OSLO AND AKERSHUS UNIVERSITY COLLEGE OF APPLIED SCIENCES

Comparing sentiment and behavioral based leading indexes for industrial production in Germany: A novel running local test

Yunus Yilmaz, Knut Lehre Seip, and Michael Schröder
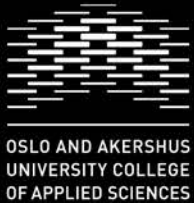
Comparing sentiment and behavioral based leading indexes for industrial production in Germany: A novel running local test

Yunus Yilmaz,[1] Knut Lehre Seip,[2] and Michael Schröder[3]

Abstract

We use a novel method to compare two survey based leading indexes and two behavioral based index to industrial production, IP, Germany 1991-2015 . The sentiment based ifo – index (managers) performs best in predicting the general changes in IP (-0.596, -1.0 being best). The ZEW – index (financial experts) is close (-0.583), and is better in predicting 6 recessions and 5 recoveries. On a third place comes, somewhat unexpectedly, the unemployment index (-0.564) and lastly comes order flow, OF (-0.186). The four indexes all scored low during time windows around 1997 and 2005. Both periods correspond to anomalous episodes in German economy.

Keywords: Leading indexes, industrial production, prediction skill, survey based indexes, Germany

JEL classification E3, C2, G2

[1] Ruhr-Universität Bochum, faculty of management and economics, Universitätsstraße 150 44801 Bochum, E-mail: Yunus Yilmaz <yunus.yilmaz091@gmail.com>
[2] Oslo and Akershus University College for Applied Sciences. Pilestredet 35, Oslo, Norway. E-mail knut.lehre.seip@hioa.no
[3] Centre for European Economic Research (ZEW), Mannheim, Germany. E.Mail: schroeder@zew.de

1. INTRODUCTION

We compare the accuracy and timing of four candidate indexes in Germany for the period January 1991 to September 2016 with a novel running (rolling) local application of the leading –lagging, LL – method developed by Seip and McNown (2007). The method estimates leading – lagging, LL-strengths, rolling cycle times and rolling phase shifts. We also identify periods where LL – relations become contrary to expectations. The LL –method offers a rapid and detailed screening of component series for the construction of composite leading indicators. Lastly, we suggest ways to improve the learning set for e.g., AR forecasting, based on movements in the paired leading and target series.

We examine the three leading indexes and one coinciding/ lagging index to Industrial production, IP, with respect to their leading- lagging, LL- relationships. The ifo business climate index is a survey-based sentiment index for Germany from the ifo institute for economic research, (IFO 2016). The ZEW business cycle index is from the ZEW (Centre for European Economic Research) (ZEW 2016). The third index is a series for "order flow", OF, that we treat as a behavioral based index. The fourth index is the standard unemployment index, that is a coinciding/ lagging index to IP (Enders 2010; Heij et al. 2011; Balcilar et al. 2013). The basic idea for this study was conceived by Yilmaz (2016).

In the present application, we calculate running average LL – strength over 3 observations, and then over a longer time window (9 to 13 observations) to obtain a significance measure. By doing this, we can identify dates where LL- strength becomes weaker or stronger, or change sign. There are several alternative methods for identifying leading- lagging relations, e.g., Hüfner and Schröder (2002) on cross correlation and Granger causality tests and Carstensen et al. (2011) on rolling regressions. Forni et al. (2001) uses a spectral density algorithm to identify cycle lengths of the EURO coincident indicator. However, to our knowledge, the present LL- method is the only method that allows calculation of running average LL- strength that does not require stationary cycle times even over short time spans. Thus, the method can be used both to measure the strength of general business cycle phases and for turning point analysis, e.g. as in Levanon et al. (2015).

OECD identifies 6 recession periods for Germany during this period 1991 to 2016. A recession is defined as the period between the peak and the following through. We discuss the performance of the indexes for the whole period, for periods in front of the recession periods and for periods before the recovery. However, the first period 1991:4 to 1993:8 identified by OECD did not show negative values for IP when detrended.

IP –growth. Since there are evidence that leading indexes may be better predict the growth in IP, we also took its first derivative and compared the leading indexes to this series.

There are 5 categories of indexes that relate to a target index in economic forecasting literature, (Abel et al. 1998). All categories refer to a common cycle time, λ, for a pair of cyclic time series. Before and after will often refer to peaks or troughs in the series, but may also refer to slopes, or to any series of three consecutive and synoptic points on the paired series. Although there is no consistent definition of the categories, a categorization could be as follows: A leading index, LI, is less than ½ λ before the target series. A lagging, or trailing index, TI; is less than ½ λ after the target series, a pro cyclic, or coincident index, CI, is leading or trailing the target cycle with less than ¼ λ. A counter cyclical index is more than ¼ λ from the target series. A fifth category is called acyclic and do not show a consistent pattern. The two first categories will show opposite rotating trajectories if the paired time series are plotted in phase space. The two next categories will show a positive and a negative regression coefficient for the scatter plots if plotted in phase space. (Seip and Grøn 2017)

Hypotheses

We develop and test five hypotheses for the relationships between the three leading indexes and the lagging index and with regard to their performance. The present study also includes an assessment of the availability and timing of the data for the leading indexes.

Firstly, we hypothesize that the survey- or sentiment-based indexes will perform better than order flow and employment in predicting IP as the latter two indexes will exhibit only a small lead, if at all, whereas the sentiment indexes are intended by construction to have a lead of about 6 months.

Secondly, we hypothesize that the leading indexes will behave better (more accurate and giving longer leading time) during normal business cycles than before recessions or recovery periods. This should apply in particular to the "great" recession in 2008, a period that was rather hard to predict by the conventional leading indicators. (Ferrara et al. 2015).

Thirdly, we hypothesize that unemployment, which is most likely to be lagging variable to economic growth (Banerji et al. 2006), will perform well during the same time windows where the leading indexes perform well. The rationale is that unemployment, as a lagging index, may confirm the more complex economic reasons for an increase or a decrease in the business cycle, e.g., Granger (1989), and maximize the intensity of turning points in a composite leading indicators (OECD 2012).

