized as the official chronology of turning points, there are some problems associated with using these dates to formally analyze financial variables. In particular, one of the NBER dating rules is that recessions must correspond to a general downturn in various sectors of the economy for a minimum duration of 6 months. Thus, the NBER recessions do not reflect periods of short-lived contractions or periods in which only some sectors of the economy might experience a downturn. There is no reason to believe that the stock market only responds to recessions as defined by the NBER.

---

4 Chauvet and Potter (1999) propose an equilibrium asset pricing model under learning that formalizes this argument.

As is well known, there have been more bear markets than recessions, although recessions are always associated with a bear market. In this paper, we explore the idea that stock markets may also react to sectoral or shorter-lived contractions in the economy not accounted for by the NBER dating, and so the frequency of bull and bear markets may be higher than the frequency of business cycle expansions and contractions.

In terms of results, the dynamic factor that we estimate exhibits distinct bull and bear market behavior with high amplitude and a financial cycle of shorter duration than the business cycle. In fact, using the estimated Markov probabilities of bull/bear states for dating the stock market cycle, we find that bear markets generally start a couple of months before an economic contraction and end before the trough of a recession, anticipating economic recovery (see Fig. 1). That is, the stock market cycle leads the business cycle and seems to be generated from expectations about changes in future economic activity. The factor is highly correlated with the excess return on the CRSP valued-weighted index, but is more highly autocorrelated and, therefore, more predictable, than the excess return series itself. In most related studies, there is evidence that idiosyncratic noise makes it difficult to detect the presence of predictable components. The dynamic financial market factor is, by construction, an indicator that captures only contemporaneous comovements underlying the financial variables and, therefore, separates information driven by common shocks from information driven by idiosyn-
catic shocks. In addition, the indicator is correlated mainly with variables that forecast future economic activity, such as changes in the Conference Board leading economic indicator, the spread between the rate of return on commercial paper and treasury bills (T-bills), and the bear/bull phases of the stock market. The leading financial indicators produced in this paper yield very good within-sample forecasting performance compared to alternative models, which also holds for out-of-sample performance. The results provide evidence on the importance of the state of financial conditions in predicting future market conditions and future excess returns.

The paper is organized as follows. Section 2 presents and interprets the dynamic factor model and the autoregressive system for the leading indicator. In Section 3, the empirical results are presented for both the coincident and leading indicators, and the model is tested for out-of-sample performance. Section 4 concludes and suggests direction for future research.

2. The dynamic stock market model

2.1. The coincident financial indicator

We propose modeling the coincident financial indicator, \( F_t \), as an unobservable autoregression whose mean and conditional volatility are functions of a Bernoulli random variable \( S_t \), which represents the state of financial markets, either bull or bear:

\[
F_t = \alpha_0 + \alpha_1 S_t + \phi F_{t-1} + \eta_t \quad S_t = 0,1
\]

\[
F_t \sim \{ \mu_{S_t}, \theta_{S_t} \} \quad \eta_t \sim \text{i.i.d. } N(0, \sigma_{\eta_t}^2).
\]

where \( F_t \) is the dynamic factor that proxy for the market risk premium and \( \theta_{S_t} \) is the state dependent conditional volatility of \( F_t \). \(^8\) Both the mean and the volatility of the dynamic factor switch between states, governed by the transition probabilities of the first order two-state Markov process, \( p_{ij} = \text{Prob}[S_t = j|S_{t-1} = i], \sum_{j=0}^{1} p_{ij} = 1, i,j = 0,1. \) That is, financial markets can be either in an expansion

\(^8\) The number of factors underlying the variables was tested by the usual method undertaken in factor analysis, that is, through examination of the eigenvalues of the common factors correlation matrix. The magnitude of the eigenvalues for each factor reflects how much of the correlations among the observable variables a particular factor explains. The procedure indicates strong evidence for the single factor specification for the underlying variables used here.
period (bull market), $S_t = 0$; or in a contraction state (bear market), $S_t = 1$. Note that the conditional volatility could either be higher or lower in the bear market than in the bull market.\(^9\)

Our dynamic factor approach starts with a vector of financial variables that move contemporaneously with the stock market cycles. This vector includes variables that reflect public information about financial conditions such as the excess stock return (defined as the difference between continuously compounded returns on the CRSP valued-weighted index and the 3-month T-bill rate), the squared excess stock returns as a proxy for market volatility, the 3-month T-bill rate, and the S&P 500 price–earning ratio.\(^10\) These variables were selected based on the theoretical belief that they contain information on the underlying systematic risks of the economy and on their success in previous empirical work. That is, we consider variables that reflect temporal variation in the expected cash flows, and in the discount rates that price expected cash flows.

These variables are transformed to achieve stationarity. We performed a Dickey–Fuller test (Dickey and Fuller, 1979) for the presence of unit roots in each of the variables and it was not able to reject the null hypothesis of integration against the alternative of stationarity at the 5% level for both the interest rates and the price–earning ratio. In addition, we find a remarkably high correlation between excess returns and the growth rate of these two variables. Finally, the proposed factor model will impose a common component in the autocovariance structure of the time series. Thus, rather than using the levels we use their first differences. Let $Y_t$ be the $n \times 1$ vector of financial variables exhibiting simultaneous movements in the $n$ vector of financial variables:

$$Y_{it} = \beta_i F_t + \varepsilon_{it}, \quad i = 1, \ldots, n \quad \varepsilon_{it} \sim i.i.d. N(0, \Sigma),$$

(2)

where the parameters $\beta_i$ are the factor loadings for each financial variable, which measure the sensitivity of the $i$-th series to the financial market indicator, and $\varepsilon_{it}$ is the individual idiosyncratic noise for each variable. We set the factor loading for the market excess return equal to one to provide a scale for the unobservable variable $F_t$.\(^11\)

---

\(^9\) We examined a specification in which the autoregressive process is also state dependent, $\phi_t$. However, these additional coefficients were not significant at the 1% level. We also considered other variations on the basic models, such as allowing only the factor mean to switch regimes, and holding the variance constant. The results are discussed in Section 3.

