In Search of Leading Indicators of Economic Activity in Germany

HARM BANDHOLZ AND MICHAEL FUNKE*
Department of Economics, Hamburg University, Germany

ABSTRACT

In this paper we present two new composite leading indicators of economic activity in Germany estimated using a dynamic factor model with and without regime switching. The obtained optimal inferences of business cycle turning points indicate that the two-state regime switching procedure leads to a successful representation of the sample data and provides an appropriate tool for forecasting business conditions.

KEY WORDS business cycles; leading indicators; turning points; Markov switching; Germany

. . . the first law of forecasting: give them a prediction or give them a date, but never give them both. (Eichengreen, 2000)

INTRODUCTION

Economic policy requires a high degree of foresight because policy actions typically take effect only after a long lag. In the light of these policy needs, it is now routine for policy makers, central banks and economic research institutes to report and analyse an array of so-called leading indicators that are supposed to give a clear indication of the status of the real economy. Since economies move in cycles, the obvious question is where are we on the clock.1 An extensive literature exists which attempts to find reliable forecasting tools for the business cycle turning points, from the early landmark study by Burns and Mitchell (1946) to the more sophisticated papers of Stock and Watson (1989, 1991, 1993). Both approaches assume that the business cycle is characterized by simultaneous co-movements in many economic activities. We interpret similarities and co-movements among macroeconomic variables as a challenge to applied economic research, suggesting the development of composite leading indicators of economic activity. The composite leading indicator is a key element in an analytic system designed to anticipate the direction in which the economy is heading. Composite indicators tend to smooth out a good part of the volatility of the individual series and thereby serve as handy summary measures of the business cycle.

* Correspondence to: Prof. Dr. Michael Funke, Universität Hamburg, FB Wirtschaftswissenschaften, Von-Melle, 5, 20146 Hamburg, Germany. E-mail: funke@hermes1.econ.uni-hamburg.de.

1 To be fair, economic forecasting is even harder than weather forecasting. At least weathermen know what the temperature is right now. Economists, in contrast, have to forecast the immediate past, which is constantly being revised. Orphanides (1998) has highlighted the uncertainty due to the revisions that are made to the real-time estimates of real GDP.
An important feature of our paper is that we focus on the model’s ability to forecast turning points. Generally speaking, the major task in economic forecasting is to anticipate turning points in economic activity—points when an expansion will reach its zenith, or peak, or when a contraction will reach its nadir, or trough. A reliable method for forecasting turning points has continued to elude forecasters, despite advances in mathematical and statistical techniques. Forecasting turning points is a crucial issue since policy makers are often very much interested in answers to simple questions like ‘Is GDP going to take off next quarter?’ Another important motivation for this exercise comes from the recognition that GDP forecasts are playing an increasing role in the development of stock markets. To address this question, we use dynamic factor analysis, Kalman filtering and regime-switching techniques. Our goal is to consider a general-enough framework which permits the data to determine the stance of the business cycle and recession probabilities at some future unspecified date.

The remainder of this paper is organised as follows. In the next section the basic dynamic common factor model is described, while in the third section a dynamic Markov switching factor model is suggested. In the fourth section the alternative measures are evaluated in terms of their ability to predict future business cycle developments by means of non-nested tests. The final section concludes and suggests areas for future research.

THE DYNAMIC COMMON FACTOR MODEL

There are two main purposes for this section. First, we want to develop a composite leading indicator model for the German economy. Second, an additional goal is to evaluate the measure in terms of its ability to forecast business cycle developments. This section briefly describes the empirical methodology. It consists of constructing a common unobserved factor from a group of indicators, following the tradition of dynamic factor analysis. This common factor is assumed to represent the shared influence of the state of the economy on the leading indicator. To the best of our knowledge, this is the first attempt to construct such a composite leading indicator for Germany using the dynamic common factor model.

An important issue in the construction of a leading indicator includes the selection of variables. Using ‘economic theory screening’, we have initially tried a number of alternative specifications, not reported in the paper. In particular, we have considered new orders manufacturing, industrial production, the stock of finished goods, oil prices, the interest rate spread, new orders for residential buildings, the nominal effective exchange rate as well as a broad share price index to the empirical model as candidate variables. Because of the large number of possible specifications, we have applied various testing procedures to help with model selection. The specifications were evaluated based on their within-sample performance. It generally turned out that the turning point predictions were considerably less satisfactory for larger models than those reported in the text.

---

2 A composite indicator is said to be leading by \( x \) time periods, when information up to time \( t - x \) is required to calculate a point estimate of real GDP at time \( t \). The leading indicator is then said to have a lead time of \( x \).


4 For example, using oil prices, share prices, interest rates or exchange rates as candidate series led to various ‘false alarms’, i.e. turning points which did not occur subsequently. We have also avoided the selection of closely related series.
Moreover, dropping additional variables simplified the estimated Kalman filter and led to more robust and plausible estimates. Our final list therefore includes only two variables.\textsuperscript{5} Let $Y_{1t}$ be the fourth difference of the index of new orders total manufacturing $(1995 = 100)$ and $Y_{2t}$ the finished goods stock level ($\%$ balance).\textsuperscript{5} Unit root tests for both series suggest that one can reject the hypothesis of both variables being I(1). Borrowing from Stock and Watson (1989), we therefore consider the following bivariate dynamic factor model:

\begin{align*}
Y_{1t} &= D_1 + \gamma_{11} I_t + e_{1t} \\
Y_{2t} &= D_2 + \gamma_{21} I_t + \gamma_{22} I_{t-1} + e_{2t} \\
(I_t - \delta) &= \phi(I_{t-1} - \delta) + \omega_t, \quad \omega_t \sim i.i.d. N(0, 1) \\
e_{it} &= \psi_{i,i} e_{i,t-1} + \psi_{i,d} e_{i,t-4} + \epsilon_{it} \quad \epsilon_{it} \sim i.i.d. N(0, \sigma_{i}^2) \quad i = 1, 2
\end{align*}

where $I_t$ is the common component (leading indicator) which enters equations (1) and (2) with different weights.\textsuperscript{7} These weights $\gamma$ indicate the extent to which each series is affected by the common component, $I_t$, which arises from a single source. In (3) it is assumed that the unobserved component follows a stationary first-order autoregressive process. The autoregressive structure of the idiosyncratic component is given in (4). The main identifying assumption in the model is that $(e_{1t}, e_{2t}, I_t)$ are mutually uncorrelated at all leads and lags. Stock and Watson (1991) have shown that the parameters $D_i$ and $\delta$ are not separately identified. Therefore they have suggested to write the model in deviations from means, thus concentrating the likelihood function.

\begin{align*}
y_{1t} &= \gamma_{11} I_t + e_{1t} \\
y_{2t} &= \gamma_{21} I_t + \gamma_{22} I_{t-1} + e_{2t} \\
i_t &= \phi_i I_{t-1} + \omega_t, \quad \omega_t \sim i.i.d. N(0, 1) \\
e_{it} &= \psi_{i,i} e_{i,t-1} + \psi_{i,d} e_{i,t-4} + \epsilon_{it} \quad \epsilon_{it} \sim i.i.d. N(0, \sigma_{i}^2) \quad i = 1, 2
\end{align*}

where $y_{it} = Y_{it} - \bar{Y}_t$ and $i_t = I_t - \delta$. In general, for the Kalman filter estimation, the model is expressed in its state-space form. The latter is composed of two parts: the measurement and the transition system. The following two equations describe our model in deviations from means in a particular state-space representation.

\textsuperscript{5} The insignificance of monetary and financial variables is in harmony with the results in Fritsche and Stephan (2000). The exact data definitions and data sources are given in the Appendix.

\textsuperscript{6} We have constructed a quarterly index because monthly GDP data for Germany are not available. Another reason is that month-to-month movements in leading indicators tend to be characterized by a high noise-to-signal ratio. Stock and Watson (1989) have therefore elected to smooth the resulting monthly series to improve the indicator’s forecasting performance.

\textsuperscript{7} We have lagged $I_t$ in equation (2) in order to take into account the phase shift between $Y_{1t}$ and $Y_{2t}$. 

Copyright © 2003 John Wiley & Sons, Ltd.

In the above state-space representation, the ten rows of the transition equation describe identities. Estimation of the model will allow the unobserved components ('states') to be uncovered. The parameters and the variances of the state-space model given in (9) and (10) can be estimated using the MLE method. Given the parameters the Kalman filter recursions can be employed to obtain the time-varying parameters $\beta_t$. The Kalman filter estimation is composed of two stages, a filtering procedure and a smoothing procedure. The first is a recursive process that computes the optimal estimate for the state variables at a moment $t$ using information available up to $t-1$, minimizing the forecast error by maximum likelihood. The second procedure smoothes the obtained estimate on the basis of the information available over the whole sample period. Parameter estimates of the state-space model are reported in Table I.8

$\beta_t = F \beta_{t-1} + \nu_t$  [transition equation]

$y_{1t} = \begin{pmatrix} \gamma_{11} & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \gamma_{21} & \gamma_{22} & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} i_t \\ i_{t-1} \\ e_{1t} \\ e_{1,t-1} \\ e_{1,t-2} \\ e_{1,t-3} \\ e_{2t} \\ e_{2,t-1} \\ e_{2,t-2} \\ e_{2,t-3} \end{pmatrix}$  [measurement equation]
At this point it is useful to consider the estimates in somewhat closer detail. All parameters are consistent with that predicted by theory and are statistically significant at the 5% level. In other words, according to the results, the model seems to fit the data quite well. The estimated autoregressive coefficient $\phi$ is positive and significant. This suggests a great deal of persistence in business cycle fluctuations. In order to check the adequacy of the model specification, we analyse the disturbances $e_i$. If the model is correctly specified, then the residuals are serially uncorrelated and normally distributed. The Ljung–Box tests for residual autocorrelation are satisfactory, their results providing evidence that does not allow to reject the null hypothesis of uncorrelated distributed residuals. On the other hand, the Doornik and Hansen (1994) normality test indicates departure from normality. The table also displays various BDS tests which suggest some neglected non-linearity remaining in the disturbances. Despite these minor failures, the overall impression is that the model works very well. Given these parameter estimates, we get $i_{g_{t1}}$ and $I_{g_{t2}}$, $t = 1, 2, \ldots, T$, by running the Kalman filter.\(^9\)

The first step in our evaluation of the resulting index is to look at how well the measure does at tracking GDP over the sample period for which we have data. Following a common practice, we