Fourthly, we hypothesized that smoothed time series will give better predictions than raw (and therefore noisy) series. *Lastly*, we hypothesize that IP- growth will be better predicted than IP itself.

We show that the best overall leading index is the ifo- index which is based on company manager's forecasts (-0.596, -1.0 is best), next comes the ZEW – index based on the forecasts of financial analysts (-0.583). Unemployment, although negatively associated with IP, is also a leading index to IP (-0.564). Lastly comes order flow, OF (-0.186). However, the ZEW index is in 83 % of the time a leading index to the ifo index. We also identify two periods, one around 1997 and one around 2005 where the ifo and ZEW indexes performed badly. These two periods do not correspond to reported recessions, recoveries, or structural breaks in the German economy, but still appear to correspond to anomalous events. Both industrial production, IP, and the leading indexes were smoothed with the locally weigthed LOESS algorithm corresponding to about 24 months running window. With this smoothing, we obtained an average cycle time of 20-30 months for all pairs of variables.

We organize the rest of the paper as follows. We present the two survey based and the two behavioral based indexes in section 2 on materials. In section 3, we present the methods used in the study

with emphasis on the running window leading – lagging, LL- strength method. In section 4, we present the results and in section 5 we discuss data availability, prediction power and prediction lead times. Section 6 concludes.

2. MATERIALS

Figure 1a shows the five time series in their raw format, and Figure 1b and c shows the time series detrended and smoothed with the LOESS smoothing algorithm, fraction used f = 0.1, 0.2 and polynomial degree p = 2, (see method section on smoothing). Figure 1 d shows power spectral density of the time series.

Industrial production. The data for industrial production, IP, in Germany were retrieved from Statistisches Bundesamt. The publication lag for IP is about six weeks. (Hüfner and Schrøder 2002). IP is our target index for which we seek a leading index.

Recessions. OECD recorded recessions in Germany during the period 1991 to 2016. The dates designate the period from the peak through the trough. The data were obtained from the internet page: https://research.stlouisfed.org/fred2/series/DEUREC. The recession dates are shown in Table 1. In the table, we also estimate the seriousness of the recession by reporting the average deviation from the linearly detrended IP series 11 months around the trough of a recession period. A potential difficulty for the predictive power of leading indexes are structural breaks in the economy. However, Schrimpf and Wang (2010) found a structural break for Germany only in 1987, that is, before our study period begins. A second difficulty for predictions are a high volatility in the leading indexes. Caglayan and Xu (2016) show that this would occur for several leading indexes from about 2005 to 2012 in Germany and Camba-Mendez et al. (2001) suggest that volatile periods would require rich models including several leading indicators.

Figure 1 in here, data

Table 1 in here, recession periods

Survey based leading indexes. Each month about 7000 companies are asked by the ifo institute for business research about their current business situation (good, satisfactory, poor) and their expectations for their business for the next 6 months (favorable, unchanged, more unfavorable). The index is released the same month as the survey is taken. The ifo institute reports that the expectation index tends to lead industrial production with about two to three months. The interpretation is that if the ifo expectation gauge turns up, then odds are that it will be followed by an acceleration in factory output. (IFO 2016).

The ZEW business cycle expectations is also a survey based leading indicator in Germany. Each month about 300 analysts and financial experts of capital markets are asked about their expectations for the business cycle development in the next 6 months. (ZEW 2016).

The order flow data were obtained from Statistisches Bundesamt (Auftragseingangsindex). Order flow is assumed to be a leading index for GDP. It is published each month about 9 weeks after the data is collected. The index of order flow is discussed in Ozyildirim et al. (2010) p. 18. It is part of the OECD leading indicator, (OECD 2012) as well as the Conference Board's composite leading index (CLI), (Heij et al. 2011).

The unemployment index, UE, and was taken from Statistisches Bundesamt (Arbeitslosenquote). To characterize German economy, we used Monetary supply, M2 (Germany´s contribution to Euro basis), the consumer price index, CPI (Seasonally and calendar adjusted), Fibor-3 month (Frankfurt Interbank Offered Rate; monthly average), unemployment: % civilian labour, and the US ISM Purchasing managers index (PMI) for manufacturing. The ifo – index (business expectations) and US unemployment (inverted) is used as componernts in the Euro Area –wide leading indicator, ALI. (de Bondt and Hahn 2014)

3. METHOD

Accuracy is a measure of the degree that a positive / negative movement in IP follows a positive /negative movement in the leading index. Timing is calculated as the average time before a movement in the leading variable is reflected in a corresponding movement in IP. The timing is a function of the series' cycle length, CL, that ideally are identical for the leading index and its target, IP. With the nomenclature used here, a perfect leading index to IP has a value close to -1 and a perfectly lagging index to IP has a value close to +1. A leading lagging LL- index =-1 would mean that the leading index is leading through the whole time series, and technically, trajectories in the phase plot with IP on the x-axis and the candidate-leading index on the y-axis would always rotate clock-wise. Visually, the peak (trough) of the leading index will come before the peak (trough) of the target series, but less than ½ cycle length.

3.1 The running average leading – lagging method.

The method consists of 5 steps and is explained with reference to Figure 2 and follows closely the description given in Seip and Grøn (2016) and Seip and Grøn (2017). The first part of the method, step 2 below, has a counterpart in electrical engineering in the Lissajous curves, see e.g., https://en.wikipedia.org/wiki/Lissajous_curve. The second part, step 3 and Eq. 1, has a counterpart in the calculation of magnetic fields around a wire, e.g., https://en.wikipedia.org/wiki/Biot%E2%80%93Savart_law.

At the basis of the method is the dual representation of paired cyclic time series, x (t) and y (t), in time representation and as phase plots. As time series the x- axis represents time and the x(t) and y(t) variables are plotted on the y –axis. As phase plot, the paired time series are depicted on the x-axis and the y-axis on a 2D graph, Figure 2. If one series leads another series with less than ½ a cycle length (in some cases by contributing a causal effect on the other), then we will have persistent rotational direction of the series trajectories in the phase plot. Figure 2a and b give an example with x (t) = sin t and y (t) =sin (ωt + φ) = sin (t + 0.785). As an example, since it is well known that sun intensity

peaks before sea surface temperature, SST, on the western hemisphere, the first series, x, could represent SST, normally peaking in July – August and denoted by T in the graph.