\(^10\) More information about the data is detailed in Appendix A.

\(^11\) Notice that only the scale of the factor is affected by this normalization. None of the time series properties of the dynamic factor or the correlation with its components is affected by the choice of the parameter scale. The estimation of the model, setting instead the factor variance to one, yields a highly significant estimated parameter for the excess return variable ($\hat{\beta}_s = 0.127$).
The financial market indicator corresponds to the filtered unobservable common factor among the financial variables and is produced as a non-linear combination of the observable variables \( Y_{it}, i = 1, \ldots, n \). This factor has a time-varying mean and conditional variance and, therefore, should play a role in determining the time series behavior of market risk premia.

Unlike previous regression methods, our model does not have an immediate interpretation as a forecast of excess returns. This is because the dynamic factor is produced considering contemporaneous information. As in APT models where factors include macroeconomic variables (see Chen et al., 1986), the dynamic factor \( F_t \) is a proxy for systematic-wide risks of the economy, although here it corresponds to investors’ expectation about financial markets conditioned on contemporaneous information about financial variables and the latent state of the economy.

In the proposed model, financial cycles are generated from common shocks to the dynamic factor, \( \eta_t \), and all idiosyncratic movements arise from the term \( \varepsilon_{it} \). That is, we assume that \( \eta_t \) and \( \varepsilon_{it}, for \forall i = 1, \ldots, n \) are mutually independent at all leads and lags. In order for this to hold for the price–earning ratio we use a measure from the S&P 500 with prices given at the midpoint of the month. However, in terms of the estimated model the unobservability of the Markov process may produce some sample dependence.

The parameters of the coincident model are estimated as follows: the model is cast in state-space form, where Eqs. (1) and (2) are, respectively, the transition and measurement equations. A non-linear discrete version of the Kalman filter is combined with the non-linear filter of Hamilton (1989) in one algorithm. This permits estimation of the unobserved state vector as well as the probabilities associated with the latent Markov state. The non-linear filter tracks the course of the state vector (the dynamic factor), which is calculated using only observations on \( Y_t \). It computes recursively one-step-ahead predictions and updating equations of the dynamic factor and the associated mean squared error matrices. The filter provides the conditional probability of the latent Markov state at time \( t \), which permits evaluation of the conditional likelihood of the observable variable. The filter evaluates this likelihood function, which can be maximized with respect to the model parameters using a non-linear optimization algorithm. Thus, the coincident indicator is constructed as a non-linear combination of the observable variables using information available through time \( t \), \( F_{it} \). The estimation procedure is described in detail in Appendix A.

2.2. The leading financial indicators

In order to investigate whether information about the underlying state of the economic activity helps predict financial market conditions and possibly excess returns, we build leading financial indicators as forecasts of the estimated coincident financial indicator, \( F_{it} \), in two ways: by projecting it forward in time based...
on its own past history; and by constructing an ex-post leading indicator using the 
history of the dynamic factor and other variables to predict its future values in a 
VAR system.

The multivariate approach consists of modeling the unobserved state of the 
financial cycle, \( F_t \), and some variables that forecast the state of the stock market, 
\( X_t \), as a system of autoregressive equations:

\[
\begin{align*}
F_t &= \mu_f + \beta_{f1}(L) X_{t-1} + u_f, \\
X_t &= \mu_x + \beta_{x1}(L) X_{t-1} + \beta_{x2}(L) X_{t-1} + u_x,
\end{align*}
\]

where \( u_f \) and \( u_x \) are serially uncorrelated error terms. The variables \( X_t \) were 
transformed as necessary to achieve stationarity. The leading financial indicator is 
then constructed as the one-step ahead forecast of \( F_t \), using Eq. (3). In order to 
minimize the risk of overfitting the data we choose to simply set the orders of the 
lag polynomials \( \beta_{f1}(L), \beta_{x1}(L), \beta_{x2}(L) \) to six.

3. Empirical results

3.1. The coincident financial indicator

3.1.1. Maximum likelihood results

We estimate the monthly coincident indicator using data from 1954:02 to 
1994:12. The estimates of the model obtained through numerical maximization of 
the conditional log likelihood function are presented in Table 1, which reports 
results for the dynamic factor model when both its mean and variance are assumed 
to switch regime. The log likelihood for a model with constant variance is 
\(-3803.77\), and for the switching variance model is \(-3796.82\). Although the

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Parameter</th>
<th>Parameter</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \alpha_0 )</td>
<td>0.028 (0.017)</td>
<td>( \alpha_0 )</td>
</tr>
<tr>
<td>( \alpha_1 )</td>
<td>-0.302 (0.196)</td>
<td>( \sigma^2_{\text{return}} )</td>
</tr>
<tr>
<td>( \phi )</td>
<td>0.381 (0.076)</td>
<td>( \sigma^2_{\text{volat}} )</td>
</tr>
<tr>
<td>( \sigma^2_{\text{return}} )</td>
<td>0.056 (0.013)</td>
<td>( \sigma^2_{\text{PB/E}} )</td>
</tr>
<tr>
<td>( \sigma^2_{\text{volat}} )</td>
<td>0.605 (0.232)</td>
<td>( \beta_{\text{sum}} )</td>
</tr>
<tr>
<td>LogL</td>
<td>-3798.819</td>
<td></td>
</tr>
</tbody>
</table>

Asymptotic standard errors in parentheses correspond to the diagonal elements of the inverse Hessian 
obtained through numerical calculation.
likelihood ratio test for comparing the Markov-switching model with a non-switching model has an unknown sampling distribution (see Hansen, 1993, for an explanation), one can evaluate the specification by comparing the mean and variance switching model with the mean only switching model using standard chi-squared sampling distributions. A likelihood ratio test rejects at the 1% level the hypothesis that the additional parameter for the state dependent variance is zero. We adopt the switching variance specification, with an AR(1) process for the factor and an AR(0) for the disturbances in the remainder of the paper.