---

\(^9\)Note that the resulting leading indicator is identified only up to an arbitrary choice of the initial value. We are not providing confidence bands around $I_t$ because it is well known that deriving bands around state variables in a Kalman filter estimation is far from being trivial. This fact stems from the many sources of uncertainty, which include uncertainty on the initial conditions, the sample dependence of the estimated hyperparameters and the relative flatness of the likelihood function.
consider real GDP as the relevant measure of economic activity. Economist have, however, quibbled about what ‘recession’ means. The popular definition in the USA—two consecutive quarters of GDP decline—is too crude. It might make more sense to define a recession as a period when GDP falls significantly below its potential trend. We therefore check whether our new leading indicator has any predictive power for the German output gap series (the difference between actual and trend output in percent). In order to calculate the output gap, real GDP must be detrended. Unfortunately, estimating output gaps is more an art than a science. A wide variety of both conceptual and empirical methods have been proposed to estimate potential output and make the notion of an output gap operational. The selection of a detrending procedure is a complex task because there are several possible ways of doing so, each of which is pertinent in certain circumstances and has its own implications. On the empirical side, the two most habitual methods are the Hodrick and Prescott (1997) filter, HP, and the band-pass filter. The first method has been widely used, specially in those contributions interested in the coherence of business cycles, but several authors have warned of the consequences of applying this filter. In particular, when studying co-movements in HP-filtered series, there exists one potential problem, which may be important. A common tool in the analysis of comovements of time series is the estimation of cross-correlations, but the standard errors of these estimates may be large. For HP-filtered independent random walks, Harvey and Jaeger (1993) report asymptotic standard deviations of the sample cross-correlations much higher than those obtained when at least one of the filtered series is white noise. Hence the danger of finding spurious co-moving if correct standard errors are not used. Given all these warnings and the potential problems, the GDP data will be detrended using the band-pass filter suggested by Baxter and King (1999) which has been designed to mitigate these problems. The band pass filter isolates the component of GDP that lies between 8 and 32 periods, a typical business cycle frequency range with quarterly data. The estimated output gaps are typically in the range of ±3%. Figure 1 plots the resulting new leading indicator \( I_t \) in levels against the output gap, expressed in percentages. A rise (fall) in the leading indicator points to an improvement (deterioration) in the business climate.

German business cycles have not been regular, no two cycles are identical. The main movements were a trough in 1974/5 (following the first oil price shock), a peak in 1979/80, a trough in 1981/82 (after the second sharp rise in oil prices), a peak in 1990/91 (following the exceptional event of German unification) and finally a trough in 1992/3. In recent years there has been a tendency to observe lower amplitudes. Since the beginning of the 1970s, expansions (measured from trough-to-peak) have varied in length from 12 quarters to the about 32 quarters. Recessions (measured from peak-to-trough) have ranged from 6 quarters to 12 quarters. The deepest recessions were in 1974/5 and 1992/3, when the output gaps were about 3%. Our main results with respect to the \( I_t \) variable can be summarized as follows. Most expansions and recessions have been correctly preceded by a rise or fall of our leading indicator. In other words, the index has done reasonably well in forecast-
ing the business cycle. This finding offers support for the reasonableness, reliability, and accuracy of the leading index forecasts.\textsuperscript{14}

In Figure 2 our new leading index is depicted against the well-known business climate measures of the Ifo-Institute for Economic Research in Munich and the Centre for European Economic Research (ZEW) in Mannheim. Both indicators are based upon surveys, i.e. the approach is judgemental and the indexes do not have well-defined statistical properties. In addition, movements in the indicators do not lend themselves to a straightforward interpretation. Although the indicators may be easy to understand at a conceptual level, it is not entirely clear what they are actually measuring.

A visual inspection suggests that all three indices move together synchronously, i.e. they are characterized by a very similar fluctuating pattern. In particular, the similarity between the $I$, and the $I$-\textit{Ifo}, indices is striking. In terms of explaining and forecasting, the ‘headline’ \textit{Ifo} survey measure therefore does not make a significant addition to the information already contained in the published new order and finished stock series.

Another index which has been good at detecting turning points for the UK and the American economy is the so-called ‘\textit{R}-word’ index published by \textit{The Economist} since 1992.\textsuperscript{15} Using a com-

\textsuperscript{14}A change in the direction of the composite leading index does not signal a cyclical turning point unless the movement is of significant size, duration and scope. The main value of the leading index is in signalling that either the risk of a recession has increased or that a recession may be coming to an end.

\textsuperscript{15}See \textit{The Economist}, 5 April 2001 and http://www.economist.com/displaystory.cfm?Story_ID=566293
puter database, the quarterly $R$-word index counts how many stories in British newspapers, in *The New York Times* and *The Washington Post* include the word ‘recession’. An obvious advantage is that the index is instantly available. On the other hand, when counting the word ‘recession’ it is impossible to see to which country it refers. There is also no way to distinguish between different uses of the word ‘recession’. This means that the $R$-word index is probably just a fun way of picking up the mood in an economy. But how does the index perform in Germany? Using the idea of *The Economist*, the HypoVereinsbank has recently constructed an $R$-word index for Germany counting the number of times stories appearing in the newspaper *Handelsblatt* mentioning the word ‘Rezession’\(^{16}\) (see Figure 3).

The $R$-count index, unrigorous as it is, had pinpointed the three phases of slowdown since 1986, including the recession of 1992/3. On the other hand, the $R$-word index gave a clearly wrong signal after German unification and provided only a late signal of the recession of 1992/3. In other words, the $R$-word index does not capture future economic activity in a timely manner. The poor performance of the ‘recession index’ $I-R_t$ is not puzzling: *a priori* we would have expected that this measure, which discards no new information, should not outperform the other measures.