Figure 2 in here, method

The second series, y, could represent sun insolation peaking in June and denoted by CC in the graph. Since Sun insolation is associated with heat transfer to sea surface, CC is a candidate cause for T. Thus, CC should peak before T, as it does in the figure. Real pairs of sun insolation and SST do the same (Seip 2015).

We explain the leading – lagging LL method in four steps.

Step. 1. Detrending and smoothing. Since the period we study is relatively short, we detrended the data by calculating the residuals after removing a linear regression against time. To remove noise we smoothed the data with the LOESS smoothing algorithm. We used four fractions of the series as running average periods: f = 0.02, 0.06, and 0.1 and f = 0, 2, and we always interpolated with a second order polynomial function, p = 2. The detrending and smoothing of the indexes are intended to mimic numerically the visual processes that are used in real life applications.

Step 2. Rotational directions in phase space. We then calculated the angles θ between two successive vectors $\mathbf{v}_1$ and $\mathbf{v}_2$ through 3 consecutive observation [4]s:

(1) $\qquad \theta = sign(\mathbf{v}_1 \times \mathbf{v}_2) \cdot \text{Arccos}\left(\frac{\mathbf{v}_1 \cdot \mathbf{v}_2}{|\mathbf{v}_1||\mathbf{v}_2|} \right).$

The rotational direction for the paired series in Figure 1c, upper part, is shown in the lower part as positive bars (counter clock-wise rotations) and negative bars (clock-wise rotations).

---

[4] It can be implemented in Excel format: With v_1 = (A1,A2,A3) and v_2 = (B1,B2,B3) in an Excel spread sheet, the angle is calculated by pasting the following Excel expression into C2: =SIGN((A2-A1)*(B3-B2)-(B2-B1)*(A3-A2))*ACOS(((A2-A1)*(A3-A2) + (B2-B1)*(B3-B2))/(SQRT((A2-A1)^2+(B2-B1)^2)*SQRT((A3-A2)^2+(B3-B2)^2))).

Step 2. The strength, LL - strength, of the mechanisms that cause two variables to either rotate clock-wise or counter clock-wise in a phase portrait is measured by the number of positive rotations minus the number of negative rotations, relative to the total number of rotations over a certain period.

(2) $LL = (N_{pos} - N_{neg})/(N_{pos} + N_{neg})$.

This means that we can assess the persistence of the rotational direction. We use the nomenclature: $LL(x, y) \in [-1, 1]$ for leading- lagging strength: $LL(x, y) < 0$ implies that y leads x, $y \rightarrow x$; $LL(x, y) > 0$ implies that x leads y, $x \rightarrow y$. In a range around $LL(x,y) = 0$ no LL- relations are significant.

Significance levels were calculated with Monte Carlo simulations for the LL-strength measure and to distinguished variables in principal component analysis, PCA, plots. We found the 95% confidence interval for the mean value (zero per definition) to be ± 0.32 for n = 9, that is, in a phase plot the series cycle persistently clock-wise or significantly counter clock-wise corresponding to significantly leading - lagging signatures for the series.

Principal component analysis, PCA, produces two major plots. The loading plot will in our study show similarities between economic states. The score plot will show how the variables that define the states relate to each other.   Significance levels for the PCA were identified by adding random numbers to the columns and the rows of the PCA matrix and then calculating confidence levels for the random numbers. Adding random numbers distorts the PCA algorithm to a certain degree, so the confidence estimates are only guiding values.

Step 3. The cycle length, CL, of two paired series that interact, can be approximated as:

(3) $CL = n \times 2\pi /(\sum_{2}^{n-1} \theta_{i-1,i,i+1})$.

$\theta_{i-1,i,i+1}$ is, as the angle between two consecutive vectors, that is, three consecutive observations. The number of angles that close a full circle corresponds to the cycle length

Step 4. The timing. TL. The regression slopes, s, or the β – coefficients, will for cyclic series give information on the shift, or time lag, between the series. For a linear regression applied to paired time

series that are normalized to unit standard deviation, the regression coefficient, r, and the β – coefficient (the slopes) will be identical. If the two series co-vary exactly, their regression coefficient will be 1, and the time lag zero. If they are displaced half a cycle length, the series are counter-cyclic, and the correlation coefficient is r = - 1. Lead or lag times, TL, are estimated from the correlation coefficient, r, for sequences of 5 observations, TL (5*)*. With λ as cycle length, an expression for the time lag between two cyclic series can be approximated by:

(4) $TL \approx \lambda/2 \times (\pi/2 - \text{Arcsine }(r))$

An expanded explanation of the method is given in Seip and Grøn (2017). The method is implemented in Excel and requires only the pasting of new datasets into two columns. The data set and all calculations are available from the authors.

For the whole period, we first calculate the LL- relation for 3 consecutive months and then calculate the running average LL- relations for 9 months. The LL- strength for the period 1991 to 2016 is the average LL- strength of the 308 observations calculated with Eq. 2.

Since many of the leading indicators aim at finding turning points in the economy, e.g., (Banerji et al. 2006; OECD 2012), we found the LL- strength for the periods before and a little into the recessions and the recoveries. We examined the leading relationship for 9 months, with 6 months before the recession peak and 3 months after the peak and correspondingly for the recovery trough.

3.2 Smoothing

Economic time series will normally be a superposition of several sub series that represent different mechanisms. Many series, like IP, will have a trend that are caused by factors that act over multidecadal time scales, there might be decadal effects associated with business cycles or growth cycle mechanisms, and there are noise. It is also quite likely that there also exists effects of dynamic chaos in the series. (Sugihara and May 1990; Tømte et al. 1998). Thus, although several authors, e.g., Florin et al. (2010) suggest that filtering should not be done without strong prior about the artefacts to be removed, smoothing the raw time series may be required both to remove possible artefacts and to

identify mechanisms that act on a certain time scale. To smooth the variables we use the LOESS lo-cally weighted smoothing algorithms. The algorithm is available in many statistical packages. We use SigmaPlot©. The smoothing algorithm has two variables. The first, f, shows how large the fraction of the series is used for calculating the running average. The second, *p*, is the order of the polynomial function used to make interpolations. To find a reasonable degree of smoothing, we used the time series 1994 to 2014. We summarized the result with principal component analysis, PCA.