The adequacy of the model specification is verified through analysis of the disturbances in the observable variables. We use the BDS test for non-linear models of Brock, Dechert, and Scheinkman (1987) to check the i.i.d. assumption for the disturbances. The diagnostic tests indicate that the specification selected is adequate for all equations. The BDS test fails to reject the i.i.d. hypothesis for the residuals. In addition, the autocorrelation functions for the disturbances $e_{it}$ are within the limit of two times their asymptotic standard deviation, and the pairwise covariance between the disturbances is nearly zero.

State 0 is characterized by a positive long run mean rate ($\mu_0 = 4.5\%$ per annum) and a low variance ($\sigma_0^2 = 0.06$), which is clearly associated with bull markets. In state 1, there is a large negative long run mean rate ($\mu_1 = -53.3\%$ per annum) with a high variance ($\sigma_1^2 = 0.60$), which describes bear markets. Notice that the volatility in bear markets is 10 times higher than the volatility during bull markets. That is, the model is capturing the empirical observation of asymmetries in the stages of stock market cycles, in which bear markets are more fidgety and associated with steep and short contractions, while bull markets are more gradual. Fig. 2 plots the behavior of the studied financial components variables across the phases of the stock market cycle.

The transition probabilities $\text{Prob}[S_t = i | S_{t-1} = i] = p_{ii}, i = 0,1$ are the probabilities of staying in a state $i$ given that financial markets are in state $i$. Their estimates are highly significant and the probability of staying in a bull market, $p_{00} = 0.97$, is higher than the probability of staying in a bear market, $p_{11} = 0.66$ and more precisely estimated. This confirms the observation that the average duration of bull markets is longer than the duration of bear markets. Notice that the last two bear markets (October 1987 and June 1990) were very short-lived, substantially reducing the estimated probability of staying in a contraction, $p_{11}$.

---

12 We also estimate a model in which neither the factor mean or variance switch regimes. The log likelihood obtained is $-3824.197$. A likelihood ratio test as proposed by Garcia (1998) based on Hansen (1993), comparing the Markov-switching and the one-state specifications strongly favors the two-state model.

13 By ‘long run’ we mean if the state no longer changed. The estimated unconditional mean is approximately zero (this is found by averaging the two ‘long run’ means by the ergodic probabilities of states 0 and 1). Further, in the sample, the financial indicator has a positive mean (see Table 2).
The factor loadings characterize the direct structural relation between the unobservable variable $F_r$ and the observable variables $Y_{it}$. The $\beta_i$ coefficients measure the sensitivity of $Y_{it}$ to a one-unit change in $F_r$. The financial market

Table 2
Statistics for the coincident financial indicator ($F_{it}$) and its components

<table>
<thead>
<tr>
<th>Statistics</th>
<th>$F_{it}$ (% annum)</th>
<th>Excess return (% annum)</th>
<th>Volatility</th>
<th>$\Delta D_{yield}$</th>
<th>$\Delta P/E$</th>
<th>$\Delta TB3$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sample mean</td>
<td>0.018</td>
<td>0.054</td>
<td>0.254</td>
<td>-0.140</td>
<td>0.118</td>
<td>0.306</td>
</tr>
<tr>
<td>Standard deviation</td>
<td>0.253</td>
<td>0.503</td>
<td>0.549</td>
<td>3.535</td>
<td>4.321</td>
<td>8.153</td>
</tr>
<tr>
<td>$\rho$ with $F_{it}$</td>
<td>1</td>
<td>0.802</td>
<td>-0.391</td>
<td>-0.838</td>
<td>0.871</td>
<td>-0.166</td>
</tr>
<tr>
<td>$\rho$ with BEAR</td>
<td>-0.450</td>
<td>-0.367</td>
<td>0.117</td>
<td>0.407</td>
<td>-0.362</td>
<td>0.025</td>
</tr>
<tr>
<td>$\rho$ with NBER</td>
<td>-0.054</td>
<td>-0.051</td>
<td>0.149</td>
<td>0.038</td>
<td>-0.016</td>
<td>-0.016</td>
</tr>
<tr>
<td>$\rho$ with JANEFF</td>
<td>0.081</td>
<td>0.044</td>
<td>0.082</td>
<td>-0.076</td>
<td>0.103</td>
<td>-0.067</td>
</tr>
<tr>
<td>$\rho$ with DEF</td>
<td>0.173</td>
<td>0.105</td>
<td>0.117</td>
<td>-0.163</td>
<td>-0.190</td>
<td>-0.219</td>
</tr>
<tr>
<td>$\rho$ with $\Delta D_{lead}$</td>
<td>0.446</td>
<td>0.289</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\rho$ with CP</td>
<td>-0.278</td>
<td>-0.144</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\rho$ with $\Delta IP$</td>
<td>0.015</td>
<td>-0.112</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

$\rho$ stands for correlation. BEAR, NBER and JANEFF are dummy variables that take the value 1, respectively, in bear markets, economic recessions, and in the month of January. DEF is the default spread (Baa–Aaa), CP is commercial paper 6-month — 3-month T-bill. $\Delta IP$, $\Delta D_{yield}$, $\Delta P/E$, $\Delta TB3$, and $\Delta D_{lead}$ are, respectively, the log first difference of the industrial production, dividend yield, price–earnings ratio, 3-month T-bill and of the leading economic indicator from the Conference Board.
Fig. 3. The coincident financial indicator, \( F_{it} \), and the excess returns.

The coincident financial indicator has the same scale as the excess return, since we set its factor loading to one, as explained above. Thus, the sign of the other factor loadings indicates the direction of the correlation of the financial variables with the financial market indicator. Both the market volatility and the changes in interest rates are negatively related to the financial market indicator. That is, in bear markets the volatility is high and the marginal change in interest rates is positive. On the other hand, the growth in the price–earning ratio is negative in bear markets. The model estimates are consistent with the actual behavior of these financial time series across stock market cycles, as observed in Fig. 2.

3.1.2. Comparison of the financial market indicator with its components

Table 2 reports some basic summary statistics for the filtered coincident financial indicator and its components. The financial market indicator is particularly highly correlated with the excess returns (0.80) and with growth in the price–earning ratio (0.87). The growth rate of the dividend yield is also highly correlated with the financial indicator (−0.84), even though it is not directly included in its construction. These high correlations indicate that the structure of the financial market indicator is not merely imposed on the financial variables by assuming large idiosyncratic noise terms.