Ideally we should expect an acceptable leading indicator model to be a model which passes specification tests for structural stability. We have therefore reestimated $I$, recursively starting in 1991:1 (Figure 4).
A DYNAMIC MARKOV-SWITCHING FACTOR MODEL

As far as the time pattern of the recursively estimated parameters is concerned, our estimation provides evidence of a consistent and hence quite robust behaviour over the period 1991:1–2001:1, as the ‘eyeball econometrics’ evidence from Figure 4 shows. The above-mentioned results suggest that the leading indicator already possesses two interesting properties, namely its seemingly acceptable forecasting properties and robustness to structural change in the sample. It therefore seems promising to pursue the analysis further, extending it to turning-point predictions since it is well known that leading indicators rarely succeed in predicting recessions. In order to calculate turning-point predictions for our leading indicator we have augmented the dynamic factor model suggested above with a bivariate Markov-switching model in the tradition of Diebold and Rudebusch (1996), Kaufmann (2000) and Kim and Nelson (1999). In other words, for the purpose of forecasting expansions and recessions, we establish a procedure which translates movements in the leading indicator into a signal about future turning points in economic activity. We again consider the model in deviation from mean form. The mean growth rate of the leading index now depends upon an unobserved dichotomous latent state variable \( S \), taking values according to a first-order Markov chain. Formally, the above model is altered as follows:

\[ \begin{align*}
\text{17 Neftci (1982) was the first to develop a statistical methodology to transfer movement in the composite leading indicator into a measure of the probability of a cyclical turning point. The BDS tests in Table I provide some empirical evidence in favour of such a non-linear specification.} \end{align*} \]
\[ y_t = \gamma_{11} \delta_t + \epsilon_{1t} \]  
\[ y_{2t} = \gamma_{21} \delta_t + \gamma_{22} \delta_{t-1} + \epsilon_{2t} \]  
\[ (1 - \phi L)(i_t - \mu_{S_t}) = \omega_t, \quad \omega_t \sim i.i.d. N(0, 1) \]  
\[ \epsilon_{it} = \psi_{i1} \epsilon_{i,t-1} + \psi_{i4} \epsilon_{i,t-4} + \epsilon_{it}, \quad \epsilon_{it} \sim i.i.d. N(0, \sigma_i^2) \quad i = 1, 2 \]  
\[ \mu_{S_t} = \mu_0 (1 - S_t) + \mu_1 S_t, \quad S_t = \{0, 1\}, \mu_1 > 0 \]  
\[ \Pr[S_t = 1|S_{t-1} = 1] = p, \quad \Pr[S_t = 0|S_{t-1} = 0] = q \]  

where \( L \) is the lag operator. A state-space representation is given by:

Figure 4. Recursive estimates of the parameters, 1991:1–2002:2
Economic Activity in Germany

\[
\begin{pmatrix}
    y_{1t} \\
    y_{2t}
\end{pmatrix} =
\begin{pmatrix}
    \gamma_{11} & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
    \gamma_{21} & \gamma_{22} & 0 & 0 & 0 & 1 & 0 & 0 & 0
\end{pmatrix}
\begin{pmatrix}
    i_t \\
    e_{1t} \\
    e_{1,1-1} \\
    e_{1,2-1} \\
    e_{1,3-1} \\
    e_{2t} \\
    e_{2,1-1} \\
    e_{2,2-1} \\
    e_{2,3-1}
\end{pmatrix}
\]

\[ j_t = H^* \beta_t \]

\[
\begin{pmatrix}
    i_t \\
    i_{t-1} \\
    e_{1t} \\
    e_{1,1-1} \\
    e_{1,2-1} \\
    e_{1,3-1} \\
    e_{2t} \\
    e_{2,1-1} \\
    e_{2,2-1} \\
    e_{2,3-1}
\end{pmatrix} =
\begin{pmatrix}
    \mu_{St} - \phi \mu_{St-1} \\
    0 \\
    0 \\
    0 \\
    0 \\
    0 \\
    0 \\
    0 \\
    0 \\
    0
\end{pmatrix}
+ \begin{pmatrix}
    \phi & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
    1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
    0 & 0 & \psi_{11} & 0 & 0 & \psi_{14} & 0 & 0 & 0 & 0 \\
    0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
    0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
    0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
    0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
    0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\
    0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1
\end{pmatrix}
\begin{pmatrix}
    i_{t-1} \\
    e_{1,1-1} \\
    e_{1,2-1} \\
    e_{1,3-1} \\
    e_{2t} \\
    e_{2,1-1} \\
    e_{2,2-1} \\
    e_{2,3-1}
\end{pmatrix}
+ \begin{pmatrix}
    0 \\
    0 \\
    0 \\
    0 \\
    0 \\
    0 \\
    0 \\
    0 \\
    0 \\
    0
\end{pmatrix}
\]

\[ \beta_t = R^*_s + F^* \beta_{t-1} + \nu_t \]

To make the turning point model tractable, the econometrician must specify a stochastic process for the state variable $S_t$. The above model assumes that two different regimes exist in the data-generating process characterized by either expansions ($S_t = 1$) or recessions ($S_t = 0$), and that the process is subject to discrete (sporadic) regime shifts governed by a two-state Markov process. This feature of regime-dependence implies that the model is able to mimic the asymmetric behaviour of macroeconomic variables over the business cycle we often observe. The mean parameter $\mu_S$ refers to the Markov-switching deviation of the leading indicator from its long-run growth rate $\delta$. The probability of being in each state is determined by the transition equation (14).