3.3 Power spectral density

We apply a power spectral density algorithm (SigmaPlot©) to the single time series and compare the cycle lengths identified by this method to the common cycle lengths for paired series identified by the LL- method.

4. RESULTS

We discuss the results for different degrees of smoothing of the time series and thereafter compare the performance of the two survey based indexes, ifo and ZEW with the order flow index. Lastly, we examine the unemployment index that is generally considered to be a lagging index to GDP.

4.1 Smoothing macroeconomic series

We use smoothing to remove noise from the five series. The LL- strength of the series increase with the smoothing degree applied, Table 2. The raw series and series smoothed over 5 months gave very low LL- strength. However, smoothing over two years gave a reasonably good LL- strength and a cor-respondingly high probability for predicting correct movements of IP. We used f = 0.1 p = 2 to investi-gate running average leading properties of the four indexes.

Table 2, results

11

The detrended and smoothed series for IP with LOESS parameters f = 0.1, p = 2 shown in Figure 1b

show peaks : 1995:6; 1998:4; 2000:11; 2002:12; 2004:7; 2008:3; 2011:7 and 2014:4 with cycle times

in months: 39, 32, 24, 19, 44, 40 and 35 giving an average cycle time of 33 months.

4.2   Leading – lagging relations.

Leading and lagging relationships for the two sentiment based indexes, ifo and ZEW are shown in Fig-

ure 3 and for the two behavioral based indexes in Figure 4. The first row of the figures shows the

paired time series, detrended and normalized to unit standard deviation. In this graph it is possible to

identify visually the LL- relations between the series. The second row shows the leading lagging-

strength (shaded bars in the range -1 to +1) as a function of time for the ifo – index (left panel) and

the ZEW – index (right panel). Dashed lines show confidence limits for LL- strength. The black bars

show the angles, θ. Negative angles represent clock-wise rotations in the phase plots and a leading

role for the candidate leading indexes. The block line at the bottom of each panel shows the reces-

sion periods for the German economy defined by OECD. Lower values show stronger recessions. The

lower row figures show estimated cycle times and estimated phase shift for the two indexes ifo or

ZEW. The phase shift represent the leading time if the index is leading IP. The LL- algorithm identifies

common cycle times of about 2 years, which can also be seen in the time series graphs in the upper

row. This is a little less than the average cycle time for the IP series. The leading times are 5 to 7

months. Corresponding graphs to those in Figure 3 is shown for OF and UE in Figure 4.

Figure 3 in here, ifo and ZEW index

Figure 4 in here OF and UE index

The LL- strength pattern in Figure 3 and in Table 3 shows that the ifo index has the longest leading

period, the ZEW index comes on a second place, the UE – index on a third place and the OF - index

last. There are two time windows where in particular the sentiment-based indexes do not perform

well. A period around 1997 and a period from about 2006 to 2007. The latter time window is some-what before the 2008 recession in Europe.

The index for unemployment is normally regarded as a lagging index for industrial production, (Enders 2010; Ball et al. 2015). However, with moderate smoothing it came out as a leading index, Figure 4.

A summary of the results in Figures 3 and 4 is shown numerically in Table 3. The results for the average recession and recovery periods show that recession periods are generally better predicted than the following recovery periods (-0.70 versus -0.32) The ifo-index predicted overall best for the whole period, but the ZEW index was better before recession and recovery periods. The UE –index was surprisingly good in predicting the combined industrial production for manufacturing and construction, IP(M+C).

We also compared the ZEW index to the ifo index and to UE. The ZEW index was a leading index to the ifo index in 83 % of the time. ZEW was largely a lagging index to UE during the period 1991 to 2007, but became a leading index to UE after 2008, Figure 5.

Figure 5 in here, ZEW and IFO; ZEW and UE

Figure 6 in here (German economy 1900 to 2016

4.3 Relations among detrended time series

The results for leading- lagging relations can be summariced in a principal component plot. However, for cyclic series the interpretation is different than for time series in general. Figure 1e showed a loading plot for 10 sine –functions that are shifted fractions ½ to 1/16 of a cycle length relative to each other. A sine function that is shifted ¼ of a cycle length (ϕ = ¼ CL) relative to a reference sine function (ϕ = 0) will in a phase plot show a perfect circle, an ordinary linear regression will show an

explained variance, $r^2 = 0$ and a probability, $p = 0$. A PCA loading plot for the detrended time series are shown in Figure 1f. It shows that the two components of industrial production: manufacturing and construction are closely associated (IPMC is close to IPM in the figure) . As anticipated, the ifo indexes are associated with IP, and the association is smaller for ifo – expected, IFOE, than for ifo – current, IFOC. The ZEW index and UE are both negatively associated with IP, that is, they are counter cyclic, suggesting that they lead IP with more than ¼ λ.

There are in particular two time domains where the leading indexes fail. Figure 6 shows how economic states in Germany 1995 to 2016 are connected. (The numbers identify the years two last digits, that is, "9" is 2009.) The years 1995 to 1998 forms an "island" with high unemployment UE, low Bundesbank rate, FF and low industrial production, IP. The years around 2005 has low monetary supply, M2 and scores low on the PMI index. It is also a year where Capital control restrictions on output growth rate increased considerably, Chakraborty et al. (2016 Figure E5) and Fernandez et al. (2016).

5. DISCUSSION

We first discuss the numerical results for the three candidate leading indexes and compare their performance. Thereafter we discuss how the leading indexes should be used with respect to smoothing of the series and LL-strength before reliable predictions can be made. We then discuss the accuracy and timing of the indexes. The candidate lagging index, UE, turned out to be a leading index for 73 % of the time.