Using Granger causality and spectral analysis we find that the indicator is coincident with excess returns and leads the price–earning ratio series and dividend yield. In addition, the time series path of the financial indicator is remarkably similar to the excess return series. Figs. 3 and 4 plot the financial

---

14 The estimated excess returns obtained from running a regression of itself on a constant and the same contemporaneous variables utilized in constructing the dynamic factor can also be interpreted as a single index model. However, the resulting estimated series is a lagging (and also a more noisy) indicator for both the excess returns and the dynamic factor.
indicator, the excess return series and a dating for financial cycles. A visual inspection of the series shows that the estimated indicator tracks very closely the turning points of excess returns and, therefore, bear and bull markets.

Table 2 also reports the correlation of the dynamic factor with some proxies for time-varying risk usually found in related literature. We use the log first difference of the leading economic index, published by the Conference Board (\(\Delta D_{lead}\)), the log first difference of industrial production (\(\Delta IP\)), the spread between commercial paper 6-month and 3-month T-bill rate (CP), and the difference between the annualized yield to maturity of Baa and Aaa bonds (DEF). These variables have been shown to relate the variation through time of risk premia to business conditions. We also use three dummy variables: one that takes a value of 1 when the economy is in a recession and 0 otherwise (NBER), a dummy which takes a value of 1 in the month of January and 0 otherwise (JANEFF), and a dummy which takes a value of 1 in bear markets and 0 otherwise (BEAR).

The estimated financial indicator is strongly related to variables that are found to predict future economic activity. In particular, the indicator presents the highest correlation with \(\Delta D_{lead}\), the BEAR dummy, CP, and DEF. On the other hand, the indicator is not correlated with the January effect dummy or with current economic conditions, as measured by the NBER-dated recession or the growth rate of

---

15 For graphical analysis, it is easier to examine smoothed version of the variables. We run a Hodrick–Prescott filtering on each variable with \(\lambda = 10\) to smooth the series. We choose the Hodrick–Prescott filtering because it does not affect turning points. The smoothed financial indicator and the excess return are plotted in Fig. 4.

16 Here we use the stock market growth cycle turning points obtained from Niemira and Klein (1994). Although our model also generates probabilities of a bear market that could be used for dating stock market cycle turning points (see Chauvet and Potter, 1998), we choose to compare our results with reference dates obtaining outside the model as a check for robustness.
industrial production ($\Delta$IP). Together all these variables explain 38% ($R^2$) of the variation in the dynamic factor indicator.

### 3.1.3. Persistence: coincident financial indicator and excess returns

The Ljung–Box statistics for 12th-order serial correlation of the monthly excess returns on the CRSP valued-weighted index are not significant, with a $p$-value of 0.35. The excess returns has a first order autocorrelation, $\hat{\rho}(1)$, of 6%. However, for the dynamic factor the Ljung–Box statistics for 12th-order serial correlation are significant at any conventional level. Also the first-order autocorrelation coefficient ($\hat{\rho}(1) = 41\%$), is seven times higher than the coefficient for the excess returns series. That is, although the estimated coincident indicator is highly correlated with monthly excess returns, it is also more persistent and, therefore, more predictable than the excess return series itself. Note that by construction the autocovariance structure of the indicator and excess returns have to be the same in population. The higher persistence as measured by the autocorrelation function in the indicator is mainly caused by its lower variance than the excess return. 17

In most related studies, there is evidence that idiosyncratic noise makes it difficult to detect the presence of predictable components. By construction, the dynamic model separates idiosyncratic noise from a common component underlying the observable financial variables. The dynamic factor is this common component and it summarizes the contemporaneous movements of financial variables that proxy for changes in risk.

We saw from Figs. 3 and 4 that our financial indicator, which is generated from common shocks to the financial variables, is somewhat smoother and more persistent than the excess return series. The coincident indicator does not have an immediate interpretation as a forecast of excess returns. This is because the dynamic factor is produced considering contemporaneous information. However, the information in the dynamic factor forecasts the financial cycle or the state of the financial conditions. For example, four lags of the coincident indicator explain 19% of the variation in the BEAR market dummy.

### 3.2. Leading financial indicators

In this section, we obtain three leading indicators based on forecasts of the coincident indicator $F_{it}$. One leading indicator is obtained by fitting an ARMA(1,1) process for the coincident financial indicator $F_{it}$, and building a one-step ahead forecast of $F_{it}$, denoted by $\hat{F}_{it}$.

Our other two leading financial indicators are constructed by jointly modeling variables $X_i$ and the unobserved state of the financial cycle, $F_{it}$, as a vector

17 The sample autocorrelation of excess return implied by Eq. (2) is $\hat{\rho}_{\text{return}} = \hat{\rho}_{\text{F},\text{return}}$ / $\hat{\rho}_{\text{return}}^2$. 


autoregressive system. Our state vector $X_t$ includes three key leading variables in addition to $F_{it}$: the excess returns series, the filtered probabilities of bear markets generated by the model (PBEAR), which can be regarded as a real-time dating of stock market cycle turning points, and the log first difference of the commodity spot market price index ($\Delta PSCCOM$). We transformed the spot index based on the result that Dickey–Fuller test with four lags cannot reject the unit root hypothesis at the 10% level.

It is interesting to notice that the excess return series provides additional useful predictive information to the dynamic factor, given that the coincident model already includes the excess returns. That is, excess return is both a coincident and a leading variable to financial market conditions.

The introduction of the filtering probabilities in the VAR system has a particular role in the analysis. This additional non-linear structure is included in order to study whether the knowledge about the state of financial conditions helps predict future $F_{it}$. Notice that the series $F_{it}$ represents the history of investors’ real-time prediction about the state of the financial conditions. We find that the inclusion of the BEAR market filtered probabilities do not affect significantly predictions of $F_{it}$, although the history of the dynamic factor and of stock market phases (PBEAR) help predicting future stock returns.