The calculated probability can therefore be viewed as a direct prediction of the underlying state of the economy in the near future. To estimate the model, it is necessary to make inferences about both the unobserved common factor and the latent Markov state. Hamilton’s (1989) paper popularized the use of bivariate Markov regime switches, but the methodology precludes the estimation of multivariate unobservable dynamic models. On the other hand, the dynamic factor models proposed by Stock and Watson (1989, 1991) is governed by a linear stochastic process and can therefore be estimated using the Kalman filter. The non-linearity in the transition equation of the model given by
(9) to (16), however, implies that the usual Kalman filter cannot be applied directly. The following estimation procedure therefore consists of a combination of Hamilton’s (1989) algorithm and a non-linear version of the Kalman filter, as proposed by Kim (1994). This permits estimation of the unobserved factor as well as the probabilities associated with the latent Markov state.\textsuperscript{18} Kim (1994) also provides a fast approximation algorithm for the full sample smoother which substantially reduces computational time.\textsuperscript{19} Table II contains information on how well the augmented model fits the historical data.

The empirical results provide support for the two-state Markov switching model. Note first that many of the same patterns observed in Table I are also present in Table II. There is a significantly positive deviation from Trend in state 1 and a significantly negative deviation from Trend in state 0. The estimated transition probabilities are highly significant and the probability of staying in expansion, $p$, is higher than the probability of staying in a recession, $q$. This confirms previous findings that the average duration of expansions is larger than the duration of recessions. The reason is that the occurrence of recessions is irregular, which means that expansions usually last longer than contractions. The expected duration for expansion $[1/(1-p)]$ and recession $[1/(1-q)]$ implied by the model is 16.64 and 4.17 quarters, respectively.

Figure 5 plots the composite leading indicators implied by the dynamic factor model without ($I_t$) and with ($I_{t-MSt}$) Markov switching, respectively. Both indices show a very similar pattern regarding amplitude, timing, and duration of fluctuations. Another graphical way of evaluating the two series is the scatter diagram in Figure 6. If $I_t$ and $I_{t-MSt}$ were exactly the same, they would lie on the

\begin{table}[h]
\centering
\caption{Parameter estimates of the dynamic factor model with Markov switching: sample period: 1971:1–2002:2}
\begin{tabular}{lcc}
\hline
Parameters & Estimates & Asymptotic $t$-values \\
\hline
$\gamma_{11}$ & 0.186 & 4.5 \\
$\gamma_{21}$ & -0.113 & 4.5 \\
$\gamma_{22}$ & -0.218 & 5.0 \\
$\mu_0$ & -1.563 & 1.8 \\
$\mu_1$ & 0.395 & 2.6 \\
$\phi$ & 0.773 & 10.0 \\
$\psi_{11}$ & 0.584 & 7.6 \\
$\psi_{14}$ & -0.332 & 4.7 \\
$\psi_{51}$ & 0.846 & 17.1 \\
$\psi_{54}$ & -0.477 & 13.8 \\
$p$ & 0.940 & 25.2 \\
$q$ & 0.761 & 3.8 \\
$\sigma_{1}^2$ & 0.348 & 7.5 \\
$\sigma_{2}^2$ & 0.012 & 1.1 \\
\hline
Log-likelihood & -87.95 \\
\hline
\end{tabular}
\end{table}

\textsuperscript{18}McCulloch and Tsay (1994) have demonstrated that the posterior probabilities of recession are sometimes quite sensitive to the number of lags on the autoregressive component of the common component. This implies that the correct specification of the number of terms to include is important when estimating regime-switching models.

\textsuperscript{19}An intuitive explanation of the computational burden is the following. Each iteration of the Kalman filter produces a twofold increase in the number of cases to be considered. Since $S_t$ takes on two possible values in each time period, there would be $2^T$ possible paths to consider in evaluating the conditional log likelihood (see Kim, 1994).
Figure 5. The leading indicators with ($I$-$MS$) and without Markov switching ($I_t$)

Figure 6. Leading indicators $I_t$ and $I$-$MS_t$
45° line shown. As Figure 6 shows, the leading indicators were quite close. However, there seems to be a tendency that more points lie below the 45° line. This indicates that I-MS, forecasted on average slightly lower gaps than did $I_t$. Contemporaneous correlation between the two series is 0.99.

Next we have calculated the posterior recession probabilities from the Markov-switching model. Figure 7 depicts the probability that the German economy was in the recession regime. The general impression is that the posterior recession probabilities (right scale) are in close agreement with the German business cycle when the cut-off rate is set to 0.50, i.e. any probability that exceeds 0.50 is counted as a signal for a recession. How can the empirical results be interpreted? As shown in Figure 7, our leading indicator correctly pointed to storm after the first and the second oil price shock and anticipated the recession in 1992/3. Another interesting feature is that the ‘mini-recessions’ of 1995/6 and 1998/9 are not classified as ‘regular’ recessions. Thus, overall the I-MS leading indicator appears to score relatively well in forecasting German business cycles and therefore the I-MS decision rule approach is a promising development.20 Finally, in 2000 the recession probability started to flash red and Germany was in a mild recession most of 2001.21

The evidence presented so far only proves one thing: from an ex-post point of view the dynamic factor Markov-switching model correctly anticipates recessions. Although this result is encouraging, we should not forget that there is a growing consensus in the recent forecasting literature that the in-sample evaluation of leading indicators is likely to be misleading in analysing the predictive power of an indicator time series. In other words, forecasters agree that forecasting methods should also