5.1 Comparing ifo and ZEW to the behavioral based index, OF.

In agreement with our *first* hypothesis, the sentiment-based indexes gave the best predictions, followed by the unemployment index and the OF index. The ifo – index, based on industry management opinions, performed best of the two sentiment based indexes, but closely, and not significantly different came the ZEW – index based on the opinions of financial experts. (-0.596, -0.583, -0.564 and -

0.19 respectively, -1.0 being the best performance and +1 the worst performance. ("+1" means that the series would behave as a perfectly lagging index.) The sequence for the LL- strength of the recession and the recovery periods were a little different, with the ZEW index best for recessions, and UE best for recoveries. On the average for all indexes, recession periods were predicted better than the recovery periods (- 0.70 and - 0. 32 respectively, Table 3). Hüfner and Schröder (2002) compared the ifo and the ZEW index and found the ZEW index to provide better forecasts for the period 1994M1 to 2002M3. However, they showed that the ifo – index was better than the ZEW index for the shorter period 1998M1 to 2002M3, thus making the comparison strongly dependent on the selected period. Since the LL- method shows running performance of LL-strength, we can compare this statement to the results shown in Figure 3. There seems to be no reason why the ifo index should become better than the ZEW –index by removing the four years 1991 to 1998 from the test set. This result illustrates the advantage of the running LL- strength method compared to methods that require stationary time series.

The ZEW index was also included in a test of 8 leading indexes for the Euro area 1992M12 to 1999M12 by Carstensen et al. (2011), but came out as #2 to #6 of 8 indicators in a series of tests. Our results are in line with results by Christiansen et al. (2014) on the role of sentiment-indicators. They found that the consumer sentiment index (their pseudo R^2 = 0.26; based on 500 households) as well as the Supply Management's Purchasing Manager's index (pseudo R^2 = 0.47; 400 industrial companies and 20 manufacturing companies) 1975- 2013 were superior to 3 classical recession predictors, e.g., the term spread, federal funds rate and stock market returns. Angelini et al. (2011) found that sentiment based (soft) indexes were better that hard indexes for longer time horizons. The unemployment index, that is supposed to be a lagging index, showed an overall leading index signature in our study, i.e., 73 % of the times. However, the German unemployment index have a different relation to the output than in many other countries, Tang and Bethencourt (2017) and the β - coefficient in Okun's law is much smaller than in for example US, Ball et al. (2015). Forni et al. (2001) did not include UE in their core set of LL- indicators for Germany, but found employment to be a significant

lagging indicator. However, in other studies UE (non-agriculture) is termed a coinciding index (Heij et al. 2011). Thus, it appears that the result for UE are characteristic for the economy studied, and may give important information for employment policies.

Periods: Recession, recoveries, index volatilities. The recovery in 1997 and the recession in 2011 was most difficult to predict, whereas the 2008 recession was predicted well by all indexes. The good prediction of the 2008 recession may be due to warning signals from the US economy that showed a peak in December 2007 and a trough in June 2009. On average, recessions would be predicted best, the overall economic growth next best and recoveries worst. This result contrasts with our *second* hypothesis that recessions and recoveries would be predicted less well than movements under normal economies. Caglayan and Xu (2016) suggest that high volatility in the indexes may affect stock returns, but that high volatility do not translate into worse than average predictions of IP.

Time windows with anomaly predictions. In the present context, an anomalous prediction means that the candidate-leading index appear as a lagging index or there is no significant LL- signature. The leading index will appear as a lagging index if it is further than ½ cycle length from the peak in IP following it, and closer than ½ cycle length to the peak in IP preceding it. We found two time windows where the two indexes failed. The first was around 1997 (ifo: 1996M8 - 1997M1; ZEW: 1997M3 - 1997M12; OF: 1995M5-1997M2; UE: 1996M8-1997M2) and the second around 2005 (ifo: 2006M5 - 2007M1; ZEW: 2006M5-2007M3; OF: 2005M4-2007M6, UE: 2006M12-2007M7). We have no definitive suggestions why these two periods turned out to be difficult to predict correctly, however, 1997 designated the end of "the great moderation" in US, McNown and Seip (2011) and it was at the end of a minor depression, that appears as an "island" in German economy, Figure 6. Around 2005 oil prices increased from $10 to $60 a barrel, there were few orders, monetary supply was low, and capital control measures increased in Europe. From Figures 3 it is seen that the upturns probably were predicted too early in 1997 and that the downturns were predicted too early in 2005.

When we replaced IP- growth as target variable instead of IP, the results were inferior to using IP (results not shown) and is opposite to our *last* hypothesis, that IP –growth would be easier to predict. However, the calculation of growth rates most often increases the noise to signal ratio. (Seip and McNown 2007).

Cycle times were 20 to 30 months, that is, around 2 years. This time is a little less than the first peak in the power density functions shown in Figure 1d. It corresponds with the cycle times that can be identified visually from the smoothed IP series shown in Figure 1b. The cycle times identified in this study are shorter than the normal estimates of business cycle times that often are set to between 2 and 8 years (Zarnowitz and Ozyildirim 2006). However, since the time series are linearly detrended, the cycles are more characteristic for IP – growth cycles than for business cycles.

The average lead times for the indexes were 4.7 to 7.5 months, but varying over time. The lead time compares well with the lead times reported for Euro Area-wide leading indicator showing 7 (0-21) months for peaks and 6 (2-24) months for troughs.(de Bondt and Hahn 2014). The unemployment index had the longest leading time, giving the observers the longest warning time for changes in the business cycle and the longest period for assessing, and smoothing, the index. However, ifo index gave the best prediction, but only 0.7 months after predictions could be made with the UE index.

5.2   Smoothing and outlier removals

We have shown that smoothing the indexes is required for using the indexes as predictors. This supports our *third* hypothesis, that smoothing of the indexes as well as the target series is a requirement for giving good predictions. The degree of smoothing depends on how much noise it is in the data. However, there is a tradeoff between the number of observations required for smoothing, and the available time for making predictions , e.g., discussions in Ozyildirim et al. (2010). In contrast to for example Camba-Mendez et al. (2001), that used an intervention model to a priori filter out particular anomalous events. With the LL- method, such events are detected by the LL- algorithm if they cause a change LL- relations. The LL- method is a rolling window approach. Alternative detrending and

smoothing algorithms are most often global, e.g., using low or high pas filters like the Hodrich-Prescot filter.