Given these findings, we consider two leading indexes of the financial coincident indicator out of the VAR system: the one-step ahead forecast of $F_{it}$, denoted by $\hat{F}_{it}^{VAR}(1)$, and the one-step ahead forecast of the excess returns variable, $\hat{r}_t^{VAR}(1)$, which also conveys information about future state of financial cycles, $F_{it}(1)$.

Table 3 presents statistics for the six-lagged autoregressive system for the full sample period, from 1954:08–1994:12, and for two subsamples, 1954:08–1979:09 and 1979:10–1994:12. The subsamples are chosen to separate the periods before and after the change in Federal Reserve operating procedures in October 1979. The level and volatility of interest rates increased considerably during the 1979–1982 period.

The four-variable autoregressive system can forecast over one-third of the variance of $F_{it}(1)$. The VAR(6) yields an $R^2$ between $\hat{F}_{it}^{VAR}(1)$ and $F_{it}(1)$ equal to 0.34 for the full sample, 0.36 for 1954:08–1979:09, and 0.33 for 1979:10–1994:12. From the system’s variance decomposition, the most significant evidence of predictive power for the coincident financial indicator comes from its past values and from the history of excess returns. The one-step ahead forecast of the

---

19 Since the VAR system uses six lags of the variables, the full sample used to compare the leading indicators is from 1954:08 to 1994:12.
Table 3
Performance of the ARMA model ($F^{\text{ARMA}}_{t\mid t}$), and of the VAR system, ($F^{\text{VAR}}_{t\mid t}$ and $\rho^{\text{VAR}}_{t\mid t}$) for forecasting $F_{t\mid t}$ and $r_{t\mid t}$. Variables in the VAR system are: the coincident indicator, $F_{t\mid t}$, excess returns ($r_{t}$), filtered probabilities of bear market, and the log first difference of the commodities spot market price index

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$R^2$ with $F_{t\mid t}$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ARMA(1,1), $F^{\text{ARMA}}_{t\mid t}$</td>
<td>0.215</td>
<td>0.201</td>
<td>0.206</td>
</tr>
<tr>
<td>VAR(6)</td>
<td>$F^{\text{VAR}}_{t\mid t}$</td>
<td>0.343</td>
<td>0.356</td>
</tr>
<tr>
<td></td>
<td>$\rho^{\text{VAR}}_{t\mid t}$</td>
<td>0.171</td>
<td>0.195</td>
</tr>
<tr>
<td>$R^2$ with $r_{t\mid t}$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VAR(6)</td>
<td>$F^{\text{VAR}}_{t\mid t}$</td>
<td>0.069</td>
<td>0.075</td>
</tr>
<tr>
<td></td>
<td>$\rho^{\text{VAR}}_{t\mid t}$</td>
<td>0.141</td>
<td>0.143</td>
</tr>
</tbody>
</table>

$F^{\text{ARMA}}_{t\mid t}$ is the one-step ahead forecast from fitting an ARMA(1,1) process to the coincident indicator $F_{t\mid t}$. $F^{\text{VAR}}_{t\mid t}$ and $\rho^{\text{VAR}}_{t\mid t}$ are, respectively, the one-step ahead forecast of the coincident indicator ($F_{t\mid t}$) equation and of the excess returns ($r_{t}$) equation from the VAR system. $R^2$ is the adjusted coefficient of correlation.

excess returns, $F^{\text{VAR}}_{t\mid t}$, explains 17% of the variance of $F_{t\mid t}$ for the full sample.

The predictions implied by linear regression models could also be interpreted as a single index and compared to the dynamic factor. Many of the single indices produced from regression models of excess returns tend to reveal the behavior of the business cycle. Since there is a vast literature on predictability in excess returns, we focus on simple existing models to compare performance in predicting stock market fluctuations. We consider a linear multivariate model for the coincident financial indicator ($F_{t\mid t}$):

$$F_{t\mid t} = \beta_r Z_{t-1} + \omega_t,$$

(5)

where we choose $Z$ to be the same lagged financial variables used to construct the financial market indicator, that is, excess returns, interest rates, price–earning ratio, and volatility. We also compare the results of an AR(1) process and a linear multivariate model for the excess returns ($r_{t}$):

$$r_{t} = \beta_r r_{t-1} + v_t,$$

(6)

$$r_{t} = \beta_r Y_{t-1} + \bar{v}_t,$$

(7)

where $Y$ includes the same variables as in $Z$ except for the excess returns.
Tables 4 and 5 report statistics and forecasting performance of our three leading indicators: $F_{it}$, $\hat{F}_{it}^{\text{VAR}}$, and $\hat{F}_{it}^{\text{ARMA}}$ from the VAR and ARMA processes; and the one-step-ahead forecasts of the alternative linear regression indexes from Eq. (5), $\hat{Z}_{it}$, from Eq. (6), $\hat{r}_{it}$, and from Eq. (7), $\hat{Y}_{it}$. Our leading indicators yield very good within-sample forecasting performance compared to alternative models. The leading indicator formed as the one-step ahead forecast of the coincident indicator equation, $F_{it}^{\text{VAR}}$, is the best in forecasting...
Table 6
Leading financial indicators and alternative models: performance in forecasting the excess returns, $r_t(1)$

<table>
<thead>
<tr>
<th>Independent variable</th>
<th>$\hat{\beta}_1 r_t$</th>
<th>$\hat{\beta}_2 Y_t$</th>
<th>$\hat{F}_{VAR}(1)$</th>
<th>$\hat{r}_{VAR}(1)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$R^2$</td>
<td>0.002</td>
<td>0.015</td>
<td>0.067</td>
<td>0.141</td>
</tr>
<tr>
<td>RMS</td>
<td>0.503</td>
<td>0.499</td>
<td>0.486</td>
<td>0.468</td>
</tr>
<tr>
<td>Theil IC</td>
<td>0.890</td>
<td>0.839</td>
<td>0.750</td>
<td>0.671</td>
</tr>
<tr>
<td>Bias</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>Var.</td>
<td>0.885</td>
<td>0.748</td>
<td>0.583</td>
<td>0.463</td>
</tr>
<tr>
<td>Cov.</td>
<td>0.115</td>
<td>0.252</td>
<td>0.416</td>
<td>0.537</td>
</tr>
</tbody>
</table>

the state of financial markets as measured by $F_{1t}(1)$. It explains one third of the variation in the coincident indicator, compared to an $R^2$ of 20% from the single index $\hat{\beta}_2 Z_t$. In addition, it gives the smallest RMS error and Theil IC in forecasting the coincident indicator. There is no apparent systematic error associated with any of the models, but the leading index $\hat{F}_{VAR}(1)$ is the best in replicating the variability of the coincident indicator, with the variance proportion of the Theil IC equal to only 0.26 compared to 0.35 of the index $\hat{\beta}_2 Z_t$.