---

20 This result confirms the results in Kim (1994) and Hamilton and Pérez-Quiros (1996) who have presented evidence for the predictive power of the composite indicator model.
21 The risk of a double-dip recession by the end of 2002, however, is very small although the recent recovery in Germany looks rather anaemic.
Figure 8. Full-sample and real-time recession probabilities for the dynamic Markov-switching factor model
be assessed for accuracy using out-of-sample tests rather than in-sample tests since the performance of models on data outside that used for estimation remains the touchstone for its utility in applications. Hence, we have re-estimated the model recursively starting in 1990:4. This procedure is designed to reproduce the reality that forecasters have data only up to the starting point of their forecast. Moreover, tracking the performance of the Markov switching indicator in different years allows us to assess the robustness of its forecasting power. Figure 8 shows the probabilities of a recession given by the recursive estimates as well as the full sample results. For example, consider the recession probability in 1999:2. To compute this probability, the model is estimated using data through 1999:2, at which point the leading indicator is calculated. Moving forward two quarters, the coefficients are recalibrated using data through 1999:4 and the next leading indicator is computed. This procedure generates a series of ‘real-time’ leading indicators. One feature is particularly noteworthy. The recursively estimated real-time recession indices nearly match the full-sample results, i.e. it makes very little difference to the prediction of turning points when real-time data are used.\(^22\)

**NON-NESTED TESTS**

Indeed, when two or more leading indicators are under consideration, it is a common approach to attempt to reject one model in favour of another. In other words, in this section we would like to evaluate, in a systematic manner, the forecasting performance of different leading indicators. To test whether \(I, I-MS\), and/or other leading indicators have leading indicator properties for the business cycle we first test for Granger causality, i.e. we test whether past values of the output gap along with the leading indicators better ‘explain’ the cyclical movement of GDP than past values of the output gap \((GAP)\) alone. This of course does not imply that the leading indicators cause \(GAP\). Instead it is possibly reflecting expectations as to where GDP might be headed. In assessing the leading indicator properties, the Granger-causality test can be supplemented by non-nested tests. Such tests can be used to evaluate whether a candidate variable gives a useful contribution in forecasting, relative to a variable chosen as a benchmark. We first specify dynamic equations linking the output gap to the respective leading indicator variable. A conventional general to specific testing-down procedure was followed until parsimonious equations for the output gap were achieved. The resulting OLS results are reported in Table III. As many other researchers, we have found that the lagged output gaps helps to explain variations in economic activity. This result confirms the presence of inertia and/or duration effects in the growth pattern of economic activity. After the testing-down procedure, we are therefore left with four rival statistically significant leading indicators. The next step is to investigate whether any of these measures is superior in predicting the quarterly output gap. A straightforward procedure is to adopt non-nested tests to discriminate among the alternative gap measures.\(^23\) Alternative pairwise non-nested tests are reported in Tables IV and V. The non-nested tests are referred to as Cox’s (1961) and Pesaran’s (1974) \(N\)-test, Davidson and MacKinnon’s \(J\)-test

---


\(^23\) In appraising the comparative performance of the various leading indicators, it must be borne in mind that they are not made at precisely the same times, and therefore the information set available to the forecasters is not identical. This also opens up the possibility of ‘herding behaviour’, i.e. the first published leading indicator can exert a direct influence on later survey measures. See Trueman (1994), for example.
and Godfrey and Pesaran’s (1983) W-test. All tests are null hypothesis-specific. The null and alternative hypothesis must be reversed and the test repeated for the results to be conclusive. Thus it is perfectly possible, that both \( H_0 \) and \( H_1 \) are accepted, or that they are both rejected. If neither model is accepted (rejected), the test is inconclusive, meaning that the data do not allow discrimination between the two competing hypotheses.

What does our analysis reveal about the predictive power of the various indicators? Overall, the pairwise non-nested test statistics indicate that \( I_t \) and \( I-MS_t \) have greater predictive value than the \( I-IFO_t \) index while the results for the \( I-ZEW_t \) index are inconclusive. Our (necessarily tentative)

---

Notes: All equations have been estimated by OLS. The dependent variable is \( GAP_t \). t-values in parentheses; \( R^2 \) denotes the coefficient of determination; \( LM(n) \) denotes the Lagrange Multiplier test for \( n \)th-order serial correlation; \( JB \) is the Jarque–Bera test for residual normality. \( RESET \) is the RESET test (using the squared predictions) for omitted variables and incorrect functional form. The \( p \)-values are given in parentheses. The \( I-R_t \) indicator has not been considered because it turned out to be a coincident or even lagging indicator of economic activity.