The two institutes that publish the two leading survey-based indexes construct their indexes so that they have a lead-time of 6 months. This fits well with the prediction horizon found here.

Prediction skill for linear autoregressive forecasting models (AR) as well as non-linear forecasting algorithms can be enhanced by identifying and removing outliers. The LL –method offer one way to do that. The phase plot for paired time series that include a target variable, e.g., IP, should ideally look like figure 2b. However, in practice there will be observations that deviate from a regular elliptic form. Although generated by two random series, the points 5 and 9 in Figure 2c and d may appear as outliers and could in principle be removed or replaced to comply better with a smooth rotation (the example just illustrates the technique, it is not meaningful to apply it to paired random series.) Improving prediction skills by removing outliers will be tested in a future study, but is outside the scope of the present study.

In actual applications the leading series are smoothed over a time window t_0 to t, and then predictions are made for times t+1, t+2.. . Our LL –algorithm identifies running average phase shifts between the leading index and the target index. This would make it easier to identify how many steps ahead a forecast is made. Our results suggest that smoothing over 4 months (f = 0.02) is too short, whereas 20 months (f = 0.1) may be too long for practical purposes. Our smoothing algorithm is "automatic" in that only the smoothing window has to be determined. Other smoothing algorithms may be available that allows shorter time windows for smoothing.

It appears that the performance of leading indexes depends upon the economy studied. The LL-method supplies a tool for continuous monitoring of decline and improvement in candidate leading variables, although there appears to be incidents where none of the present methods work (e.g., the steep changes in the ifo index in 1997 and 2005).

Principal component analysis arrange cyclic series approximately according to the phase shift be-tween them, Figure 1f. The variables studied here are imperfectly cyclic, still their position to each other in the PCA loading plot in Figure 1e show relations that are consistent with assumptions about their leading lagging relation. For example, the ifo expectation series, IFOE, are shifted a larger dis-tance from the IP series than the ifo current series, IFOC.

6. CONCLUSION

We compare two survey-based and one behavioral-based leading indexes to industrial production, IP, for the period 1991 to 2016 in Germany. We find that - with appropriate smoothing of the indexes and IP - the sentiment based ifo index based on surveys of 7000 business managers gives the best predictions. However, the ZEW- index based on surveys among 300 financial experts is very close both in prediction strength and in timing. The behavioral based OF – index is worst, but surprisingly, the UE index is quite good. Prediction skills for recession periods were better than the overall predic-tion skill, whereas prediction skill for recoveries were worse. Using the indexes requires more than 4 months period to smooth both the indexes and the time series for IP. We found that there are time windows where all leading indexes failed and that these periods coincided with abnormal periods in German economy. However, the periods where the leading indexes failed my give support for im-provement in the prediction methods. We believe that the rolling leading –lagging method de-scribed here will give rapid and accurate description of candidate indexes for the construction of leading – indicators.

Literature

Abel, A., B. Bernanke and R. McNabb (1998). Macroeconomics. Harlow, Addison- Wesley.

Angelini, E., G. Camba-Mendez, D. Giannone, L. Reichlin and G. Runstler (2011). "Short-term forecasts of euro area GDP growth." Econometrics Journal **14**(1): C25-C44.

Balcilar, M., R. Gupta, A. Majumdar and S. M. Miller (2013). "Forecasting Nevada gross gaming revenue and taxable sales using coincident and leading employment indexes." Empirical Economics **44**(2): 387-417.

Ball, L., J. T. Jalles and P. Loungani (2015). "Do forecasters believe in Okun's Law? An assessment of unemployment and output forecasts." International Journal of Forecasting **31**(1): 176-184.

Banerji, A., P. Dua and S. M. Miller (2006). "Performance evaluation of the new Connecticut leading employment index using lead profiles and BVAR models." Journal of Forecasting **25**(6): 415-437.

Caglayan, M. and B. Xu (2016). "Sentiment volatility and bank lending behavior." International Review of Financial Analysis **45**: 107-120.

Camba-Mendez, G., G. Kapetanios, R. J. Smith and M. R. Weale (2001). "An autmoatic leading indicator of economic activity: forecasting growth for European countries." Econometrics journal **4**: S56-S90.

Carstensen, K., K. Wohlrabe and C. Ziegler (2011). "Predictive Ability of Business Cycle Indicators under Test A Case Study for the Euro Area Industrial Production." Jahrbucher Fur Nationalokonomie Und Statistik **231**(1): 82-106.

Chakraborty, I., R. Hai, H. Holter and S. Stepanchuk (2016). A mulit-country equilibrium analysis of Europe. Carnegie-Rochhester-NYU Conference on public policy.

Christiansen, C., J. N. Eriksen and S. V. Moller (2014). "Forecasting US recessions: The role of sentiment." Journal of Banking & Finance **49**: 459-468.

de Bondt, G. J. and E. Hahn (2014). "Introducing the Euro Area-wide Leading Indicator (ALI): Real-Time Signals of Turning Points in the Growth Cycle from 2007 to 2011." Journal of Forecasting **33**(1): 47-68.

Enders, W. (2010). Applied econometric time series, 3 edition, John Wiley & sons.

Fernandez, A., M. W. Klein, A. Rebucci, M. Schindler and M. Uribe (2016). "Capital Control Measures: A New Dataset." Imf Economic Review **64**(3): 548-574.

Ferrara, L., M. Marcellino and M. Mogliani (2015). "Macroeconomic forecasting during the Great Recession: The return of non-linearity?" International Journal of Forecasting **31**(3): 664-679.

Florin, E., J. Gross, J. Pfeifer, G. R. Fink and L. Timmermann (2010). "The effect of filtering on Granger causality based multivariate causality measures." Neuroimage **50**(2): 577-588.

Forni, M., M. Hallin, M. Lippi and L. Reichlin (2001). "Coincident and leading indicators for the EURO area." Economic Journal **111**(471): C62-C85.