3.2.1. Forecasts of excess return
Although our main goal is to forecast stock market fluctuations, a noteworthy result is that this autoregressive system also explains a substantial fraction of the variance of the excess returns series, compared to extant literature. Thus, we also explore the ability of the system in forecasting the excess return as well. Table 6 reports the ability of the three leading indicators in forecasting excess returns. The leading indicator formed as the one-step ahead forecast of the excess return,

![Fig. 5. Alternative model: estimated AR(1) process for excess returns, $\hat{\beta}_1 r_t$ (---) and actual excess returns (-----).](image-url)
\( \hat{r}_t^{VAR} \), explains 14% of the variation in excess returns. In contrast, a range between 6% and 12% for the monthly excess return on the CRSP value-weighted index is usually found in the literature. The coincident financial indicator and the probabilities of bear markets are especially important in this regard. The single indexes obtained from Eqs. (6) and (7), \( \hat{\beta}_r, \) and \( \hat{\beta}_Y, \) respectively, explain only 0.2% and 1.5% of the variation in excess returns. Further, the variance proportion from Theil IC is only 0.46 for \( \hat{r}_t^{VAR} \), compared to 0.88 for \( \hat{\beta}_r \), and 0.75 for \( \hat{\beta}_Y. \)

Figs. 5–7 plot excess returns and one-step ahead predictions \( \hat{r}_t^{VAR} \) from the VAR system, and the monthly indexes \( \hat{\beta}_r \) and \( \hat{\beta}_Y. \) As discussed above, the most striking distinction is that the leading indicator \( \hat{r}_t^{VAR} \) mimics closely the volatility of excess returns, and therefore the amplitude of its oscillations, while the other two indexes are much smoother. The leading indicator \( \hat{r}_t^{VAR} \) also closely tracks the turning points of the excess return series.
3.3. Out-of-sample performance

In this section we examine the model performance in two out-of-sample exercises. In the first, the parameters of the coincident indicator, \( F_{t} \), and subsequently of the leading indicator, \( F_{VAR}(1) \), were estimated using data up to 1989:12. The in-sample estimates were then used to generate out-of-sample forecasts of the filtered dynamic factor from 1990:01 to 1994:12, denoted \( \hat{F}_{89:12} \). Given the sample selection, the exercise amounts to testing the model for an unusual period, corresponding to the economic downturn in 1990 and the unparalleled slow recovery of the economy in 1991–1992. In the second exercise, the parameters were estimated using data up to 1979:09, and the out-of-sample forecasts of the filtered dynamic factor from 1979:10 to 1994:12 are denoted \( \hat{F}_{79:10} \). The idea is to verify whether the model is able to predict out-of-sample even when major events such as the change in Federal Reserve operating procedures during the 1979–1982 period and the 1987 October crash are excluded from the sample.

Table 7 reports the model out-of-sample performance for both subperiods. We compare our model performance in predicting the coincident indicator \( F_{t} \) with the alternative specification \( 5 \), built using information up to 1979:09, \( \hat{\beta}_{t}Z_{79:09} \). The coefficient of determination between each of the out-of-sample factors individually and the full sample factor is above 86%. In contrast, the \( R^{2} \) between the index \( \hat{\beta}_{t}Z_{79:09} \) and the coincident indicator \( F_{t} \) is only 23%. Also, our model presents a much smaller RMS and Theil IC compared to the alternative index.

In addition, our model exhibits a better performance in predicting out-of-sample the excess returns, \( r_{t} \). The \( R^{2} \) between the factor and the excess returns using information up to 1979:09 is 0.51, compared to only 0.02 between the single index from Eq. (7), \( \hat{\beta}_{t}Y_{79:09} \) and the excess returns. Also, the RMS and Theil IC are substantially smaller for the switching dynamic factor specification.

Figs. 8 and 9 plot the full sample filtered dynamic factor and the ones obtained using data up to 1989:12 and up to 1979:09, respectively. For both out-of-sample periods, the full sample dynamic factor and the out-of-sample factors are very similar. In particular, both out-of-sample factors predict the 1990 bear market and the ‘jagged’ period in the stock market during the uncertain period of the economic recovery, between 1991 and 1992. It is interesting to notice that the out-of-sample factor using data up to 1979:09 predicts a smaller fall in the financial markets in October 1987 compared to the full sample dynamic factor. That is, the full sample dynamic factor built from estimates that take into consideration the October 1987 Crash and the 1990 bear market reveals a deeper fall in the stock market during those periods than the out-of-sample factor.