---

### Table III. Baseline models for the German output gap

<table>
<thead>
<tr>
<th>Regressor</th>
<th>(a)</th>
<th>(b)</th>
<th>(c)</th>
<th>(d)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>-0.368</td>
<td>-0.354</td>
<td>-0.335</td>
<td>-0.287</td>
</tr>
<tr>
<td>( (4.5) )</td>
<td>( (4.3) )</td>
<td>( (4.0) )</td>
<td>( (2.5) )</td>
<td></td>
</tr>
<tr>
<td>( GAP_{t-1} )</td>
<td>0.445</td>
<td>0.455</td>
<td>0.463</td>
<td>0.510</td>
</tr>
<tr>
<td>( (5.4) )</td>
<td>( (5.5) )</td>
<td>( (5.6) )</td>
<td>( (4.2) )</td>
<td></td>
</tr>
<tr>
<td>( GAP_{t-4} )</td>
<td>-0.144</td>
<td>-0.131</td>
<td>-0.110</td>
<td>-0.040</td>
</tr>
<tr>
<td>( (2.1) )</td>
<td>( (1.9) )</td>
<td>( (1.5) )</td>
<td>( (0.3) )</td>
<td></td>
</tr>
<tr>
<td>( I_{t-1} )</td>
<td>0.004</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>( (4.6) )</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( I-MS_{t-1} )</td>
<td>—</td>
<td>0.004</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>( (4.3) )</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( I-IFO_{t-1} )</td>
<td>—</td>
<td>—</td>
<td>0.003</td>
<td>—</td>
</tr>
<tr>
<td>( (4.0) )</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( I-ZEW_{t-1} )</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>0.003</td>
</tr>
<tr>
<td>( (2.5) )</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Diagnostics:

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>( R^2 )</td>
<td>0.48</td>
<td>0.48</td>
<td>0.47</td>
<td>0.44</td>
</tr>
<tr>
<td>( LM(1) )</td>
<td>0.03</td>
<td>0.05</td>
<td>0.71</td>
<td>1.12</td>
</tr>
<tr>
<td>( (0.86) )</td>
<td>( (0.83) )</td>
<td>( (0.40) )</td>
<td>( (0.29) )</td>
<td></td>
</tr>
<tr>
<td>( LM(8) )</td>
<td>12.24</td>
<td>13.24</td>
<td>10.70</td>
<td>5.01</td>
</tr>
<tr>
<td>( (0.14) )</td>
<td>( (0.10) )</td>
<td>( (0.22) )</td>
<td>( (0.76) )</td>
<td></td>
</tr>
<tr>
<td>( RESET )</td>
<td>0.00</td>
<td>0.00</td>
<td>0.30</td>
<td>1.21</td>
</tr>
<tr>
<td>( (0.96) )</td>
<td>( (0.99) )</td>
<td>( (0.58) )</td>
<td>( (0.28) )</td>
<td></td>
</tr>
<tr>
<td>( JB )</td>
<td>73.35</td>
<td>71.40</td>
<td>96.74</td>
<td>17.58</td>
</tr>
<tr>
<td>( (0.00) )</td>
<td>( (0.00) )</td>
<td>( (0.00) )</td>
<td>( (0.00) )</td>
<td></td>
</tr>
</tbody>
</table>

Notes: All equations have been estimated by OLS. The dependent variable is \( GAP_t \). t-values in parentheses; \( R^2 \) denotes the coefficient of determination; \( LM(n) \) denotes the Lagrange Multiplier test for \( n \)th-order serial correlation; \( JB \) is the Jarque–Bera test for residual normality. \( RESET \) is the RESET test (using the squared predictions) for omitted variables and incorrect functional form. The \( p \)-values are given in parentheses. The \( I-R_t \) indicator has not been considered because it turned out to be a coincident or even lagging indicator of economic activity.

---

24 Godfrey and Pesaran (1983) have shown that the traditional Cox-test overrejects in small samples. This led them to derive a modified version of the Cox-statistic which is used in this paper. Both statistics have the same asymptotic properties, so no ambiguities should arise.

25 It is worth mentioning that in evaluating the \( ZEW \) survey indicator caution is in order because the sample period used for the non-nested tests is relatively short, due to the short history of the \( ZEW \) indicator.
Table IV. Non-nested tests of $I_t$ versus $I$-Ifot $(1972:1$–$2002:2)$ and $I$-ZEW, $(1992:2$–$2002:2)$

<table>
<thead>
<tr>
<th>Test statistic</th>
<th>$I_t$ versus $I$-Ifot</th>
<th>$I$-Ifot versus $I_t$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$N$</td>
<td>-0.81</td>
<td>-2.44</td>
</tr>
<tr>
<td></td>
<td>(0.42)</td>
<td>(0.02)</td>
</tr>
<tr>
<td>$W$</td>
<td>-0.69</td>
<td>-2.14</td>
</tr>
<tr>
<td></td>
<td>(0.49)</td>
<td>(0.03)</td>
</tr>
<tr>
<td>$J$</td>
<td>0.74</td>
<td>1.99</td>
</tr>
<tr>
<td></td>
<td>(0.46)</td>
<td>(0.05)</td>
</tr>
</tbody>
</table>

Schwarz’s BC criteria $I_t$ versus $I$-Ifot = 1.74 favours $I_t$

<table>
<thead>
<tr>
<th>Test statistic</th>
<th>$I_t$ versus $I$-ZEW,</th>
<th>$I$-ZEW, versus $I_t$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$N$</td>
<td>-2.05</td>
<td>-2.36</td>
</tr>
<tr>
<td></td>
<td>(0.04)</td>
<td>(0.02)</td>
</tr>
<tr>
<td>$W$</td>
<td>-1.33</td>
<td>-1.53</td>
</tr>
<tr>
<td></td>
<td>(0.19)</td>
<td>(0.13)</td>
</tr>
<tr>
<td>$J$</td>
<td>1.33</td>
<td>1.46</td>
</tr>
<tr>
<td></td>
<td>(0.19)</td>
<td>(0.14)</td>
</tr>
</tbody>
</table>