Heij, C., D. van Dijk and P. J. F. Groenen (2011). "Real-time macroeconomic forecasting with leading indicators: An empirical comparison." International Journal of Forecasting **27**(2): 466-481.

Hüfner, F. P. and M. Schrøder (2002). Forecasting economic activity in Germany - How useful are sentiment indicators? Mannheim, Germany, Zentrum für Europeische Wirtschaftsforschung GmbH**: 14.

Hüfner, F. P. and M. Schröder (2002). Forecasting Economic Activity in Germany - How Useful are Sentiment Indicators? (September 15, 2002). Available at SSRN: , ZEW institute, Centre for European economic research.

IFO. (2016). "German IFO business survey."   Retrieved January 7, 2016, from https://books.google.no/books?id=wj79gd3m8cUC&pg=PA330&lpg=PA330&dq=ifo+institute+for+economic+research+ifo+index&source=bl&ots=td1PeV-QHz&sig=qqaaphAxdC_DWwMZh39R3O4uPaY&hl=no&sa=X&ved=0ahUKEwjortKk2JfKAhXh73IKHZkdBoIQ6AEITzAH#v=onepage&q=ifo%20institute%20for%20economic%20research%20ifo%20index&f=false.

Levanon, G., J. C. Manini, A. Ozyildirim, B. Schaitkin and J. Tanchua (2015). "Using financial indicators to predict turning points in the business cycle: The case of the leading economic index for the United States." International Journal of Forecasting **31**(2): 426-445.

McNown, R. and K. L. Seip (2011). "Periods and structural breaks in US economic history 1959-2007: A data driven identification." J. Policy Modelling **33**: 169-182.

OECD (2012). OECD system of composite leading indicators.

OECD. (2016). "OECD based Recession Indicators for Germany."   Retrieved January 21 2016, from https://research.stlouisfed.org/fred2/series/DEUREC.

Ozyildirim, A., B. Schaitkin and V. Zarnowitz (2010). "Business Cycles in the Euro Area Defined with Coincident Economic Indicators and Predicted with Leading Economic Indicators." Journal of Forecasting **29**(1-2): 6-28.

Schrimpf, A. and Q. W. Wang (2010). "A reappraisal of the leading indicator properties of the yield curve under structural instability." International Journal of Forecasting **26**(4): 836-857.

Seip, K. L. (2015). "Investigating possible causal relations among physical, chemical and biological variables across regions in the Gulf of Maine." Hydrobiologia **744**: 127-143.

Seip, K. L. and O. Grøn (2016). "Leading the Game, Losing the Competition: Identifying Leaders and Followers in a Repeated Game." Plos One **11**(3).

Seip, K. L. and Ø. Grøn (2017). "A new method for identifying possible causal relationships between CO_2, total solar irradiance and global temperature change." Theoretical and Applied Climatology **127**: 923-938.

Seip, K. L. and R. McNown (2007). "The timing and accuracy of leading and lagging business cycle indicators: a new approach." International journal of forecasting **22**: 277-287.

Sugihara, G. and R. M. May (1990). "Nonlinear forecasting as a way of distinguishing chaos from measurement errors in time series." Nature **344**: 731-741.

Tang, B. and C. Bethencourt (2017). "Asymmetric unemployment-output tradeoff in the Eurozone." Journal of Policy Modeling **39**(3): 461-481.

Tømte, O., K. L. Seip and N. Christophersen (1998). "Evidence that loss in predictability (and possibly dynamic chaos) increase with increasing trophic level in aquatic ecosystems." Oikos **82**: 325-332.

Yilmaz, Y. (2016). Die bessere Glaskugel -Vor- und Nachteile von Früh - vs. Sp¨tindkatoren zur Prognose von Konjunkturentwicklungen in Deutchland" Ein literaturbasiertr Vergleich. Bachelor, Ruhr-Universitet Bochum.

Zarnowitz, V. and A. Ozyildirim (2006). "Time series decomposition and measurement of business cycles, trends and growth cycles." Journal of Monetary Economics **53**(7): 1717-1739.

ZEW. (2016). "ZEW Economic Sentiment Index ", from http://www.tradingeconomics.com/germany/zew-economic-sentiment-index.

Table 1 OECD based Recession Indicators for Germany from the period following the Peak through the Trough. Rescission seriousness (OECD 2016)

Start Period /peak	1995:7	2001:6	2008:3	2011:6	2014:3
End period/Trough	1997:3	2005:4	2009:5	2013:3	-
Recessen seriousness	- 1,69	-0,76	-13,4	-1,77	-1-,05

Table 2 Effects of smoothing the leading indexes and IP. f Is fraction of series used as rolling average. We always interpolate with a second order polynomial function, p = 2.

	Accuracy 1994-2014, Best = -1, worst = +1 (lagging); Smoothing				
Index	Raw	f = 0.02, n = 5 months	f = 0.06, n = 15 months	f = 0.1, n = 25 months	f = 0.2, n = 50 months
ifo	-0.136	-0,136	-0,184	-0,504	-0.704
ZEW	-0.056	-0,048	-0,360	-0.608	-0.784
OF	-0.12	-0,096	-0,464	-0.720	-0.816
UE	-	-	-	-0.08	-

Table 3. Leading – lagging strength for IP(M+C) versus . Characteristics for the whole period 1991 - 2016, for the average of the 6 recession periods and for the average of the 5 recovery periods. Results with LOESS smoothing f = 0.1, p = 2. Number in parentheses are lags found by Hüfner and Schrøder (2002) See text

	Index	LL- strength			Leading time		
1		1991-2016	reces-sion	recov-ery	Cycle time, months	Timing, 1994 - 2014 months	Percentage significant
2	ifo (managers)	-0,596	-0,778	-0,156	33,6	7,3 (2)	78
3	ZEW (financial experts)	-0,583	-0,867	-0,289	27.3	6.8 (5)	77
4	Unemployment (behavioral)	-0,564	-0,733	-0,556	32,3	8,0	73
5	Order flow, OF (behavioral)	-0,186	-0,422	-0,289	29.7	5.7	46
	Average	-0.48	-0.70	-0.32	30.73	6.95	68.5

Figure 1.