The similarities of the financial indicators in this out-of-sample exercise highlight the robustness of the model to major switching events and the role of the Markov process in the model. For comparison, we estimate a dynamic factor in which neither the factor mean or its variance switch. The factor obtained from this
Table 7
Out-of-sample performance

<table>
<thead>
<tr>
<th>Independent variable</th>
<th>$F_{02}$</th>
<th>$r_t$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\hat{F}_{1989:12}$</td>
<td>0.983</td>
<td>0.521</td>
</tr>
<tr>
<td>$\hat{F}_{1979:09}$</td>
<td>0.865</td>
<td>0.509</td>
</tr>
<tr>
<td>$\hat{\beta}<em>Y</em>{1979:09}$</td>
<td>-</td>
<td>0.021</td>
</tr>
<tr>
<td>$\hat{\beta}<em>Z</em>{1979:09}$</td>
<td>0.227</td>
<td>-</td>
</tr>
</tbody>
</table>

|                      |         |       |
| $\hat{F}_{1989:12}$  | 0.015   | 0.275 |
| $\hat{F}_{1979:09}$  | 0.114   | 0.366 |
| $\hat{\beta}_Y_{1979:09}$ | -       | 0.522 |
| $\hat{\beta}_Z_{1979:09}$ | 0.287   | -     |

$R^2$ is the filtered factor obtained using data up to 1989:12. The out-of-sample analysis for this factor is from 1990:01 to 1994:12. $\hat{F}_{1979:09}$ is the filtered factor obtained using data up to 1979:09. The out-of-sample analysis for this factor is from 1979:10 to 1994:12. $R^2$ is the adjusted coefficient of correlation.

Theil IC

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$\hat{F}_{1989:12}$</td>
<td>0.041</td>
<td>0.388</td>
</tr>
<tr>
<td>$\hat{F}_{1979:09}$</td>
<td>0.182</td>
<td>0.402</td>
</tr>
<tr>
<td>$\hat{\beta}<em>Y</em>{1979:09}$</td>
<td>-</td>
<td>0.847</td>
</tr>
<tr>
<td>$\hat{\beta}<em>Z</em>{1979:09}$</td>
<td>0.582</td>
<td>-</td>
</tr>
</tbody>
</table>

$\hat{F}_{1989:12}$ is the filtered factor obtained using data up to 1989:12. The out-of-sample analysis for this factor is from 1990:01 to 1994:12. $\hat{F}_{1979:09}$ is the filtered factor obtained using data up to 1979:09. The out-of-sample analysis for this factor is from 1979:10 to 1994:12. $R^2$ is the adjusted coefficient of correlation.

one-state specification is less correlated with the excess returns and, as the linear specifications (5), (6), and (7), is a lagging (and also a more noisy) indicator for

![Fig. 8. Out-of-sample coincident financial indicator using data up to 1989:12, (— — —), full sample filtered coincident financial indicator (——) and bear markets (shaded area).](image-url)
both the excess returns and our switching financial indicator. This result supports the idea that investors’ perceptions about the state of financial conditions is relevant in forming forecasts of future excess returns and future market conditions.

4. Conclusions

In this paper, we consider the idea that stock markets not only react to expected future economic recessions but also fluctuates as a function of future sectoral or shorter-lived contractions in the economy not accounted for by the NBER dating. We model fluctuations in stock market at the monthly frequency as a function of investors’ reactions towards changes in unobserved market risk factors. The non-linear proxy for the market risk premia is constructed as an unobservable factor whose first moment and conditional volatility are driven by a two-state Markov variable. The estimated coincident indicator reflects common movements in the stock market driven by common shock to fundamentals and separates out idiosyncratic noise. It can be interpreted as the investors’ real-time belief about the state of financial conditions.

In terms of results, the dynamic factor that we estimate exhibits distinct bull and bear market behavior with high amplitude and a financial cycle of shorter duration than the business cycle. The factor is highly correlated with the excess return on the CRSP valued-weighted index, but is more persistent and, therefore, more predictable, than the excess return series itself. In addition, the indicator is correlated mainly with variables that forecast future economic activity, such as changes in the Conference Board leading indicator, the spread commercial paper-bill rate, and the bear/bull phases of the stock market.

We also explore an approach in which investors use their perceptions about the state of financial conditions to form forecasts of future market conditions. To investigate this, we build leading indicators as forecasts of the estimated coinci-
dent financial index. The results support this idea and the relevance of the state of stock markets in forming forecasts not only of financial conditions but also of excess returns. The leading indicators produced in this paper yield very good within-sample forecasting performance compared to alternative models. These results also hold for out-of-sample performance.

In the future, we are interested in the production of real-time forecasts. This would require a Bayesian analysis of the dynamic factor model through a Gibbs sampling algorithm. The study of the properties of real-time forecasts would also allow examination of how quickly the financial markets react to new information compared with the results obtained here. These issues should be the subject of future research.

Acknowledgements

The views expressed in this paper are those of the authors and do not necessarily reflect the views of the Federal Reserve Bank of New York or the Federal Reserve System. We thank Doug Steigerwald, Robert Engle and participants in seminars at U.S.C., U.C. Davis, U.C. Santa Barbara, U.C. San Diego, and U.C. Riverside for helpful comments.

Appendix A

A.1. Data


Excess return \( r_t \): continuously compounded value-weighted return index (VWRD from CRSP) minus continuously compounded 3-month T-bill:

\[
Excess \ return \; r_t = \log\left(\frac{P_t + D_t}{P_{t-1}}\right) - \log(1 + R_f) \\
= \log(1 + VWRD) - \log(1 + FYGM3).
\]

Proxies for volatility:

1. volatility 1 = \((r_t - \bar{r})^2\) where \( \bar{r} \) is the mean of the excess returns
2. volatility 2 = \(\sqrt{\frac{1}{2} |r_t - \bar{r}|}\)
3. volatility 3 = \(\sqrt{\left(\frac{1}{N(t)} - 1\right) \sum_{i=1}^{N(t)} \left(\frac{d_{i,t} - \bar{d}_t}{N(t)}\right)^2}\)

where \( N(t) \) is number of trading days within a month, \( d_{i,t} \) is the return on the \( i \)-th trading day of the \( t \)-th month, and \( \bar{d}_t \) is the mean of the daily returns during month \( t \).

DLEAD is composite leading index from the Conference Board, obtained from the Federal Reserve Board’s databank (August 1995).
A.1.2. From Citibase 1996

FYGM3 (TB3) 3-month T-bill rate (% per annum)
FYCP commercial paper rate (% per annum)
CP commercial paper spread = FYCP – FYGM3
FSPXE (PEAR) S&P composite common stocks: price–earnings ratio
FSDXP (D_{yield}) S&P composite common stocks dividend yield (% per annum)
FYBAAC Moody’s Baa bond yield (% per annum)
FYAAAC Moody’s Aaa bond yield (% per annum)
DEF default spread = FYBAAC – FYAAAC
PSCCOM spot market price index — all commodities: BLS and CRB (1967 = 100)

A.2. Estimation procedure

We use method developed by Kim (1994) to estimate our model.