Schwarz’s BC criteria $I_t$ versus $I$-ZEW, = 0.21 favours $I_t$

Table V. Non-nested tests of $I$-MS, versus $I$-Ifot $(1972:1$–$2002:2)$ and $I$-ZEW, $(1992:2$–$2002:2)$

<table>
<thead>
<tr>
<th>Test statistic</th>
<th>$I$-MS, versus $I$-Ifot</th>
<th>$I$-Ifot versus $I$-MS,</th>
</tr>
</thead>
<tbody>
<tr>
<td>$N$</td>
<td>-0.97</td>
<td>-2.12</td>
</tr>
<tr>
<td></td>
<td>(0.33)</td>
<td>(0.05)</td>
</tr>
<tr>
<td>$W$</td>
<td>-0.83</td>
<td>-1.86</td>
</tr>
<tr>
<td></td>
<td>(0.41)</td>
<td>(0.06)</td>
</tr>
<tr>
<td>$J$</td>
<td>0.87</td>
<td>1.77</td>
</tr>
<tr>
<td></td>
<td>(0.38)</td>
<td>(0.08)</td>
</tr>
</tbody>
</table>

Schwarz’s BC criteria $I$-MS, versus $I$-Ifot = 1.21 favours $I$-MS,

<table>
<thead>
<tr>
<th>Test statistic</th>
<th>$I$-MS, versus $I$-ZEW,</th>
<th>$I$-ZEW, versus $I$-MS,</th>
</tr>
</thead>
<tbody>
<tr>
<td>$N$</td>
<td>-2.21</td>
<td>-2.24</td>
</tr>
<tr>
<td></td>
<td>(0.03)</td>
<td>(0.03)</td>
</tr>
<tr>
<td>$W$</td>
<td>-1.41</td>
<td>-1.43</td>
</tr>
<tr>
<td></td>
<td>(0.16)</td>
<td>(0.15)</td>
</tr>
<tr>
<td>$J$</td>
<td>1.38</td>
<td>1.40</td>
</tr>
<tr>
<td></td>
<td>(0.17)</td>
<td>(0.16)</td>
</tr>
</tbody>
</table>

Schwarz’s BC criteria $I$-MS, versus $I$-ZEW, = 0.02 favours $I$-MS,

Notes: $p$-values are given in parentheses. The decision rule always is to reject $H_0$ in favour of $H_1$ for large values of the test statistics. The different sample periods are given in the headline.
conclusion therefore is that our leading indicators do better at tracking the German business cycle than the well-known Ifo index. There is therefore good reason to be wary of placing too much confidence in forecasts of economic activity that rest exclusively on survey climate indicators.

CONCLUSIONS

Both policy makers and the private sector maintain a keen interest in understanding the state of business affairs and the most likely path the economy will take over the near future. In this paper we have therefore constructed a composite leading economic indicator of economic activity in Germany using the dynamic factor modelling approach over the 1971–2002 period. In order to refine our forecast of future recessions and expansions, we then outlined a methodology for the construction of a recession and expansion index. The historical performance of our indices suggests that the information they convey about the timing and the likelihood of a recession or expansion shows great promise and represents an improvement over the information offered by other ‘headline’ survey measures which are closely monitored by the financial and popular press. Another advantage of our methodology is that forecasts can be made quickly and are based on a small set of series only. Such assessment should of course trigger further research, part of it being quite straightforward, namely a comparison exercise for Euroland. We leave this topic for future research.

APPENDIX: DATA USED

Time series for unified Germany exist only as of 1991. We have therefore joined the two series (pan-German, West German) after having re-scaled the ‘old historical’ data to the ‘new’ German series. In addition, due to the recent changeover to ESA95, it was necessary to backdate the National Account series. The ‘old’ series were re-based and joined to the ‘new’ series, applying the same method used to overcome the German unification problem. All series were seasonally adjusted (if necessary) using X12. Quarterly data are simple averages of monthly data.

<table>
<thead>
<tr>
<th>Time series</th>
<th>Details and source</th>
</tr>
</thead>
<tbody>
<tr>
<td>New orders manufacturing (index 1995 = 100)</td>
<td>Source: Statistisches Bundesamt, Fachserie 4</td>
</tr>
<tr>
<td>Finished goods stock level (% balance)</td>
<td>Source: Ifo-Institute (Munich); the survey measure (published monthly) is presented in the form of the difference (hence the term balance) between the percentage of firms which have noted an increase an those which have reported a decrease of finished goods stock levels.</td>
</tr>
</tbody>
</table>

Only finally revised data have been used for the estimates. Therefore, the exercise is a ‘pseudo’ forecasting exercise because, strictly speaking, due to subsequent revisions the data used here were not available to a ‘real-time’ forecaster.26

26The GDP and therefore the output gap measure is normally more affected by frequent revisions, while the two variables used to construct the leading indicator are the subject of infrequent, but perhaps larger, changes in definitions and classifications. GDP growth is often overstated at first when an economy slows sharply. This is because initial GDP numbers includes estimates for some components of GDP that are derived by extrapolating recent trends.
REFERENCES


**Authors’ biographies:**

**Harm Bandholz** is a Research and Teaching Assistant at the Department of Economics, Hamburg University, Germany.

**Michael Funke** is Professor of Economics at the Department of Economics, Hamburg University, Germany.

**Authors’ address:**

**Harm Bandholz** and **Michael Funke**, Department of Economics, Hamburg University, Von-Melle-Park 5, 20146 Hamburg, Germany.