The data: Industrial production, IP; Leading index, ifo, leading index, ZEW, order flow, OF. a) Raw data, b) data linearly detrended and smoothed with LOESS smoothing algorithm, f = 0.1, p = 2. c) Data linearly detrended and smoothed with LOESS smoothing algorithm, f = 0.2, p = 2. d) Power spectral density for IP and the three leading indexes. The arrow shows peaks around 33 months. e) Principal component plots, PCA. Loading plots. Time series linearly detrended, f) Principal component plot for 10 sine functions that is shifted relative to each other. Numbers indicate the fractions of a cycle length that the sine is shifted. Note that when a sine is shifted ¼ cycle length, the trajectories in phase space is a circle, and the explained variance, $r^2 = 0$ and p = 1. IPMC = industrial production including manufacturing and constructions, IPM = Industrial production, only manufacturing, IFOC = ifo index current economy, IFOE = ifo index expectations; IFOCL = ifo –index climate; ZEW = ZEW index expectations. OF = Order flow, UE = unemployment,

Figure 2

Time series (left) and phase plots (right). a) Two sine functions: CC is candidate cause and T is target. The candidate cause, CC, peaks before the target, T. b) In a phase plot with T on the x- axis and CC on the y-axis the time series rotates clock-wise (negative by definition), θ is the angle between two consecutive trajectories. Note that with two time series normalized to unit standard deviation, phase plot for the two series will show an ellipse with the long axis in the 1:1 or 1:-1 direction. The phase shift between the series is a function of the ratio between the long and the short axes. c) Upper part: time series based on random numbers drawn from a uniform distribution; lower part: running angles. d) Phase plot for the time series in c. Points on the trajectories are numbered consecutively. Notice that the first angle 0-1-2 is positive (rotate counter clock-wise). Figure redrawn after Seip and Grøn (2016).

Figure 3

Leading- lagging relations between two leading, LL- indexes and industrial production, IP (manufacturing and construction). The paired variables are LOESS smoothed with time window f = 0.1 (25 months) and polynomial degree p = 2. a) Time series for IP(M+C) and ifo –index for expectations; b) Time series for IP(M+C) and ZEW –index for expectations. c) LL- relation for IP vs. ifo (shaded area, average of 9 consecutive observations), confidence limits (dashed lines) and angles (black bars, average of 3 consecutive observations) Full broken line shows stylized OECD recession periods, depth indicate recession seriousness. d) LL- relation for IP vs. ifo other curves as in c; e) Running average common cycle times and phase shift for IP vs. ifo- index. f) Running average common cycle times and phase shift for IP vs. ZEW- index.

Figure 4

Leading- lagging relations between two leading, LL- indexes and industrial production, IP (manufacturing and construction). a) Time series for IP(M+C) and order flow, OF; b) Time series for IP(M+C) and unemployment, UE. c) LL- relation for IP vs. OF (shaded area, average of 9 consecutive observations), confidence limits (dashed lines) and angles (black bars, average of 3 consecutive observations)

Full broken line shows stylized OECD recession periods, depth indicate recession seriousness. d) LL-relation for IP vs. UE other curves as in c; e) Running average common cycle times and phase shift for IP vs. OF. f) Running average common cycle times and phase shift for IP vs. UE.

Figure 5

Leading – lagging relations between the ZEW – index and potential leading indexes. a) detrended time series for ZEW and the ifo –index (R6, expectations) b) Detrended time series for ZEW and un-employment, UE. C) LL-relation for ZEW vs. ifo; d) LL- relations between ZEW and UE. Dashed line shows recession periods.

Figure 6.

 Principal component plots for German economy 1900 to 2016. a) Score plot showing time sequence for the economy. Only two last digits in the years are shown. b) Loading plot showing the position of economic variables. FF   federal funds rate, IP = industrial production, M" = monetary supply, CPI = consumer price index, UE = unemployment, PMI = purchasers managers index.

Figure 1

a)

Industrial production, IP,
leading indexes, IFO, ZEW and order flow, OF;
unemployment, UE

b)

IP, IFO, ZEW and OF, detrended
smoothed (LOESS; 0.1,2), normalizede

c)

IP, IFO, ZEW and OF, detrended,
smoothed (LOESS 0.2;2), normalized

d)

Log(Power spectral density) for IP, IFO, ZEW and OF

e)

Principal component loading plot for
shifted sine functions

f)

Time series, detrended
and LOESS smoothed (f = 0.1, p = 2)

27

Figure 2

a)

Candidate cause +0,785 and target

b)

Phase plot: Candidate cause + 0.785 and tar

c)

Random, uniformely distributed time series
positive = counter clock-wise rotations
negative = clock-wise rotations

d)

Random series as phase plot

28

Figure 3

a)

Industrial production (M+C) and IFO expectations

b)

Industrial production (M+C) and ZEW expectations

c)

Pos. bars: IP(M+C) leads IFO(R6 expectations)
Neg. bars: IP(M+C) lags IFO(R6 expectations)

d)

Pos. values: IP leads ZEW
Neg. values: ZEW leads IP

e)

IP versus IFO (expectations),
cycle length and phase shift

f)

IP vs. ZEW(expectations)
cycle length and phase shift

Figure 4

a)

Industrial production (M+C) and Order flow, OF

b)

Industrial production (M+C) and Unemployment, UE

c)

Pos. values IP leads OF
Neg. values: OF leads IP

d)

Pos. bars: IP leads Unemployment, UE
Neg. bars: UE leads IP

e)

IP vs. OF
Cycle length and phase shift

f)

IP vs. Unemployment, UE
cycle length and phase shift

Figure 5

a)

ZEW and IFO - indexes (expectation)

b)

ZEW and unemployment, UE

c)

Pos. bars: ZEW leads IFO
Neg. bars ZEW lags IFO

d)

Pos. bars: ZEW leads UE
Neg. bars: ZEW lags UE

Figure 6

Score plot, German economy 1900 to 2016

Loading plot for German economy 1900 to 2016

31