The objective of this non-linear filter is to form forecasts of the unobserved dynamic factor, \( F_{i|t-1}^{(i,j)} \), and the associated mean squared error matrices, \( \theta_{i|t-1}^{(i,j)} \). The forecasts are based on information available up to time \( t-1 \), \( I_{t-1} = [Y_{t-1}', Y_{t-2}', \ldots, Y_{1}'] \), on the Markov state \( S_i \), taking on the value \( j \), and on \( S_{t-1} \), taking on the value \( i \):

\[
F_{i|t-1}^{(i,j)} = E(F_i|I_{t-1}, S_i = j, S_{t-1} = i),
\]

\[
\theta_{i|t-1}^{(i,j)} = E\left((F_i - F_{i|t-1}^{(i,j)}) (F_i - F_{i|t-1}^{(i,j)}') | I_{t-1}, S_i = j, S_{t-1} = i\right).
\]

The filter uses as inputs the joint probability of the Markov-switching states at time \( t-2 \) and \( t-1 \) conditional on information up to \( t-1 \), \( \text{Prob}(S_{t-2} = h, S_{t-1} = i | I_{t-1}) \); an inference about the state vector using information up to \( t-1 \), given \( S_{t-2} = h \) and \( S_{t-1} = i \), that is, \( F_{t|t-1}^{(h,i)} \); and the mean squared error matrices, \( \theta_{t|t-1}^{(i,j)} \). The outputs are their one-step updated values.\(^{20}\)

This non-linear discrete version of the Kalman filter, applied to our model is:

\[
F_{i|t}^{(i,j)} = \alpha_j + \phi F_{i|t-1}^{(i,j)} \quad \text{(prediction equations)}
\]

\[
\theta_{i|t}^{(i,j)} = \phi^2 \theta_{i|t-1}^{(i,j)} + \sigma_y^2
\]

\[
F_{i|t}^{(i,j)} = F_{i|t-1}^{(i,j)} + K_{i|t-1}^{(i,j)} \theta_{i|t-1}^{(i,j)} \quad \text{(updating equations)}
\]

\[
\theta_{i|t}^{(i,j)} = (I - K_{i|t}^{(i,j)} B) \theta_{i|t-1}^{(i,j)}.
\]

\(^{20}\)For the probabilities we use as an initial condition the probabilities associated with the ergodic distribution of the Markov chain, that is, \( \text{Prob}(S_{t-1} = h, S_{t-1} = i | I_{t-1}) = \text{Prob}(S_0 = i), i = 0, 1 \). For the state vector, its unconditional mean and unconditional covariance matrix are used as initial values, that is, \( \xi_{0|0} = E(\xi) \) and \( P_{0|0} = \text{Var}(\xi) + \sigma_y^2 \).

Where, \( K^{(i,j)}_{t-1} = \theta^{(i,j)}_{t-1} B^T \left( \tilde{Q}^{(i,j)}_t \right)^{-1} \), \( N^{(i,j)}_{t-1} = Y_t - BF^{(i,j)}_{t-1} \) is the conditional forecast error of \( Y_t \), and \( \tilde{Q}^{(i,j)}_t = B \theta^{(i,j)}_t B^T + \Sigma \) is its conditional variance. The non-linear filter allows recursive calculation of the predicted equations, given the parameters in \( \phi, B \) (the vector of factor loadings), \( \sigma^2 \), and initial conditions for \( F_{ij} \) and \( \theta_{ij} \).

In the second part of the filter, the probability terms are computed using Hamilton’s non-linear filter. The conditional likelihood of the observable variable is evaluated as a by-product of the algorithm at each \( t \), which allows estimation of the unknown model parameters. The log likelihood function is:

\[
\log f(Y_T, Y_{T-1}, \ldots | I_t) = \sum_{t=1}^{T} \log \left( \sum_{j=0}^{1} \sum_{i=0}^{1} \left( 2\pi^{-n/2} |Q^{(i,j)}_t|^{-1/2} \right) \times \exp \left( -\frac{1}{2} N^{(i,j)}_{t-1} Q^{(i,j)}_t^{-1} N^{(i,j)}_{t-1} \right) \right) \text{Prob}[S_t = j, S_{t-1} = i | I_{t-1}] .
\] (A7)

For each date \( t \) the non-linear filter computes four forecasts, which implies that at each iteration the number of cases is multiplied by 2. Thus, if the filter does not reduce the number of terms at each time \( t \), it becomes computationally unfeasible even in this simple two-state case. Kim proposed an approximation introduced through \( F_{ij} \) and \( \theta_{ij} \) for \( t > 1 \), based on the work of Harrison and Stevens (1976). The approximation consists of a weighted average of the updating procedures by the probabilities of the Markov state, in which the mixture of four Gaussian densities is collapsed, after each observation, into a mixture of two densities. That is:

\[
F_{ij}^{(i,j)} = \frac{\sum_{i=0}^{1} \text{Prob}[S_{t-1} = i, S_t = j | I_t] F^{(i,j)}_{ij}}{\text{Prob}[S_t = j | I_t]} \quad \text{(A8)}
\]

and the estimated dynamic factor is:

\[
\theta_{ij}^{(i,j)} = \frac{\sum_{i=0}^{1} \text{Prob}[S_{t-1} = i, S_t = j | I_t] \left( \theta_{ij}^{(i,j)} + \left( F_{ij}^{(i,j)} - F_{ij}^{(i,j)} \right) \left( F_{ij}^{(i,j)} - F_{ij}^{(i,j)} \right) \right) \} {\text{Prob}[S_t = j | I_t]} \quad \text{(A9)}
\]

\[
F_{ij} = \sum_{i=0}^{1} \text{Prob}[S_t = j | I_t] F_{ij}^{(i,j)} \quad \text{(A10)}
\]

References


